Best Available Copy
The Research Laboratory of Electronics Progress Report No. 136 January 1-December 31, 1993

Jonathan Allen (principal investigator)

Massachusetts Institute
77 Massachusetts Avenue
Cambridge, MA 02139

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

Approved for public release; distribution unlimited.

Security Classification of Report: UNCLASSIFIED
Security Classification of This Page: UNCLASSIFIED
Security Classification of Abstract: UNCLASSIFIED

DTIC ELECTED JUL 12 1994

SBD

UL
RLE Progress Report Number 136

Cover and title page:

This sea ice was synthesized from polarimetric multifrequency data. The data was collected by synthetic aperture radar (SAR) installed on a NASA/JPL DC-8 flying over the Beaufort Sea area in the Arctic Ocean northeast of Alaska. See part 2, section 3, chapter 1 for a further description of this research which is being carried out in RLE's Electromagnetics Group under the direction of Professor Jin Au Kong. For more information about electromagnetics and the work of the Electromagnetics Group, see RLE currents (volume 7, number 1, Fall 1993), which is available upon request from the RLE Communications Office.

Our special thanks to the following staff members of the RLE Communications Group: Mary J. Ziegler for her exceptional editing and scanning; Mary S. Greene for formatting, proofreading, and preparation of the publications and personnel chapters; Rita C. McKinnon for proofreading, and Michelle M. Scheer for formatting. We also want to thank David W. Foss, Manager of the RLE Computer Facility, for his time and invaluable technical assistance.

We thank the faculty, staff, and students of RLE for their generous cooperation.

Editor: Barbara Passero
Associate Editor: Mary J. Ziegler
Design and photography: John F. Cook
Printer: ZBR Publications, Wilmington, Massachusetts
Typesetting: This report was produced with IBM's BookMaster Software.

Copyright © 1994 by the Massachusetts Institute of Technology. All rights reserved. ISSN 0163-9218
# Table of Contents

## Introduction

1

## PART I SOLID STATE PHYSICS, ELECTRONICS AND OPTICS

### Section 1 Materials and Fabrication

#### Chapter 1 Heterostructures for High Performance Devices

*Professor Clifton J. Fonstad, Jr.*

- 1.1 Introduction .......................................................... 7
- 1.2 Growth Optimization of MBE-Grown InAlAs on InP ........................................ 7
- 1.3 Fabrication of Ridge Waveguide Distributed Feedback Lasers by X-ray Lithography .................................................. 8
- 1.4 Numerical Calculation of Coupling Coefficients in Ridge Waveguide Distributed Feedback Lasers ......................................................... 9
- 1.5 Measurement of Excited-state Lifetimes in Narrow Quantum Wells .................... 11
- 1.6 Tunable Semiconductor Lasers ........................................................................ 11
- 1.7 Integration of Vertically-emitting, In-plane Cavity Laser Diodes on GaAs VLSI Circuitry .......................................................... 13
- 1.8 Thermal Stability of GaAs MESFET VLSI Circuits ............................................ 14
- 1.9 Gas Source MBE of InGaAsP Laser Diodes on GaAs Substrates ....................... 15
- 1.10 High-Density OEIC Neural Systems Produced by Monolithic Integration of GaAlAs Light Emitting Diodes on GaAs MESFET VLSI Circuits ......................................................... 15
- 1.11 Surface-Normal Optical Input and Output Cells for High-Density, High-Speed GaAs MESFET-based OEICs ........................................... 18
- 1.12 Fiber-coupled GaAs MESFET-based OEICs ....................................................... 19
- 1.13 Low-temperature, Selective-area MBE Growth of GaAlInAs Laser Diodes and Optical Waveguides on Semi-insulating GaAs Substrates ......................................................... 20
- 1.14 Applications of Resonant Tunneling Diodes in GaAs MESFET VLSI .................. 21
- 1.15 Polarization-resolved Infrared Spectra of Very Narrow AlAs/InGaAs/InP Quantum Wells .......................................................... 22
- 1.16 Symmetry Properties of Quantum Well Subband Energy Levels and Selection Rules for Intersubband Transitions ........................................... 24
- 1.17 Investigation of Infrared Intersubband Emission from InGaAs/AlAs/InP Quantum Well Heterostructures ......................................................... 25
- 1.18 High-Frequency/High-Speed Characterization, Analysis and Modeling of Heterojunction Bipolar Transistors, Laser Diodes, and m-s-m Photodetectors ......................................................... 26
- 1.19 Damage-Free In-Situ UHV Etching and Cleaning of III-V Heterostructures Using Molecular Beams .......................................................... 26

#### Chapter 2 Physics of InAlAs/InGaAs Heterostructure Field-Effect Transistors

*Professor Jesús A. del Alamo*

- 2.1 Introduction .......................................................... 29
- 2.2 Physics of Breakdown in InAlAs/InGaAs MODFETs ........................................... 30

* Sponsored by the Joint Services Electronics Program.
Table of Contents

*2.3 Impact Ionization in InAlAs/InGaAs HFETs ......................................................... 33
*2.4 Publications and Conference Papers ..................................................................... 34

Chapter 3 Gas Source Molecular Beam Epitaxy of Compound Semiconductors ................. 35
Professor Leslie A. Kolodziejski

*3.1 Introduction ........................................................................................................... 35
*3.2 Gas Source Molecular Beam Epitaxy of ZnSe, ZnSe:Cl and ZnSe:N ......................... 37
*3.3 Novel Epitaxial III-V Buffer Layers for Wide Bandgap II-VI Visible Sources .......... 42
*3.4 Integrated Photonic Devices: The Channel-Dropping Filter ..................................... 45
*3.5 Heterovalent Interfaces Composed of II-VI/III-V Heterostructures ......................... 46
*3.6 Optoelectronic Very Large Scale Integrated Circuits ............................................. 48
*3.7 Heteroepitaxy of GaAs onto Corrugated Surfaces of Si ........................................ 48

Section 2 Quantum-Effect Devices

Chapter 1 Statistical Mechanics of Quantum Dots ......................................................... 51
Professor Boris L. Altshuler

*1.1 Project Description ............................................................................................... 51
*1.2 Publications .......................................................................................................... 52

Chapter 2 Single-Electron Transistors ............................................................................ 53
Professor Marc A. Kastner

*2.1 Project Description ............................................................................................... 53
*2.2 Publications .......................................................................................................... 56

Chapter 3 Edge Structure of a Quantum Dot In a Large Magnetic Field ......................... 57
Professor Patrick A. Lee

*3.1 Project Description ............................................................................................... 57
*3.2 Publications .......................................................................................................... 58

Chapter 4 Nanostructures Technology, Research, and Applications .............................. 61
Professor Henry L Smith

4.1 NanoStructures Laboratory .................................................................................... 61
*4.2 Scanning Electron-Beam Lithography Facility ...................................................... 61
*4.3 Spatial-Phase-Locked Electron-Beam Lithography ................................................ 62
*4.4 X-Ray Nanolithography ....................................................................................... 64
*4.5 Improved Mask Technology for X-Ray Lithography ........................................... 65
*4.6 A High Precision Mask Alignment System ......................................................... 66
*4.7 Optimization of Synchrotron-Based X-ray Lithography ....................................... 69
*4.8 Achromatic Holographic Lithography .................................................................. 70
*4.9 Ion Beam Lithography ........................................................................................ 70
*4.10 High Performance Self-aligned Sub-100 nm MOSFETs Using X-ray Lithography ... 72
*4.11 Fabrication of T-gate Devices Using X-ray Lithography ..................................... 73
*4.12 Studies of Coulomb Charging Effects and Tunneling in Semiconductor Nanostructures ............................................................. 76
4.13 Dual Electron Waveguide Device Fabricated Using X-ray Lithography ................... 79
*4.14 Novel Mesoscopic Superconducting Devices ..................................................... 79
Table of Contents

• 4.15 Channel-Dropping Filters Fabricated Using X-ray Lithography .......................................................... 79
• 4.16 Ridge-Grating Distributed-Feedback Lasers Fabricated by X-ray Lithography .............................. 83
  4.17 Fabrication of Sub-micron MSM Photodiodes by X-ray Lithography ........................................ 85
• 4.18 Submicrometer-Period Transmission Gratings for X-ray and Atom-Beam Spectroscopy and Interferometry ......................................................... 86
• 4.19 High-Dispersion, High Efficiency Transmission Gratings for Astrophysical X-ray Spectroscopy ........ 86
• 4.20 GaAs Epitaxy on Sawtooth-Patterned Silicon .................................................................................. 88
  4.21 Publications ................................................................................................................................. 88

Chapter 5  Single-Electron Spectroscopy .................................................................................................... 91
  Professor Raymond C. Ashoori

  • 5.1 Project Description ....................................................................................................................... 91
  • 5.2 Publications ................................................................................................................................ 95

Section 3  Optics and Devices

Chapter 1  Optics and Quantum Electronics ............................................................................................ 99
  Professor Hermann A. Haus, Professor Erich P. Ippen, Professor James G. Fujimoto,
  Professor Peter L. Hagelstein, Dr. Santanu Basu

  • 1.1 Additive Pulse Modelocking ........................................................................................................ 99
  • 1.2 Fiber Ring Laser .......................................................................................................................... 100
  • 1.3 Long Distance Fiber Communications ...................................................................................... 101
  • 1.4 Squeezing .................................................................................................................................... 101
  • 1.5 Integrated Photonic Components: The Channel Dropping Filter ............................................. 103
  • 1.6 Ultrafast Nonlinearities in Active Semiconductors ..................................................................... 104
  • 1.7 Femtosecond Raman Measurements in Optical Fibers ............................................................. 105
  • 1.8 Coherent Phonons in Solid-State Materials .................................................................................. 106
  • 1.9 Publications and Meeting Papers ............................................................................................... 107
  • 1.10 Ultrashort Pulse Generation in Solid State Lasers ...................................................................... 109
  • 1.11 Cavity Dumping Techniques for Modelocked Ti:Al₂O₃ Lasers .................................................... 110
  • 1.12 Spectrally Resolved Autocorrelation for Femtosecond Diagnostics ......................................... 113
  • 1.13 Ultrafast Phenomena in Materials and Devices ......................................................................... 113
  • 1.14 Laser Medicine ......................................................................................................................... 118
  • 1.15 EUV Laser Studies ..................................................................................................................... 124
  • 1.16 Ni-like Nb Studies ....................................................................................................................... 124
  • 1.17 H-like Boron Recombination X-ray Laser .................................................................................. 125
  • 1.18 Pump Laser Conversion ............................................................................................................. 129
  • 1.19 Development of a Densitometer ................................................................................................. 130
  • 1.20 Quantum-Well X-ray Detector .................................................................................................. 131
  • 1.21 Lineshape Theory ...................................................................................................................... 133
  • 1.22 Lattice-induced Reactions ......................................................................................................... 133

Chapter 2  Optical Propagation and Communication .................................................................................. 139
  Professor Jeffrey H. Shapiro, Dr. Robert H. Rediker, Dr. Ngai C. Wong

  2.1 Introduction .................................................................................................................................... 139
  2.2 Squeezed States of Light ................................................................................................................. 139
  2.3 Multiresolution Laser Radar Range Profiling ............................................................................... 140
  2.4 Optical Frequency Division and Synthesis ..................................................................................... 141
  2.5 Analog Processing of Optical Wavefronts Using Integrated Guided-Wave Optics ...................... 143
Table of Contents

Chapter 3  High-Frequency (> 100 GHz) and High-Speed (< 1 ps) Electronic Devices ............. 145

Professor Qing Hu

3.1  Facility for Millimeter-wave and THz Frequencies ......................................................... 145
3.2  Far-infrared Studies of Antenna-coupled Quantum-effect Devices .............................. 145
3.3  Femtosecond Dynamical Studies of Quantum Devices for Ultrafast Electronic Applications .......................................................... 148
3.4  High-Tc Superconducting Josephson Devices ................................................................. 149
3.5  Millimeter-wave and Infrared Superconducting Focal-plane Receiver Arrays .................. 153
3.6  Far-infrared (THz) Lasers Using Multiple Quantum Wells ........................................... 154
3.7  Publications .................................................................................................................. 158

Section 4  Surfaces and Interfaces

Chapter 1  Statistical Mechanics of Constrained Electronic Systems and Semiconductor Structures .................................................. 161

Professor A. Nihat Berker

*1.1  Introduction .................................................................................................................. 161
*1.2  Renormalization-Group Approach to Highly Correlated Electronic Systems .............. 161
*1.3  Statistical Mechanics of Mesoscopic Systems and Quantum Dots ............................. 164
*1.4  Collective Phenomena in Systems with Quenched Impurities and Frustration ............ 164
*1.5  Surface and Bulk Structures of Semiconductor Alloys ................................................ 165
*1.6  Publications ................................................................................................................ 166

Chapter 2  Synchrotron X-Ray Studies of Surface Disordering ........................................... 169

Professor Robert J. Birgeneau

*2.1  Introduction .................................................................................................................. 169
*2.2  Vicinal Semiconductor Surfaces .................................................................................... 169
*2.3  Chiral Melting of the Si(113) (3x1) Reconstruction ...................................................... 170
*2.4  Two-Dimensional Melting ............................................................................................ 171
*2.5  Publications ................................................................................................................ 171

Chapter 3  Chemical Reaction Dynamics at Surfaces ............................................................... 175

Professor Sylvia T. Ceyer

*3.1  A New Mechanism for Dissociative Chemisorption on Si: Atom Abstraction .............. 175
*3.2  A Model for Atom Abstraction by Surfaces .................................................................. 177
*3.3  Atom Abstraction and its Relationship to Thin Film Growth ..................................... 178
*3.4  Etching of Si(100) by Energetic Fluorine .................................................................... 179

Chapter 4  Semiconductor Surface Studies ............................................................................. 181

Professor John D. Joannopoulos

*4.1  Introduction .................................................................................................................. 181
*4.2  Defects on Surfaces .................................................................................................... 181
*4.3  Cross-sectional Scanning Tunneling Microscopy .......................................................... 184
*4.4  Publications ................................................................................................................ 185
Table of Contents

Chapter 5  Step Structures and Epitaxy on Semiconductor Surfaces ........................................... 187
Professor Simon G.J. Mochrie

• 5.1 Project Description ................................................................. 187
• 5.2 The Structure of the (3x1) Reconstruction of the Si(113) Surface ................................ 188
• 5.3 Morphology of Stepped Si(113) Surfaces ............................................. 189
• 5.4 Publications ........................................................................ 190

PART II  APPLIED PHYSICS

Section 1  Atomic, Molecular and Optical Physics

Chapter 1  Quantum Optics and Photonics .......................................................... 195
Professor Shaoul Ezekiel, Dr. M. Selim Shahriar

1.1 Data Storage Using Raman Induced Optical Spectral Holeburning .................... 195
1.2 Observation of Cooling Assisted Velocity Selective Coherent Population Trapping .......... 197
1.3 Suppression of Absorption of Resonance Fluorescence in a Folded
Three-Level Atom ........................................................................... 199
1.4 Inversionless Raman Lasing ....................................................................... 201
1.5 Creating a Large Angle Coherent Atomic Beamsplitter without a Magnetic Field .......... 202
1.6 Stimulated Brillouin Scattering Ring Laser Gyroscope ....................................... 204

Chapter 2  Basic Atomic Physics ........................................................................ 209
Professor Daniel Kleppner, Professor David E. Pritchard, Professor Wolfgang Ketterle

2.1 Testing Quantum Chaos with Rydberg Atoms in Strong Fields ......................... 209
• 2.2 Millimeter-Wave Frequency Measurement of the Rydberg Constant ................. 215
• 2.3 Precision Mass Spectroscopy of Ions ....................................................... 218
• 2.4 Atom Interferometry .............................................................................. 221
• 2.5 Cooling and Trapping Neutral Atoms .......................................................... 225

Section 2  Plasma Physics

Chapter 1  Plasma Dynamics ......................................................................... 231
Professor George Bekefi, Professor Abraham Bers, Professor Bruno Coppi,
Professor Jonathan S. Wurtele, Dr. Stefano Migliuolo, Dr. Abhay K. Ram,
Dr. Barrett Rogers, Dr. Linda E. Sugiyama

1.1 Relativistic Electron Beams ........................................................................ 231
1.2 Plasma Wave Interactions - RF Heating And Current Generation ..................... 237
1.3 Physics of Thermonuclear Plasmas ................................................................ 249

Section 3  Electromagnetics

Chapter 1  Electromagnetic Wave Theory and Applications ...................................... 263
Professor Jin Au Kong, Professor Terry Orlando, Dr. Robert T. Shin, Dr. Y. Eric Yang

1.1 Remote Sensing of Earth Terrain .................................................................. 263
• 1.2 Electromagnetic Waves in Multilayer Media ............................................... 265
1.3 Simulation of Electromagnetic Wave Scattering ............................................... 266
1.4 Superconducting Transmission Lines ............................................................. 269
## Table of Contents

### Section 4  Radio Astronomy

#### Chapter 1  Radio Astronomy  .................................................. 277

*Professor Bernard F. Burke, Professor David H. Staelin, Professor Jacqueline N. Hewitt, Dr. Philip W. Rosenkranz*

<table>
<thead>
<tr>
<th>Section / Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Extragalactic Radio Source Studies</td>
<td>277</td>
</tr>
<tr>
<td>1.2 Studies of Gravitational Lenses</td>
<td>279</td>
</tr>
<tr>
<td>1.3 Radio Interferometry of Nearby dMe Stars</td>
<td>280</td>
</tr>
<tr>
<td>1.4 The Detection of Astronomical Radio Transients</td>
<td>280</td>
</tr>
<tr>
<td>1.5 Algorithms for Advanced Microwave Sounding Unit Operational Use</td>
<td>282</td>
</tr>
<tr>
<td>1.6 Earth Observing System: Advanced Microwave Sounding Unit</td>
<td>282</td>
</tr>
<tr>
<td>1.7 High-Resolution Passive Microwave Imaging of Atmospheric Structure</td>
<td>283</td>
</tr>
<tr>
<td>1.8 Rapid Precision Net-Form Manufacturing</td>
<td>284</td>
</tr>
<tr>
<td>1.9 Conformal Experiment Design</td>
<td>284</td>
</tr>
<tr>
<td>1.10 Recognition of Natural Nearly Repetitive Signals</td>
<td>284</td>
</tr>
</tbody>
</table>

### PART III  SYSTEMS AND SIGNALS

#### Section 1  Computer-Aided Design

##### Chapter 1  Custom Integrated Circuits  .................................. 289

*Professor Jonathan Allen, Professor John L. Wyatt, Jr., Professor Jacob White, Professor Srinivas Devadas*

<table>
<thead>
<tr>
<th>Section / Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Custom Integrated Circuits</td>
<td>289</td>
</tr>
<tr>
<td>1.2 Analog VLSI Systems for Integrated Image Acquisition and Early Vision Processing</td>
<td>292</td>
</tr>
<tr>
<td>1.3 Parallel Algorithms for Device Simulation</td>
<td>300</td>
</tr>
<tr>
<td>1.4 Numerical Simulation of Short Channel MOS Devices</td>
<td>301</td>
</tr>
<tr>
<td>1.5 Coupled Simulation Algorithms for Microelectromechanical CAD</td>
<td>302</td>
</tr>
<tr>
<td>1.6 Numerical Techniques for Simulating Josephson Junction Arrays</td>
<td>303</td>
</tr>
<tr>
<td>1.7 Efficient 3-D Interconnect Analysis</td>
<td>303</td>
</tr>
<tr>
<td>1.8 Adaptive Gridding Techniques for Multipole-Accelerated Solution of Integral Equations</td>
<td>304</td>
</tr>
<tr>
<td>1.9 Circuit Simulation Algorithms</td>
<td>304</td>
</tr>
<tr>
<td>1.10 Techniques for Embedded System Design, Formal Verification, and Synthesis for Low Power Dissipation</td>
<td>305</td>
</tr>
</tbody>
</table>

##### Chapter 2  Computer-Integrated Design and Manufacture of Integrated Circuits  309

*Professor Donald E. Troxel*

<table>
<thead>
<tr>
<th>Section / Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>309</td>
</tr>
<tr>
<td>2.2 Principal Objectives</td>
<td>309</td>
</tr>
<tr>
<td>2.3 Principal Accomplishments</td>
<td>310</td>
</tr>
<tr>
<td>2.4 Process Flow Representation</td>
<td>311</td>
</tr>
<tr>
<td>2.5 Technology CAD Framework</td>
<td>312</td>
</tr>
<tr>
<td>2.6 Publications</td>
<td>313</td>
</tr>
</tbody>
</table>
Table of Contents

**Section 2 Digital Signal Processing**

**Chapter 1 Digital Signal Processing Research Program** ................................................................. 317

*Professor Alan V. Oppenheim, Professor Arthur B. Baggeroer, Professor Gregory W. Wornell*

1.1 Introduction ........................................................................................................................ 317
1.2 Active Noise Cancellation in Automobiles ........................................................................ 318
1.3 Single Mode Excitation in the Shallow Water Acoustic Channel ...................................... 318
1.4 Self-Synchronization of Chaotic Systems: Analysis, Synthesis, and Applications ............. 318
1.5 Algebraic and Probabilistic Structure in Fault-Tolerant Computation .............................. 319
1.6 Signal Processing Applications of Chaotic Dynamical Systems ...................................... 319
1.7 Wavelet-Based Representation and Algorithms for Generalized Fractal Signals .......... 320
1.8 Approximate Signal Processing ........................................................................................ 320
1.9 Code Division Multiple Access for Digital Storage ......................................................... 321
1.10 Estimation and Detection of a Class of Chaotic Signals .................................................. 321
1.11 Real-Time Active Noise Cancellation ............................................................................ 322
1.12 State and Parameter Estimation with Chaotic Systems .................................................. 322
1.13 Model-Based Analysis of Music ..................................................................................... 323
1.14 Nonlinear Models for Signal Processing ......................................................................... 323
1.15 Environmental Robustness in Automatic Speech Recognition .................................. 323
1.16 Active Noise Cancellation ............................................................................................... 324
1.17 Oceanographic Signal Processing .................................................................................... 325
1.18 Publications .................................................................................................................... 326

**Chapter 2 Advanced Telecommunications and Signal Processing Program** ...................... 329

*Professor Jae S. Lim*

2.1 Introduction ........................................................................................................................ 329
2.2 ATRP Facilities .................................................................................................................. 329
2.3 Very-low-bit-rate Video Representations .......................................................................... 330
2.4 Audio Compression Using Hierarchical Nonuniform Filterbanks ................................ 331
2.5 Transform Coding for High-Definition Television ......................................................... 331
2.6 Pre-Echo Detection and Reduction .................................................................................. 332
2.7 Video Source Coding for High-Definition Television ..................................................... 332
2.8 Error Concealment for an All-Digital HDTV System ..................................................... 333
2.9 Transmission of HDTV Signals in a Terrestrial Broadcast Environment ......................... 333
2.10 Position-Dependent Encoding ....................................................................................... 334
2.11 HDTV Transmission Format Conversion and the HDTV Migration Path ....................... 334
2.12 Speech Enhancement ...................................................................................................... 334

**Chapter 3 Combined Source and Channel Coding for High-Definition Television** .......... 337

*Professor William F. Schreiber*

3.1 Introduction ........................................................................................................................ 337
3.2 System Development ........................................................................................................ 337
3.3 Error Correction ............................................................................................................... 338
3.4 Channel Equalization ....................................................................................................... 338
TABLE OF CONTENTS

PART IV BIOELECTRONICS

Section 1 Genetic Analysis

Chapter 1 Genosensor Technology Development .................................................. 345

Dr. Mark Hollis

1.1 Introduction ................................................................................................. 345
1.2 Development of Genosensor Arrays for DNA Decoding ............................... 345
1.3 Microdetection Technology for Automated DNA Sequencing ......................... 348
1.4 Publications ............................................................................................. 349

PART V LANGUAGE, SPEECH AND HEARING

Section 1 Speech Communication

Chapter 1 Speech Communication ...................................................................... 355

Professor Kenneth N. Stevens, Dr. Joseph S. Perrelli, Dr. Stefanie Shattuck-Hufnagel

1.1 Introduction ................................................................................................. 356
1.2 Studies of the Acoustics, Perception, and Modeling of Speech Sounds .......... 356
1.3 Studies of Normal Speech Production .......................................................... 358
1.4 Speech Research Relating to Special Populations .......................................... 359
1.5 Speech Production Planning and Prosody .................................................... 361
1.6 Models for Lexical Representation and Lexical Access .................................. 362
1.7 Publications ............................................................................................. 363

Section 2 Sensory Communication

Chapter 1 Sensory Communication ..................................................................... 367

Professor Louis D. Braid, Nathaniel I. Durlach, Dr. William M. Rabinowitz,
Dr. Charlotte M. Reed, Dr. Mandayam A. Srinivasan, Dr. David Zeltzer, Dr. Patrick M. Zurek

1.1 Introduction ................................................................................................. 367
1.2 Hearing Aid Research ................................................................................. 367
1.3 Cochlear Prostheses ................................................................................. 374
1.4 Tactile Communication of Speech ............................................................... 375
1.5 Multimicrophone Hearing Aids ................................................................... 378
1.6 Superauditory Localization for Improved Human-Machine Interfaces ........... 379
1.7 Mechanistic Modeling of the Primate Fingerpad .......................................... 380
1.8 Peripheral Neural Mechanisms of Haptic Touch ......................................... 382
1.9 Biomechanics of Skin-Object Contact ....................................................... 383
1.10 Human and Robot Hands: Mechanics, Sensimotor Functions and Cognition .. 384
1.11 Virtual Environment Technology for Training (VETT) ................................. 388
1.12 The Virtual Sailor: Synthetic Humans for Virtual Environments .................. 390
1.13 Research on Improved Sonar Displays: A Human/Machine Processing System .... 391

Section 3 Auditory Physiology

Chapter 1 Signal Transmission in the Auditory System ......................................... 397

Professor Lawrence S. Frishkopf, Professor Nelson Y.S. Kiang, Professor William T.
Peake, Professor William M. Siebert, Professor Thomas F. Weiss, Dr. Bertrand Delgutte,
Dr. Donald K. Eddington, Dr. Dennis M. Freeman, Dr. John J. Guinan, Jr., Dr. John J.
Rosowski

x RLE Progress Report Number 136
# Table of Contents

## Section 1: Introduction

1.1 Introduction ........................................................................................................................... 397
1.2 Signal Transmission in the External and Middle Ear .......................................................... 397
1.3 Cochlear Mechanisms ........................................................................................................... 399
1.4 Stimulus Coding in the Auditory Nerve and Cochlear Nucleus ........................................ 400
1.5 Interactions of Middle-Ear Muscles and Olivocochlear Efferents ..................................... 402
1.6 Cochlear Efferent System ...................................................................................................... 403
1.7 Cochlear Implants .................................................................................................................. 404

## Section 4: Linguistics

Chapter 1 Linguistics .................................................................................................................. 411

*Professor Noam Chomsky, Professor Morris Halle*

1.1 Introduction ........................................................................................................................... 411
1.2 Abstracts of Doctoral Dissertations ..................................................................................... 411

## APPENDICES

**Appendix A** RLE Publications and Papers Presented .......................................................... 417

A.1 Meeting Papers ..................................................................................................................... 417
A.2 Journal Articles ..................................................................................................................... 431
A.3 Books/Chapters in Books/Published Proceedings ............................................................... 442
A.4 RLE Publications .................................................................................................................. 444
A.5 RLE Theses ............................................................................................................................ 444
A.6 Miscellaneous ....................................................................................................................... 446

**Appendix B** Current RLE Personnel .................................................................................... 449

**Appendix C** Milestones .......................................................................................................... 455

C.1 New Faculty and Staff .......................................................................................................... 455
C.2 Retirements ............................................................................................................................ 455
C.3 Chair Appointments .............................................................................................................. 455
C.4 Awards and Honors .............................................................................................................. 455

**Appendix D** RLE Research Support Index ........................................................................... 457

**Project Staff and Subject Index** ......................................................................................... 461
Introduction

The Research Laboratory of Electronics

The Research Laboratory of Electronics (RLE) was established in 1946 as the Institute's first interdepartmental laboratory. Originally organized under the joint sponsorship of the Departments of Physics and Electrical Engineering, RLE has broadened its interests to cover a wide range of research.

RLE Progress Report

RLE Progress Report Number 136 describes research programs at RLE for the period January 1 through December 31, 1993. Each chapter of the Progress Report contains both a statement of research objectives and a summary of research efforts for research projects listed. Faculty, research staff, students and others who participated in these projects are identified at the beginning of each project, along with sources of funding.

There are four appendices at the end of the report: Appendix A is a bibliography of RLE publications and papers presented by RLE staff during 1993; Appendix B is a roster of current RLE staff; Appendix C is a list of RLE faculty and staff milestones and honors received during 1993; and Appendix D is an index of RLE sponsors. In addition, the Project Staff and Subject Index provides access to the information in this report.

In addition to the Progress Report, RLE also publishes RLE currents, a biannual newsletter which focuses on RLE research interests; a technical report series; Speech Group Working Papers; and other information related to the Laboratory. The RLE Brochure, published in 1991 for RLE's 45th anniversary, is also available. Further inquiries may be addressed to:

Massachusetts Institute of Technology
Research Laboratory of Electronics
Communications Office
Room 36-412
Cambridge, Massachusetts 02139-4307
Tel. (617) 253-2566
Fax (617) 258-7864

Professor Jonathan Allen, Director
Research Laboratory of Electronics

The RLE environment provides both the freedom of action essential in an academic institution and the availability of large-scale laboratory facilities and services required by researchers. RLE's interdisciplinary setting offers many opportunities for creative and collaborative research. By fostering this powerful combination of research and education, RLE effectively penetrates beyond the horizon of new ideas and information.
Part 1  Solid State Physics, Electronics, and Optics

Section 1  Materials and Fabrication

Section 2  Quantum-Effect Devices

Section 3  Optics and Devices

Section 4  Surfaces and Interfaces
Section 1  Materials and Fabrication

Chapter 1  Heterostructures for High-Performance Devices

Chapter 2  Physics of InAlAs/InGaAs Heterostructure Field-Effect Transistors

Chapter 3  Gas Source Molecular Beam Epitaxy of Compound Semiconductors
Chapter 1. Heterostructures for High-Performance Devices

Academic and Research Staff
Professor Clifton G. Fonstad, Jr.

Visiting Scientists and Research Affiliates
Yuzo Hirayama,1 Kathy Meehan,2 Paul Gavrilovich,2 Sheila Prasad,3 Chulhun Seo4

Graduate Students

Technical and Support Staff
Karen Chenausky, Charmaine A. Cudjoe-Flanders, Angela R. Odoardi, Richard R. Perilli

1.1 Introduction
The broad objective of our research effort is to develop III-V quantum heterostructures for high-performance electronic, optoelectronic, and photonic devices for applications in high-speed optical communications and signal processing. To this end, we are developing: (1) new, higher-performance materials systems including InP-based InGaAlAs heterostructures and <111> oriented strained layer superlattices; (2) novel approaches to integrate laser diodes on VLSI-level electronic integrated circuits; (3) a new family of quantum-well-base, tunnel-barrier n-n-n transistors and near- and far-infrared optoelectronic devices; and (4) new damage-free in situ processing techniques for fabricating advanced quantum structure and embedded heterostructures.

The following sections describe our progress during the past year in the above research areas. Our group works closely with Professors Hermann A. Haus, Erich P. Ippen, and James G. Fujimoto to develop the optical device application, characterization, and modeling aspects of this program, with Professor Henry I. Smith to develop a novel distributed feedback laser structure, and with Professor Sylvia T. Ceyer to develop new in situ processing techniques.

1.2 Growth Optimization of MBE-Grown InAlAs on InP

Sponsors
Advanced Research Projects Agency/NCIPT
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff
Woo-Young Choi, Professor Clifton G. Fonstad, Jr.

InAlAs grown on InP is of great importance for many electrical and optical devices that utilize semiconductor heterostructures. The material quality of InAlAs, however, still leaves much to be desired. It suffers from high reactivity of aluminum with oxygen-containing residual species and alloy clustering due to the large difference in In-As and Al-As bond energies. Consequently, great care must be taken to establish growth conditions that minimize these degrading effects.

In growing InAlAs with an MBE machine that has ultraclean vacuum and high-source purity, there are two controllable growth parameters that significantly affect the resulting material quality: substrate temperature and As overpressure. The task is, then,
obtaining the optimal combination of these two that gives the best material quality. For this goal, we performed a systematic study in which InAlAs samples were grown on InP by MBE with different combinations of growth temperature and As overpressure and characterized by double crystal x-ray diffraction (DCXRD) measurements for the evaluation of crystalline quality, Hall measurements for electrical quality, and photoluminescence (PL) for optical quality.

Figure 1 shows different growth conditions under which five InAlAs samples were grown. These samples were all grown below the RHEED 2x to 4x transition temperature, as shown in the figure, above which the required As-stable surface condition cannot be maintained. Figure 2 shows the results of depend on MBE growth conditions. Particularly, high-As overpressure provides better quality than low As overpressure, probably because it minimizes alloy clustering. In addition, under a given As overpressure, higher growth temperature provides better quality because it most likely reduces impurity incorporation. If the temperature is too high, As deficiency may degrade the material quality. The similar trend is also shown in photoluminescence.

1.3 Fabrication of Ridge Waveguide Distributed Feedback Lasers by X-ray Lithography

Sponsors
Advanced Research Projects Agency/NCIPT
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff
Woo-Young Choi, Professor Clifton G. Fonstad, Jr., in collaboration with V. Wong, Professor Henry I. Smith, Kathy Meehan, Paul Gavrilovic, P. Whitney

We are developing novel InP-based ridge waveguide distributed feedback (RWG DFB) lasers which do not require any epitaxial regrowth or growth on corrugated substrates. Conventional DFB lasers have gratings buried within semiconductor materials, and this necessitates the complicated and often yield-limiting step of epitaxial overgrowth on top of gratings. Our approach relies on high-quality laser materials whose device performance, even with the ridge waveguide confinement scheme, is adequate for many applications. We simply add gratings next to the ridge waveguide after the formation of the waveguide; consequently, no complicated overgrowth step is needed and the simple and mature RWG fabrication technology can be fully utilized. Figure 3 schematically shows the structure of our RWG DFB device.

---

* Lasertron, Bedford, Massachusetts.
Numerical calculations show that the coupling coefficient values in the range of 10 to 100 cm\(^{-1}\) can be easily achieved if gratings are made as close to the ridge as possible and grating etching depth is about 0.3 \(\mu m\). To achieve these, we are utilizing x-ray lithography and \(\text{CH}_4/\text{H}_2\)-based RIE. X-ray lithography, with its large depth-of-focus and minimal proximity effect, can provide gratings of the required period on non-planar structures. RIE patterning of gratings as well as waveguides gives waveguides with vertical side-walls and gratings with large aspect ratios. With these techniques, we have successfully fabricated RWG DFB structures on monitor InP wafers as shown in the figure 4, and we are now in the process of fabricating complete RWG DFB devices.

![Figure 3. A schematic drawing of the proposed RWG DFB structure.](image)

**1.4 Numerical Calculation of Coupling Coefficients in Ridge Waveguide Distributed Feedback Lasers**

**Sponsors**

Advanced Research Projects Agency/NCIPT
Joint Services Electronics Program
Contract DAAL03-92-C-0001

**Project Staff**

Woo-Young Choi, Professor Clifton G. Fonstad, Jr., in collaboration with Jerry C. Chen and Professor Hermann A. Haus

As a part of our efforts in fabricating novel ridge waveguide (RWG) distributed feedback (DFB) lasers that do not require any epitaxial regrowth, we performed systematic numerical studies in which the coupling coefficients were calculated against many RWG structural parameters. In order to calculate the coupling coefficients, we used the coupled-mode theory which takes gratings as perturbation to grating-free two-dimensional waveguides. This requires an accurate two-dimensional field intensity profile, especially around the bottom parts of the ridge side-walls where most coupling between gratings and the evanescent fields occurs. Since simple two-dimensional waveguide solvers fail to provide accurate field intensity profile around this critical region, we used the newly developed imaginary-distance beam propagating method (IDBPM) for calculating accurate two-dimensional field intensity profiles. Figure 5 shows schematically the RWG of interest along with the parameters whose influences on the coupling coefficient were investigated. Figure 6 shows an example of guided mode intensity profile calculated by the IDBPM.

The investigation shows that, among the investigated parameters, the ridge width, the proximity of the gratings to the ridge, and the grating duty cycle have the most significant effects on the coupling coefficient. In addition, the coupling coefficients in the range of 10 to 100 cm\(^{-1}\) are easily achieved,
clearly demonstrating the feasibility of the proposed RWG DFB structure. Furthermore, calculations show that the proposed structure has advantage over the conventional DFB structure because it has more stable Bragg wavelength against the ridge width variations and the spatially-varying coupling coefficient DFB structure can be easily achievable.

Figure 5. A schematic drawing of a ridge waveguide distributed feedback laser with various parameters whose influences on the coupling coefficients are investigated.

Figure 6. A contour plot of a typical guided mode intensity profile calculated by the imaginary-distance beam propagation method.
1.5 Measurement of Excited-state Lifetimes in Narrow Quantum Wells

**Sponsor**
Joint Services Electronics Program
Contract DAAL03-92-C-0001

**Project Staff**
Lung-Han Peng, Jurgen H. Smet, Professor Clifton G. Fonstad, Jr., in collaboration with Professor Erich P. Ippen and Gadi Lenz

Recent achievement of 1.55 μm intersubband transitions using three-monolayer wide InGaAs/AlAs/InAlAs pseudomorphic quantum wells on InP has opened up a venue to study the carrier dynamics and time resolved optical spectra of such transitions in this important optical communication wavelength window. Fast relaxation times and large optical nonlinear behavior are the major merits expected from such quantum-well structures.

The 1.55 μm intersubband transition was observed in a single quantum well structure measured in a waveguide geometry to benefit from multiple passes through the quantum well and to increase the absorption to 5 percent. Nevertheless, to successfully carry out femtosecond pump probe measurements, modifications of the sample may be required. Current work is directed toward:

1. Improving the confinement of the incident light to the active region by incorporating the single quantum well structure into a ridge waveguide.
2. Attempting to grow a stack of these pseudomorphic single quantum wells using strain compensation techniques.
3. Adding quantum wells on both sides of the thin quantum well to improve carrier confinement in the upper level and prevent carrier leakage to the InAlAs barrier region.
4. Exploring the AlGaAs/InAs system on GaAs. The barriers are preferably thick and in this material system they are unstrained.

---


1.6 Tunable Semiconductor Lasers

**Sponsor**
Advanced Research Projects Agency/NCIPT

**Project Staff**
Paul S. Martin, Professor Clifton G. Fonstad, Jr., in collaboration with Professor Hermann A. Haus

Wavelength division multiplexing (WDM) applications such as the channel dropping filter proposed by Professor Hermann A. Haus and Bobby Y. Lai require tunable lasers and/or optical amplifiers. We have proposed a new three-terminal addressable asymmetric double quantum well (IAADQW) structure (figure 7) in which current injection into two quantum wells located within a single mode rib waveguide can be independently controlled. Simulations we have done show a continuous tuning range of more than 20 nm about a center wavelength of 1.5 μm for the structure shown below with one 95 Å well and one 100 Å well (figure 8). This is significantly larger than the continuous tuning range of both three-section distributed Bragg reflector (DBR) and tunable twin-guide (TTG) lasers that are competing for use in WDM systems. Furthermore, because the IAADQW laser is a three-terminal device, light output power can be held constant over the entire tuning range, making the IAADQW an important candidate for WDM applications.

Fabrication of the IAADQW devices has begun using a self-aligned technique we developed in which a single photolithography step defines a photoresist/silicon diode layer. We use this layer as: (1) an etch mask for the definition of an optical waveguide using reactive ion etching (RIE); (2) an ion implant mask that we use to confine the current injection in the lower quantum well; and (3) a lift-off stencil for defining the N-type contact to the center well region. Use of this technique eliminates critical lithographic alignments that can make the fabrication of three-terminal devices prohibitively complex.
Chapter 1. Heterostructures for High-Performance Devices

**IAADQW Process Summary**

1) SiO₂ Deposition
2) Photolithographic Definition of Ridge Waveguide
3) RIE of SiO₂ and P AlGaAs Clad
4) Ion Implantation of Hydrogen Current Confinement Layer
5) Au₃ Evaporation and SiO₂ Assisted Lift-off
6) SiO₂ RIE Planarization
7) AuZn P Contact Evaporation

![Diagram of IAADQW process](image)

**Figure 7.** IAADQW process description.

![Graph of tuning range](image)

**Figure 8.** IAADQW simulated tuning range.
1.7 Integration of Vertically-emitting, In-plane Cavity Laser Diodes on GaAs VLSI Circuitry

Sponsors
Advanced Research Projects Agency/NCIPT Hertz Foundation Fellowship

Project Staff
Krishna V. Shenoy, Professor Clifton G. Fonstad, Jr., in collaboration with W. Goodhue

We have demonstrated for the first time key components of a fabrication sequence for manufacturing monolithic optoelectronic very large scale integration (OE-VLSI) circuits and systems. Commercially available refractory-gate metal VLSI GaAs MESFETs have recently been shown to be stable after more than three hours at 500°C. Thus, it is now possible to regrow optical sources, detectors, and modulators on fully processed MESFET circuitry. VLSI circuits and systems can, therefore, be built with optical inputs and optical outputs without electronic circuit design rule modification. Such OE-VLSI circuits and systems are useful for high-speed optical communication and optical computing (smart pixels).

An OE-VLSI circuit can be fabricated in four steps: (1) foundry fabrication of the electronic circuit design, (2) selective epitaxial regrowth of photonic heterostructures in window regions (GaAs substrate exposed), (3) removal of inter-window polycrystalline material and photonic device processing, and (4) photonic and electronic device interconnection. By cascading these mature technologies, high-performance OE-VLSI circuits are immediately available.

An OE transceiver (generic smart pixel) was designed, regrown, and is nearing completion to demonstrate the high-density integration of LEDs, ridge lasers, and in-plane surface emitting lasers (IPSELS) with state-of-the-art VLSI GaAs MESFET transceiver designs. IPSELS are attractive for the following reasons: (1) $V_{th} < 2V$ is compatible with DCFL MESFET logic levels; (2) heterostructure thickness = 4 $\mu$m is planar with electronics, (3) low divergence angles (13° FWHM), and (4) $\eta_d$ as high as 66 percent pulsed and 48 percent cw.

Figure 9 is a cross-section of a selectively grown strained-layer (Al,Ga,In)As laser heterostructure in which vertical mirror facets, a parabolic deflection mirror, and a ridge waveguide have been fabricated using chlorine ion-beam-assisted etching. Figure 10 is a scanning electron micrograph of a completed monolithically integrated IPSEL on a transceiver chip. Preliminary optoelectronic characterization is currently underway.

---

7 MIT Lincoln Laboratory, Lexington, Massachusetts.
Chapter 1. Heterostructures for High-Performance Devices

Project Staff

Eric K. Braun, Krishna V. Shenoy, Professor Clifton G. Fonstad, Jr., in collaboration with J. Mikkelson14 with G. Hansell,15 and A. Harton16 and C. Wyatt16

Integration of optical devices with VLSI GaAs circuits by MBE regrowth has recently been demonstrated. The time temperature budget of the epitaxial layers is constrained by the thermal stability of the GaAs MESFET circuits. Investigations to date have demonstrated that commercially available GaAs VLSI discrete devices exhibit stable characteristics for three hour thermal cycles at temperatures up to 500°C. In this research we are continuing to evaluate the DC and AC performance of commercial VLSI GaAs MESFETs after thermal cycles of varying time and temperature. DC tests on refractory-gate metal MESFETs, from Vitesse Semiconductor Corporation's HGaAs2 process and Motorola Corporation's CS-1 process, show reduced MESFET performance, after three hour thermal cycles, for temperatures above 500°C (figure 11). DC parameters were extracted from enhancement mode MESFETs (EFETs), depletion mode MESFETs (DFETs) and transmission line model structures on the standard process control monitor test bars. It is observed that the EFET and DFET parameter trends are in agreement. Threshold voltage, Schottky diode ideality, Schottky barrier voltage, contact resistance, and sheet resistance all show no significant trends. However, with increasing cycle temperature, transconductance and saturation current decrease as the channel series resistance increases.

Future work on the thermal cycle response of VLSI GaAs MESFET circuits will be directed in the following three areas: (1) Additional DC tests of the latest refractory-metal gate processes from Vitesse and Motorola spanning a wider time and temperature range. (2) AC tests of the Motorola and Vitesse processes including the measurement of ring oscillators. (3) Testing of a gate-less process flow provided by Vitesse to isolate the mechanism for the increasing channel resistance.

Figure 10. Scanning electron micrograph of a completed IPSEL structure monolithically fabricated in a dielectric window on a transceiver chip designed at MIT and fabricated at Vitesse Semiconductor Corporation through MOSIS.

1.8 Thermal Stability of GaAs MESFET VLSI Circuits

Sponsors

Advanced Research Projects Agency/NCIPT
National Science Foundation Fellowship
Hertz Foundation Fellowship

14 Vitesse Semiconductor Corporation, Camarillo, California.
16 Motorola Inc., Schaumberg, Illinois.

14 RLE Progress Report Number 136
1.9 Gas Source MBE of InGaAsP Laser Diodes on GaAs Substrates

Sponsor
Advanced Research Projects Agency/NCIPT

Project Staff
Joseph F. Ahadian, Professor Clifton G. Fonstad, Jr., in collaboration with Professor Leslie A. Kolodziejski

The collaborative effort between Professors Clifton G. Fonstad, Jr., and Leslie A. Kolodziejski is aimed at the development of Al-free compound semiconductor lasers as part of a program to develop commercially viable OptoElectronic VLSI circuit (OE-VLSI) technology. Lasers are ultimately to be regrown on commercially fabricated GaAs MESFET VLSI circuits. These VLSI circuits have been found to withstand temperatures up to 500°C for several hours without degradation. While this condition counterindicates the use of conventionally grown AlGaAs based lasers, which traditionally require growth temperatures greater than 600°C to achieve high-quality material, the InGaAsP system is normally grown at a substrate temperature of approximately 500°C.

At present, multiple quantum well InGaAs/GaAs/InGaP lasers are under development. The lasers are to be grown by GSMBE in Professor Kolodziejski’s laboratory. To date, high-quality InGaP epilayers on GaAs substrates have been achieved, and doping studies are underway. Phase I, the growth of a preliminary double heterostructure laser structure, began in February 1994. Phase II will involve growth of quantum well active regions and Phase III will build on this structure to achieve state-of-the-art graded index, quantum well performance.

1.10 High-Density OEIC Neural Systems Produced by Monolithic Integration of GaAlAs Light Emitting Diodes on GaAs MESFET VLSI Circuits

Sponsor
Advanced Research Projects Agency/NCIPT

Project Staff
Krishna V. Shenoy, Professor Clifton G. Fonstad, Jr., in collaboration with A. Grot17 and D. Psaltis17

Recently it has been shown that commercially available, self-aligned VLSI GaAs MESFETs are stable after more than three hours at 500°C.18 Thus, it is now possible to use molecular beam epitaxy techniques to regrow optical sources, detectors, and modulators on fully processed, high-density MESFET circuitry without electronic design rule modifications. Such optoelectronic arrays are potentially extremely useful in optical neural systems where planes of electronic “neurons” are optically interconnected in the third dimension.

A first generation optoelectronic “bump” neuron was fabricated to demonstrate the integration of two phototransistor for optical inputs (E-MESFETs), nonlinear thresholding electronics (E/D-MESFETs),

17 California Institute of Technology, Pasadena, California.

and a regrown AlGaAs LED for optical output. The desired optoelectronic thresholding function was achieved while maintaining good electronic circuit performance and LED efficiencies.

Based on this initial success, three additional generations of optoelectronic neural chips have been designed, regrown, and fabricated incorporating increasingly complex neural circuits and more refined processing techniques. These techniques include regrowth in dielectric windows as small as 10 μm x 10 μm, bottom side n+ LED contacts made by regrowth on n+ source/drain ion-implanted regions, and monolithic LED-probe pad interconnect metallization. Figure 12 shows a cross section of a Generation 4 neural chip with an LED regrown on a source/drain ion implanted region in a dielectric window. Winner-take-all (WTA) circuits, where the only LED that is on is the output associated with the neuron in an array with the largest optical input signal, were successively demonstrated. Figure 13 shows the output power of two competing branches when the optical power on unit 1 was fixed at 30 nW while the power on unit 2 varied from 0 to 100 nW. A Generation 5 neural chip is in progress, which will demonstrate a 100-element optoelectronic WTA array and will provide a significantly higher speed alternative to spatial light modulator arrays currently used in optical neural systems.

---

Figure 12. Cross-section of a Vitesse/MOSIS chip with LED grown in the dielectric via.

---

Chapter 1. Heterostructures for High-Performance Devices

Figure 13. Output power of two competing branches of a winner-take-all circuit.

Figure 14. Transceiver block diagram (above) showing ECL input and output levels, a MSM detector with transimpedance amplifier, a MBE regrown laser diode, and a cross point switch.
1.11 Surface-Normal Optical Input and Output Cells for High-Density, High Speed GaAs MESFET-based OEICs

Sponsor
Advanced Research Projects Agency/NCIPT

Project Staff
Krishna V. Shenoy, Professor Clifton G. Fonstad, Jr., in collaboration with P.R. Nuytkens

Optoelectronic signal processors have become major components of advanced communication systems, real-time signal processors, and high-performance sensing systems. Integrated optoelectronics offers potential to significantly advance these systems as well as to enable the commercialization of board-to-board optical interconnects. In order to fully exploit the potential of optoelectronic devices, semiconductor processes and technology must be designed to monolithically integrate photodetectors, laser diodes, optical modulators and VLSI transistor circuits. Such monolithic optoelectronic device technology results in lower cost, reduced power consumption, higher bandwidth, lower weight, increased reliability and increased performance of optoelectronic systems.

An optoelectronic integrated circuit (OEIC) transceiver has been designed (Nuytkens, MIT), fabricated in the Vitesse Semiconductor, Inc. HGaAs$_3$ technology through MOSIS, molecular beam epitaxially regrown, and in-plane surface-emitting laser diodes (IPSELs) have been fabricated. The OEIC transceiver accepts either electrical or optical signal inputs and generates corresponding electrical or optical signal outputs. The OEIC incorporates an ECL electrical input receiver and output driver, an MSM photodetector, a transimpedance amplifier, a cross point switch, and control logic. The OEIC transceiver demonstrates four possible signal conversions (figure 14): (1) Electrical signal at ECL level input, through intermediate DCFL electrical level, to ECL level output; (2) Electrical signal at ECL level input, through laser current driver, to optical signal at 0.98 µm wavelength output (via regrown IPSELs); (3) Optical signal at 0.86 µm wavelength input, through transimpedance amplifier, to electrical signal at ECL level output, and (4) Optical signal at 0.86 µm wavelength input, through transimpedance amplifier and laser current driver, to optical signal at 0.98 µm wavelength output. Figure 15 shows the layout of the transceiver circuitry and the dielectric windows in which the IPSELs are regrown. Preliminary optoelectronic characterization is underway.

![OEIC Transceiver Layout](image)

Figure 15. Each 5 mm x 5 mm chip contains twenty four individual transceivers (right), each of which is 500 µm x 1500 µm including cascade probable pads and two 400 µm x 100 µm windows for (AlGaIn)As laser diode MBE regrowth.

---


1.12 Fiber-coupled GaAs MESFET-based OEICs

**Sponsor**
Advanced Research Projects Agency/NCIPT

**Project Staff**
Eric K. Braun, Yakov Royter, Professor Clifton G. Fonstad, Jr.

Recently, due to the fiberoptic telecommunications boom, much work has been done in the area of integration of semiconductor optical devices, such as modulators, detectors, and lasers, with driving electronics made up of MESFETs, MODFETs, and HBTs. However, the necessity to optimize both the optical and electronic devices simultaneously makes this approach slow. Our project is to integrate mature GaAs electronic ICs with optical components, by selectively growing using MBE optical devices on open areas of commercial quality IC chips. This has already been successfully achieved for integrating surface emitting laser diode array with such electronics. Our goal is to demonstrate a working optical switching circuit fabricated on and integrated with a GaAs IC chip (figure 16). The basic optical components of this system are passive waveguides, waveguide couplers and Y-junctions, phase modulators, and detectors. The electronics will consist of driving and amplifying circuit for the modulators, and detectors, as well as signal detection and routing decision digital circuits.

The first version of the circuit will consist of conventional double heterostructure waveguides and electro-optic modulators, with total internal reflection mirror coupled MSM detectors. A schematic of integrating a waveguide with a modulator or a detector is shown in figure 17. A quantum well heterostructure version of the system will also be attempted in order to take advantage of the strong electro-optic effect in QW structures. Finally, a structure with an InGaAs detector layer will be attempted, which will enable the system to operate at the industry standard wavelength of 1.35 μm.

One of the biggest challenges of this project is the fabrication of the optical circuits on the exposed areas of the chip without causing drastic deterioration of the IC performance. According to recent studies performed on the ICs, this means that all of epitaxial growth has to be performed at temperatures below 520°C. Thus, to quantify the effect of low temperature grown AlGaAs, we are performing a study of waveguide loss in this material. At the same time, we are working on the optimization of our optical circuit using BPM simulation software. Finally, we are addressing the issues of size reduction of the optical system, as well as the waveguide-fiber coupling issue.

Figure 16. A schematic representation of the proposed OEIC network signal router.

---

Chapter 1. Heterostructures for High-Performance Devices

1.13 Low-temperature, Selective-area MBE Growth of GaAlInAs Laser Diodes and Optical Waveguides on Semi-insulating GaAs Substrates

**Sponsor**
Advanced Research Projects Agency/NCIPT

**Project Staff**
Eric K. Braun, Yakov Royter, Krishna V. Shenoy, Professor Clifton G. Fonstad, Jr.

Integration of optical devices on commercially available VLSI GaAs circuits requires MBE growth of heterostructures at temperatures at or below ~500°C to avoid electronic device and circuit performance deterioration. Al containing MBE grown lasers at lowered growth temperatures (<600°C) exhibit substantially reduced performance with threshold current densities typically an order of magnitude or more higher than at optimal growth temperatures (~700°C). However, a mechanism for the reduced performance has not been clearly isolated. This research is intended to identify, as either optical or electrical, the nature of the dominant degradation mechanism.

To decouple the electrical and optical contributions to the increasing threshold current density as growth temperature is lowered, waveguide loss measurements on the same structure have been proposed. The experimental setup for measuring waveguide loss is shown in figure 18. A narrow linewidth laser (17 MHz) is focused into one end of the waveguide cavity and the transmitted beam is collected at the other end by microscope objectives. A heat lamp is used to modulate the index of refraction and length of the waveguide resulting in Fabry-Perot oscillations of the transmitted intensity. The ratio of the maximum to minimum intensity in the oscillation pattern can be related to the waveguide loss coefficient, as shown in figure 19. Using this setup the waveguide loss can be determined to an accuracy of 0.5 dB/cm. Waveguide loss measurements will be performed on GaAlInAs strip loaded waveguides, with both doped and undoped cladding layers, over the growth temperature range of 350°C to 700°C. Laser diodes, fabricated from the doped waveguides, will then be characterized. Using this information, it will be possible to design new heterostructure and doping profiles to improve lowered-temperature MBE grown AlGaAs laser performance. These proposed experiments will also be essential for the proper design of integrated passive optical components, including waveguides and modulators.
1.14 Applications of Resonant Tunneling Diodes in GaAs MESFET VLSI

Sponsor
U.S. Army Research Office
Grant DAAL03-92-G-0251

Project Staff
Rajni J. Aggarwal, Professor Clifton G. Fonstad, Jr., in collaboration with P.R. Nuytkens

We are investigating the integration of Resonant Tunneling Diodes (RTDs) with GaAs MESFETs. Our focus is on the determination of the optimum diode electrical characteristics for RTD based the SRAM cell shown in figure 20. Based on previous theoretical work, the desired diode has a low reso-
nance voltage, a high peak to valley current ratio (PVCR), and an extended low current valley.

Using molecular beam epitaxy, we have grown GaAs/AlAs, GaAs/InGaAs/AlAs and relaxed buffer InGaAs/AlAs resonant tunneling structures. In these structures, the smaller bandgap of InGaAs is used to reduce both the resonance voltage and valley current of the devices. The relaxed buffer devices use thick InGaAs layers of increasing indium fraction to change the lattice constant of the substrate, enabling the growth of higher percentage indium containing diodes. We have successfully demonstrated working diodes using all three structures. The 7:1 @ 300K (17.5:1 @ 77K) PVCR obtained on a relaxed buffer diode is the highest current. Future work will involve examining the dynamics of the switch in an effort to quantify more specifically the influence of diode current density, capacitance, and resistance on the switching process.

Diodes were tested in the SRAM configuration. Initial tests were performed to verify the bistable nature of a series diode chain. In these measurements, the diode supply voltage was varied as the node voltage was measured. Subsequent measurements focused on understanding the memory cell switching process. In these measurements, a programmable current source was used to simulate the MESFET. Figure 20 (right) shows a trace of a cell going through a write HI, hold HI, write LO, hold LO cycle. For a given cell, the maximum current necessary for switching will be the difference in the peak and valley currents for the diode with the largest PVCR. However, the switching current is a function of the supply voltage and can be reduced by adjusting the bias of the cell. There is an inherent tradeoff between switching and static current. Future work will involve examining the dynamics of the switch in an effort to quantify more specifically the influence of diode current density, capacitance, and resistance on the switching process.

Resonant Tunneling Diode Based Memory Cells

1.15 Polarization-resolved Infrared Spectra of Very Narrow AlAs/InGaAs/InP Quantum Wells

Sponsor
National Science Foundation
U.S. Army Research Office
Grant DAAL03-92-G-0251

Project Staff
Lung-Han Peng, Jurgen H. Smet, Yuzo Hirayama, Professor Clifton G. Fonstad, Jr.

The application of III-V quantum well (QW) C1-to-C2 intersubband transitions for far (8 ~ 10 μm) and mid (1 ~ 5 μm) infrared (IR) detection have, owing to their quantum structure-enhanced dipole transition, fast response speed, and reliance
on more mature growth and process technologies, substantial advantages over the bulk mercury-based materials, which are known to suffer serious problems of device uniformity and yield.

To date, the IR device application of QW intersubband transitions has been hindered by the inappropriate geometry used to study their optical selection rules. Our group has developed a polarization-resolved IR techniques for Fourier transform infrared (FTIR) and near infrared grating (NIR) spectroscopic study of QW optical transitions in the wavelength range from 400 nm to 20 μm.

Figure 21 shows the device structures, experimental setup, and polarization-resolved IR absorption data used in our study of AlAs/InGaAs QW intersubband transitions. By engineering the composition and width of the InGaAs QW, we have succeeded in bringing the intersubband absorption peak wavelength into the NIR (1.55 μm ~ 3 μm) regime. Our data also shows that by employing strained InGaAs/AlAs QWs, intersubband transitions exhibit (1) equally strong TE- and TM-polarization activity and (2) strain-induced polarization splitting.

Future work in strain InGaAs/AlAs QW waveguide devices will focus on their applications in integrated optics by using (1) linear intersubband absorption for fast IR detectors and modulators in the 1.55 μm spectral region for fiber optics communication systems and (2) nonlinear effects of QW intersubband and interband transitions for surface emitting of visible sum-frequency generation.

Figure 21. Polarization-resolved IR absorption data for strain InGaAs/AlAs single quantum wells measured in a waveguide configuration. Samples (a) ~ (h) are n° InGaAs wells of 3, 4, 5, 6, 10, 15, 17, and 22 monolayers in thickness.
Chapter 1. Heterostructures for High-Performance Devices

Figure 22. Linear dependence of QW intersubband polarization splitting with QW strain.

1.16 Symmetry Properties of Quantum Well Subband Energy Levels and Selection Rules for Intersubband Transitions

Sponsor
National Science Foundation

Project Staff
Lung-Han Peng, Professor Clifton G. Fonstad, Jr., in collaboration with Professors R.V. Jones and Victor Ehrenrich

According to an oversimplified picture of a double-barrier quantum well (DB-QW) structure in the energy-position space, it is widely assumed this DB-QW is of inversion symmetry in the growth direction as well as in the QW plane. The eigen wave functions of DB-QW subbands are typically written as a product of sin or cos functions in the growth (z) direction with those of plane waves solutions in the QW plane (x-y directions). As a result, the intersubband transitions in DB-QW are believed to be active only for light polarized with electric field along the z direction (TM-polarization).

Our polarization-resolved intersubband data on DB-QW (figure 21), however, clearly resolved (1) equally TE- and TM-polarization intersubband activity and (2) strain induced polarization splitting. Intersubband data of multiple quantum wells measured in the conventional Brewster angle configuration were confirmed to be corresponding to the TE-active signals measured in the polarization-resolved waveguide configuration. The experimental discrepancy in the Brewster angle configuration was caused by the large index of refraction in the semiconductors, as a result, 90 percent of the refracted light were TE-polarized into the QW plane.

We consider the DB-QW intersubband optical selection rules from the symmetry point of view. Near the k=0 point in the Brillouin zone, DB-QW of zinc blende III-V materials has D2d symmetry that does not include an inversion center. The Bloch states of C1 and C2 QW conduction subbands are S- and P-like at k=0 instead of being both S-like in the old theory. The corresponding wavefunction at k=0 point reduces to those of Bloch states and can be written as cos X cos Y cos Z, and p(sin X cos Y cos Z), where p stands for the permutations of the X,Y,Z coordinates in the parenthesis.

Our model analysis leads to the conclusions that C1-to-C2 intersubband transitions (1) are of equal absorption strength for both polarizations (parallel and perpendicular to the QW plane) and (2) have strain induced polarization splitting. The later predication was confirmed in figure 22, where a linear dependence of intersubband polarization splitting versus QW strain was clearly demonstrated.

Our model further predicates the possibility of using counter-propagating TE- and TM-active intense coherent intersubband nonlinear excitations to generate surface-emitting sum-frequency. Experimental work is underway to confirm this proposal.

---

23 Harvard University, Cambridge, Massachusetts.
1.17 Investigation of Infrared Intersubband Emission from InGaAs/AlAs/InP Quantum Well Heterostructures

Sponsor
U.S. Army Research Office
Grant DAAL03-92-G-0251

Project Staff
Jurgen H. Smet, Professor Clifton G. Fonstad, Jr., in collaboration with Professor Qing Hu

In this work, we study the possibility of achieving population inversion between two quantum well subbands in the conduction band, with an electrically pumped far infrared intersubband laser as the ultimate goal.

Until recently, proposals in the literature favored intra-well transitions between the first two levels of the quantum wells in superlattices or the middle quantum well of triple quantum-well structures to achieve this goal, because these transitions account for 96 percent of the total available oscillator strength. The adjacent quantum wells in these superlattices or triple quantum wells ensured selective injection into the second level and selective removal from the bottom subband. However, obtaining a sufficiently large population inversion and current density while maintaining all energy levels in resonance conditions appears difficult in both schemes.
Chapter 1. Heterostructures for High-Performance Devices

Recent spontaneous emission experiments in the 45 m wavelength range reported by Faist et al. have exploited inter-well intersubband transitions between off-resonance ground levels of double quantum wells. At the expense of a lower oscillator strength and uncertainty in the lasing wavelength, this approach is more suitable to attain a large population inversion and current density and is being explored for our wavelength of interest (50 μm).

1.18 High-Frequency/High-Speed Characterization, Analysis and Modeling of Heterojunction Bipolar Transistors, Laser Diodes, and m-s-m Photodetectors

Sponsor
National Science Foundation

Project Staff
Sheila Prasad, Professor Clifton G. Fonstad, Jr.

Emitter-down InGaAs/InAlAs/InP heterojunction bipolar transistors have been characterized at frequencies up to 40 GHz and small-signal equivalent circuit models have been obtained using the commercial Touchstone software as well as the simulated annealing (SA) algorithm. The bias dependence of the small-signal equivalent circuit elements was determined by measuring the scattering parameters at 39 different bias points. These results were used for the large-signal modeling of the HBT. The model obtained was based on first-order device physics and this was verified experimentally. When standard large signal modeling techniques such as the Harmonic Balance method are used, it was observed that the first-order model was sufficiently accurate for medium power levels with respect to the device under consideration. At higher power levels where the transistor is driven into hard saturation and cutoff, a more elaborate model is necessary to obtain a high third-order intermodulation product (IP3) and hence device linearity. Representative results are shown in Figure 23. Work on model improvement is in progress. Optimization of the HBT for high third-order intermodulation product is the goal of this work. Characterization of laser diodes and m-s-m photodetectors using the automatic network analyzer and on-wafer probe station is planned and modeling methods which have been effectively applied to HBTs will be used.

1.19 Damage-Free In-Situ UHV Etching and Cleaning of III-V Heterostructures Using Molecular Beams

Sponsors
AT&T Bell Laboratories Fellowship
Advanced Research Projects Agency/NCIPT
National Science Foundation
Grant ECS 90-07745

Project Staff
Isako Hoshino, Professor Clifton G. Fonstad, Jr.

The development of damage-free ultrahigh vacuum (UHV) etching, cleaning, and regrowth techniques compatible with molecular beam epitaxy (MBE) and ex-situ processing of III-V heterostructures is a major challenge facing device researchers. The ability to selectively pattern, etch, and overgrow quantum heterostructures is crucial to the effective realization of integrated optical circuitry and quantum effect electronic structures. Present techniques to do this involve relatively high-energy ions or plasma sources which cause substantial subsurface or structural damage, much of which is impossible to remove or repair, especially on compound semiconductors which are more susceptible to such damage than silicon.

As a solution to the problem of process-induced damages, we have begun a project investigating the use of molecular beam (hot neutral beam, or also known as kinetic beam) techniques to etch and clean III-V substrates and heterostructures with a minimum of surface damage and allowing maximum flexibility in attaining various etch profiles. Depending on the etchant gas, it is anticipated that low energy (0.5 to 30 eV) kinetic beams can be used to (1) both directionally and isotropically etch-pattern III-V heterostructure wafers with no damage; (2) clean surfaces allowing epitaxial growth on wafers that have been removed from the UHV environment for external processing; and (3) selectively remove masking materials and clean surfaces suitable for subsequent overgrowth.


Currently, the construction of the differentially pumped UHV Kinetic Beam Etch (KBE) system designed to use a methane-hydrogen gas mixture using a supersonic beam source is near completion. The system is designed such that it can be connected to the existing Riber 2300 solid source MBE system in the future through a transfer mechanism of special design.

The initial function tests of the KBE system are scheduled to be performed in the near future. Subsequently, a full-scale characterization of the ability of the KBE system will be performed, mainly concentrating on etch rate, etch profile, and surface damage assessment. The extent of surface damage, if any, will be determined through photoluminescence and carrier mobility measurements using various III-V heterostructures grown by MBE.

1.19.1 Publications

Journal Articles


Chapter 1. Heterostructures for High-Performance Devices


Meeting Paper

Chapter 2. Physics of InAlAs/InGaAs Heterostructure Field-Effect Transistors

Academic and Research Staff
Professor Jesús A. del Alamo

Graduate Students
Sandeep R. Bahl

Undergraduate Students
Akbar A. Moolji

Technical and Support Staff
Charmaine A. Cudjoe-Flanders, Angela R. Odoardi

2.1 Introduction

Sponsors
Charles S. Draper Laboratories
Contract DL-H-441694
Joint Services Electronics Program
Contract DAAL03-92-C-0001
Texas Instruments
Agreement dated 08/14/91

The goal of this project is to explore the suitability of InAlAs/InGaAs heterostructure field-effect transistors for power applications at very high frequencies. This material system has recently emerged as a promising candidate for low-noise applications at ultra-high frequencies where world record results have been demonstrated.1 Unfortunately, before InAlAs/InGaAs HFETs can also be used for high-frequency power applications, the physics of the breakdown voltage of these devices needs to be better understood. This has been the objective of our project for the last few years.

Over the last few years, we have been studying ways to improve the breakdown voltage of InAlAs/InGaAs HFETs, that is, the maximum voltage that the device can handle. Our work has resulted in technological design criteria such as employing AlAs-rich InAlAs pseudosilts2 and carrying out selective recessed-mesa sidewall isolation3 that are now widely used in industry. Two years ago we identified the detailed physical mechanisms responsible for breakdown in InAlAs/n-InGaAs HFETs fabricated at MIT.4 We found that the breakdown path involves two different processes in series. First, electron thermionic field emission takes place from the gate over the InAlAs barrier into the InGaAs channel. This is followed by hot electron relaxation in the channel with impact ionization of electron-hole pairs.

Over the last year, in collaboration with the Daimler-Benz Research Center in Ulm, Germany, we have been studying the physics of breakdown in state-of-the-art quarter-micron gate length InAlAs/InGaAs modulation-doped field-effect transistors (MODFETs) fabricated at Daimler Benz. To our surprise, we found that the physics of breakdown in these devices is qualitatively identical to that of MIT's InAlAs/n-InGaAs HFETs in spite of the radically different geometries and layer structure. There were important quantitative differences which help us trace the origin of the chronically low break-

down voltage of MODFETs in this material system. In a separate effort, we have also studied the occurrence of impact ionization in InAlAs/n+InGaAs HFETs under regular operating conditions, that is, with the channel on and away from breakdown. We indeed found that substantial impact ionization is taking place in these devices at typical operating biases. This is important for low-noise microwave and photonic applications where excess gate current due to hole generation in the channel is a serious noise contribution. The following sections describe in more detail our technical findings and conclusions.

2.2 Physics of Breakdown In InAlAs/InGaAs MODFETs

In collaboration with Daimler-Benz Research Center, we have performed the first comprehensive experimental study of off-state breakdown in InAlAs/InGaAs MODFETs. The heterostructure of these devices is shown in figure 1. The cap was designed to be surface-depleted for higher breakdown voltage. Three different heterostructures were characterized in this study, with different InAs mole-fractions of x=0.53 (lattice-matched), x=0.62, and x=0.70 in the channel. For the device with x=0.70, the lower 250 Å of the channel was grown with a lattice-matching composition of x=0.53 to avoid excessive strain. Processing was carried out according to the sequence described in Dickmann et al.\(^5\) The cap was etched selectively, and mesa-sidewall isolation was used in all heterostructures as we originally proposed in 1992.\(^6\) MODFETs with \(L_0 = 0.28 \mu m\) and \(W_0 = 80 \mu m\) were characterized. For reference, the lattice-matched device had \(I_{d}(\text{max}) = 275 \text{mA/mm}, \quad g_m(\text{peak}) = 400 \text{mS/mm}, \quad f_t = 83 \text{GHz}, \quad \text{and} \quad f_{max} = 140 \text{GHz}.

These are excellent values for devices with high BV. In this study, both drain-source breakdown voltage, \(BV_{DS}\), and drain-gate breakdown voltage, \(BV_{DG}\), were measured at several temperatures using the drain-current injection technique we have innovated.\(^7\) The test current was 1 mA/mm.

A plot of \(BV_{DS}\) and \(BV_{DG}\) versus \(T\) is presented in figure 2. At 300 K, \(BV_{DS}\) was 8.9 V, 6.3 V, and 5.1 V for x=0.53, 0.60, and 0.70 respectively. These are very high values for InAlAs/InGaAs MODFETs. In all devices, \(BV_{DS}\) and \(BV_{DG}\) show a negative temperature coefficient. The drain-current injection technique indicated that \(V_{DG}\) is limited by gate breakdown for the entire range of temperatures in all devices. \(BV_{DS}\) and \(BV_{DG}\) track each other with a difference of about \(V_T\), consistent with the above finding. \(BV_{DS}\) increases from 7.9 V at 340K to 14.4 V at 220 K for x=0.53. Since impact-ionization has a positive temperature coefficient, it follows that breakdown cannot be a simple impact-ionization phenomenon, as might be expected for a narrow bandgap channel with long mean-free path.

We also investigated the temperature dependence of the gate current approaching breakdown. \(I_g\) was found to be thermally activated around room-temperature. Figure 3 shows an Arrhenius plot of \(I_g/T^2\) at \(V_{DG}=4 \text{V}\) for the three devices. It is clear that as x is increased in the channel, \(I_g\) increases and \(E_A\) is reduced. For \(V_{DG}=4 \text{V}\), \(E_A\) is 0.17 eV, 0.09 eV, and 0.06 eV for x=0.53, 0.62, and 0.70 respectively.


Figure 2. Drain-source and drain-gate breakdown voltages $B_{DSS}$ and $B_{DAS}$ vs. temperature as a function of InAs mole fraction, $x$, in the channel.

Figure 3. An Arrhenius plot of $\log(I_D/T^2)$ vs. $1000/T$ for $V_{DG}=4$ V, as a function of $x$.

In a side-gate measurement, we verified that holes are produced in the channel at breakdown. Since specially designed sidegate structures did not exist in this wafer, this measurement was carried out by positioning a probe very close to the drain of the device and applying a reverse voltage of -80 V. In this manner, the probe was able to capture a few holes generated in the channel.
All these findings are qualitatively identical to those that we previously obtained on InAlAs/n⁺-InGaAs HFETs. This is rather remarkable since the present devices differ considerably from those in Bahl and del Alamo⁸ in their much shorter gate-length, different cap design, the presence of dopants in the insulator, the absence of dopants in the channel, Si₃N₄ passivation, incorporation of gate-recessing, and different buffer-layer. The similitude of experimental observations regarding off-state breakdown in such dissimilar devices reveals how fundamentally the breakdown process is associated with the materials involved.

The thermionic-field emission/Auger-generation hypothesis we proposed for the InAlAs/n⁺-InGaAs HFET⁹ can explain all the findings for the MODFET. Essentially, electrons going from the InAlAs insulator to the InGaAs channel suddenly gain a kinetic energy equal to ΔEc from the conduction-band step. This is in addition to the energy they already acquired from the electric-field in the insulator. For Inₐ.₅₉Ga₄.₄₁As/InₓGaₙAs, ΔEc = 0.5 eV, which is a substantial fraction of the bandgap of In₀.₅₉Ga₄.₄₁As (0.73 eV). In this case, even the presence of a low electric-field in the insulator can give rise to impact-ionization in the channel.

The combined thermionic emission/Auger-generation process, shown schematically in figure 4a, is a two-step process. First, electrons are injected from the gate edge into the high-field drain-gate region of the insulator by thermionic-field emission. Second, because of the large conduction-band offset and the electric-field in the insulator, they enter the channel hot, and immediately relax their energy through impact-ionization. The electrons flow towards the drain, and the holes can either be extracted by the gate or flow towards the source, where they recombine with electrons. This process actually occurs in two-dimensions, with electron injection likely to take place sideways from the gate into the drain-gate gap, as illustrated in a two-dimensional sketch in figure 4b.

The physical understanding gained in the course of this work provides for design criteria for future improvements in the off-state breakdown voltage of InAlAs/InGaAs MODFETs for the first time. Specifically, our findings now explain the need for an insulator that has the largest possible Schottky barrier. This barrier suppresses electron thermionic-field emission and improves breakdown. Second, the bandgap in the channel must not be too narrow since it facilitates Auger generation of electrons and holes. As a third conclusion, correct engineering of the cap also impacts breakdown because it affects

---

Figure 4. A schematic diagram showing the combined effect of electron thermionic-field emission and Auger generation. (a) Basic mechanism, (b) two-dimensional drawing. Electrons are injected sideways through thermionic-field emission from the gate into the high-field drain-gate region of the insulator (step 1). They then enter the channel hot and impact-ionize by Auger generation (step 2).

---

the electric field distribution directly underneath the gate edge on the drain-end of the channel where electron injection from the gate to the channel takes place.

We now understand the reason for the superior breakdown voltage characteristics of the MIDFET (metal-insulator-doped channel FET), that we have been studying at MIT, over the more popular MODFET. The presence of dopants inside the insulator of the MODFET bends the bands sharply down underneath the gate, favoring thermionic field emission of electrons at lower energies. In the MIDFET, the absence of dopants in the insulator results in a more "squarish" barrier relatively suppressing electron emission out of the gate. This explains the drastically different activation energies obtained in both kinds of devices, about 0.1 eV for the MODFET versus 0.45 eV for the MIDFET.9

2.3 Impact Ionization in InAlAs/InGaAs HFETs

We have examined the occurrence of impact ionization in the channel of InAlAs/n-InGaAs HFETs fabricated at MIT under regular biasing conditions, that is, with the channel turned on and away from breakdown conditions. Impact ionization in the channel results in excessive shot noise in the drain current and in additional gate leakage current which compromises the sensitivity of photonic receivers based on these devices.

We carried out sidegate current measurements on specially designed HFETs on a heterostructure studied previously (figure 5).10 By applying a sufficiently negative bias to the sidegate contact (-20 V in this experiment), it is possible to remove a small fraction of the holes that might be produced in the HFET channel as a result of impact ionization. We analyze the sidegate current in a similar manner as Hui et al.11 which dealt with the gate current of MESFETs. In our case, the gate current is not a good indicator of impact ionization because of the barrier that the insulator presents to the holes in the channel (the valence band discontinuity between insulator and channel which is about 0.2 eV).

Following the approach of Hui et al., we plot the absolute magnitude of the sidegate current (minus a small baseline electron leakage current through the gate) over the drain current as a function of the inverse voltage drop at the drain-end of the channel. A straight line in a semilog plot that is independent of $V_{DS}$ is the key signature of impact ionization.11 Indeed, as figure 6 shows, this is what we observe in our devices. This study unmisatably shows that impact ionization is taking place in the channel of InAlAs/InGaAs HFETs under a variety of bias conditions. This important fact needs to be taken into consideration when designing circuit applications around these devices, particularly, low-noise amplifiers for microwave or photonic applications.

---


Chapter 2. Heterostructure Field-Effect Transistors

![Semilog graph of sidegate current over drain current versus inverse voltage drop at the drain-end of the channel. The bunching of data in a straight line independent of $V_{DS}$ indicates the occurrence of impact ionization in the channel.](image)

2.4 Publications and Conference Papers

Bahl, S.R., B.R. Bennett, and J.A. del Alamo. "Doubly-Strained In$_{0.41}$Al$_{0.59}$As/n'-In$_{0.65}$Ga$_{0.35}$As HFET with High Breakdown Voltage." *IEEE Electron Device Lett.* 14(1): 22-24 (1993).


Theses


Chapter 3. Gas Source Molecular Beam Epitaxy of Compound Semiconductors

3.1 Introduction

Current state-of-the-art epitaxial growth techniques employ various metalorganic and hydride gases to deliver constituent species to the substrate surface, particularly high vapor pressure species such as phosphorus and sulfur. Gas source molecular beam epitaxy (GSMBE) utilizes hydride gas sources and solid elemental sources. The more conventional growth approach, molecular beam epitaxy (MBE), uses only molecular beams derived from the thermal evaporation of elemental or compound solid sources.

All the research objectives described in this chapter are concerned with layered structures composed of compounds containing As and P, or Se, S, and Te. The presence of these high vapor pressure species suggests that many advantages will be attained through the fabrication of the various multilayered device structures using the gaseous source epitaxy approach.

In the chemical beam epitaxy facility at MIT, the epitaxial growth of both II-VI and III-V compound semiconductors is underway using GSMBE growth techniques. The chemical beam epitaxy laboratory consists of two gaseous source epitaxy reactors interconnected to several smaller chambers which are used for sample introduction and in-situ surface analysis and metalization. The multichamber epitaxy system allows for the fabrication of a large number of different heterostructures completely within a continuous ultrahigh vacuum environment. The interconnected reactors enable an additional degree of freedom in device design by providing the ability to integrate the II-VI and III-V material families into a single device. The III-V GSMBE reactor utilizes solid elemental sources of Ga, In, Al, Si, and Be and gaseous hydride sources of arsenic and phosphorus. The II-VI reactor currently uses a solid Zn source and gaseous hydrogen selenide, in addition to a nitrogen plasma source and a solid ZnCl2 source to achieve n- and p-type doping. The unique design of the II-VI reactor also allows the use of metalorganic sources. Figure 1 highlights the many heterostructure systems, based on II-VI materials, III-V materials, or on a combination of II-VI and III-V semiconductors, available for exploration. Many are currently being fabricated in our epitaxy facility.

Wide bandgap II-VI heterostructures have the potential for commercially viable short wavelength (visible to ultraviolet) optical sources for use in optical recording applications and bright emissive displays, for example. A key advance in the area of Zn chalcogenides involved the successful p-type doping of ZnSe-based semiconductors using nitrogen as the acceptor species, as first reported by Park et al.1 The ability to dope the wide bandgap II-VIs led to the fabrication of a pn diode injection laser which operated at blue/blue-green wave-

Chapter 3. Molecular Beam Epitaxy of Compound Semiconductors

Figure 1. The shaded area highlights the many different II-VI and III-V semiconductors and the various heterostructure configurations which are available for investigation in the MIT chemical beam epitaxy laboratory.

lengths by Haase et al.\textsuperscript{2} and H. Jeon et al.\textsuperscript{3} The laser structures consisted of (Zn,Cd)Se narrow bandgap well layers with Zn(Se,S) cladding barrier layers, although more recent structures are composed of barrier layers of (Zn,Mg)(Se,S).\textsuperscript{4} To minimize defect generation within the active regions of the device as well as to maximize the incorporation of the nitrogen acceptor species, the reported substrate temperatures have ranged from 150$^\circ$C to 250$^\circ$C. At low growth temperatures, control of the flux ratio of high vapor pressure Group II and Group VI species is very crucial and extremely difficult. By employing gas source epitaxy technologies, control of the constituent species via precision mass flow controllers (as opposed to solid source thermal effusion ovens) is anticipated to offer a solution to the difficulties encountered in controlling the composition of ternaries and quaternaries. To address these material-related issues, we have embarked upon a II-VI-based research program which emphasizes the growth, using gas source molecular beam epitaxy, of various heterostructures which have the potential for being visible light emitters.

In the next section, we will describe our progress in the growth and doping of ZnSe using gas source molecular beam epitaxy. The II-VI effort is complemented by a program having the goal to fabricate lattice-matched epitaxial buffer layers for ZnSe that are composed of (In,Ga)P. The III-V GSMBE system is also utilized for the fabrication of (In,Ga)(As,P) waveguide devices that operate as tunable filters at 1.55 $\mu$m, which is the wavelength


of interest for optical fiber communication. New projects which are in the early stages of investigation are described at the end of the chapter and take advantage of the many capabilities available in the chemical beam epitaxy laboratory.

3.2 Gas Source Molecular Beam Epitaxy of ZnSe, ZnSe:Cl and ZnSe:N

Sponsors
Advanced Research Projects Agency
Subcontract 284-25041
Joint Services Electronics Program
Contract DAAL03-92-C-0001
National Center for Integrated Photonic Technology
Contract 542-381

Project Staff
Professor Leslie A. Kolodziejski, Dr. Irina Mnushkina, Dr. Gale S. Petrich, Christopher A. Coronado, Philip A. Fisher, Easen Ho, Jody L. House, James R. Geraci, Kuo-Yi Lim

In the case of the growth of the II-VI material family, all the constituent species are high vapor pressure elements necessitating the reproducible control of effusion cells which operate at low temperatures. The difficulty in maintaining elemental effusion cells at very low temperatures (-180 - 300°C) is eliminated by the use of mass flow controllers which precisely regulate the flow of gaseous anion species. The ZnSe-based heterostructure system has relevance to the optoelectronics community for system applications requiring optical sources operating in the green to UV spectral range. In this program, the growth and doping of ZnSe by GSMBE are investigated.

ZnSe was grown on semi-insulating, n- and p-type (001) GaAs bulk substrates using elemental Zn and gaseous H₂Se. The growth temperature ranged from 240°C to 400°C, although the majority of the films were grown at 270°C. The substrate temperature was calibrated by observing the eutectic phase transition (356°C) of 500 Å of Au that was deposited onto Ge and was continuously monitored using an optical pyrometer. The H₂Se gas flow was varied from 0.8 to 4.2 sccm and resulted in a chamber pressure of approximately 2 x 10⁻⁵ Torr during growth; the H₂Se was thermally decomposed at 1000°C. The zinc flux, as measured by a water-cooled crystal oscillator placed in the substrate position, corresponded to a zinc deposition rate ranging from 0.5 to 2.3 Å/s. Under these conditions, the Se:Zn flux ratio ranged from 1.3:1 to 0.6:1. During the epitaxial growth process, the surface stoichiometry was monitored by observing the surface reconstruction using reflection high energy electron diffraction (RHEED). A radio frequency (RF) plasma cell was used for a source of atomic nitrogen acceptor species. The RF power to the plasma cell was varied from 100 to 500 Watts. A leak valve regulated the nitrogen gas flow, while an ionization gauge in the reactor measured the equilibrium nitrogen pressure, which ranged from 6 x 10⁻⁷ to 2 x 10⁻⁵ Torr. The ZnSe was doped n-type using a ZnCl₂ source having a source temperature which ranged from 150°C to 250°C. The ZnSe was grown at rates of 0.3 to 0.7 μm/hour to layer thicknesses of 1 to 4 μm.

The optical properties of the epitaxial ZnSe layers were measured by photoluminescence (PL). PL was obtained by optically exciting the samples with the 3250 Å line of a focused He-Cd laser, providing a power density of approximately 300 mW/cm². The ZnSe luminescence was analyzed with a half-meter spectrometer and a photomultiplier. Figure 2 shows the low temperature (10K) PL spectra obtained from a series of undoped ZnSe films grown under identical growth conditions, but with different substrate temperatures. For each growth temperature investigated, the PL spectrum was dominated by an intense donor-bound exciton feature having an energy of 2.798 eV and is speculated to be due to an unintentionally incorporated donor originating from the source materials. The deep level defect-related luminescence band, which is broadly centered about 2.25 eV, was at least 100-1000 times weaker in intensity than the near-bandedge features. To elucidate the presence of the defect-related luminescence, photoluminescence was also obtained at 77K. As seen in figure 3, as a function of growth temperature, the PL spectra showed that the donor-bound exciton persisted as the dominant feature. However, at the growth temperature extremes, a small amount of luminescence originating from deep levels was observed. As another indication of the high quality of the ZnSe layers, the integrated PL intensity of the near-bandedge feature was found to decrease to one-sixtieth of its 10K value as the sample temperature was increased to room temperature; room temperature luminescence was easily observed by the naked eye.
Chapter 3. Molecular Beam Epitaxy of Compound Semiconductors

Figure 2. The 10K photoluminescence intensity as a function of energy for ZnSe epilayers grown at various substrate temperatures. The dominant donor-bound exciton had an energy of 2.798 eV. The ZnSe film that was grown at 284°C, was deposited on an (In,Ga)P buffer layer.

Figure 3. The 77K photoluminescence intensity as a function of energy for ZnSe epilayers grown at various substrate temperatures. The dominant donor-bound exciton has an energy of 2.787 eV. The ZnSe film that was grown at 284°C, was deposited on an (In,Ga)P buffer layer.
Figure 4. The 10K photoluminescence intensity as a function of energy (the inset shows a higher resolution energy scan). The spectrum in (a) reflects the PL obtained from ZnSe grown under Se-rich surface stoichiometry conditions, whereas (b) was obtained from a sample grown under conditions of Zn-rich surface stoichiometry. All other growth parameters were held constant.

During the growth, the surface stoichiometry was varied to provide either a Zn-rich or Se-rich surface by modifying the H₂Se gas flow. Changes in the RHEED pattern from a (2x1) reconstructed pattern, suggesting a Se-rich surface, to a (2x2) reconstructed pattern, suggesting a Zn-rich surface, verified the presence of a particular surface stoichiometry. The surface stoichiometry, which was maintained throughout the growth, was found to influence the near-band edge PL features, as well as the unintentional impurity incorporation. Figure 4 compares the PL spectrum, obtained from an undoped ZnSe film grown under Se-rich surface stoichiometry (figure 4a) to a sample grown under Zn-rich surface stoichiometry (figure 4b). All other growth parameters were identical.

Regardless of the surface stoichiometry, the near-band edge luminescence was dominated by a donor-bound exciton transition. However, an increase in the intensity of the free exciton feature (at an energy of 2.803 eV) was apparent in the undoped films grown under conditions which provide a Zn-rich surface stoichiometry. Differences in the undoped carrier concentration were also observed depending on the surface stoichiometry. Room temperature Hall effect measurements were performed on the undoped ZnSe samples with indium contacts in the Van der Pauw geometry. Films grown with a Se-rich surface stoichiometry had free electron concentrations in the mid $10^{16}$ to low $10^{17}$ cm$^{-3}$ range, while films grown with a Zn-rich surface stoichiometry had free electron concentrations from the mid $10^{15}$ to low $10^{16}$ cm$^{-3}$.

The RF plasma cell provided a source of both atomic and molecular nitrogen species, while a ZnCl₂ effusion cell provided a source of Cl for investigation of the p- and n-type doping, respectively, of ZnSe using the GSMBE growth technique. The nitrogen pressure was monitored by an ionization gauge within the reactor (prior to insertion of the H₂Se gas). The Cl doping level was controlled by the temperature of the ZnCl₂ effusion cell. Characterization of the doped ZnSe layers was performed using photoluminescence, Nomarski microscopy, double crystal x-ray diffraction rocking curves, secondary ion mass spectroscopy (SIMS), and capacitance-voltage (C-V) measurements. The observed structural characteristics were similar to the undoped layers. Specifically, the surface remained featureless as examined by Nomarski microscopy for all doping conditions explored thus far; the (400) peak of the x-ray rocking curves exhibited typical full width at half maximum (FWHM) ranging from 175 to 200 arc seconds.

Low temperature PL was used to optically characterize both the nitrogen- and chlorine-doped ZnSe
Chapter 3. Molecular Beam Epitaxy of Compound Semiconductors

Figure 5. The 10K photoluminescence intensity as a function of energy for various nitrogen gas flows. (b)-(d) The nitrogen flow was increased to vary the incorporation of nitrogen acceptor species. The PL obtained from undoped ZnSe (a) is shown for comparison.

epilayers. The photoluminescence obtained from ZnSe:N layers are shown in figure 5. In this figure, the 10K spectra are compared for samples grown under identical growth conditions; however, the nitrogen gas flow was systematically increased to enhance the incorporation of acceptors into the lattice. Figure 5 illustrates that the characteristic PL features of the ZnSe:N films were strongly dependent on the degree of nitrogen incorporation. The growth conditions for the series of films in figure 5 were as follows: a 270°C calibrated substrate temperature, a Se-rich surface stoichiometry, and a constant RF power of 100 Watts for the plasma source. From the comparisons of the dominant transitions observed in the 10K PL, we deduced that the nitrogen concentration increased with higher nitrogen gas flow, as expected. For the sample grown with the lowest nitrogen flow, shown in figure 5b, the near-bandedge transitions were dominated by the neutral N acceptor peak at 2.793 eV. (PL obtained from an undoped sample is shown for reference in figure 5a.)

The free electron-to-acceptor transition (FA) was present at 2.716 eV indicating a nitrogen acceptor binding energy of 109 meV, assuming a 10 K bandgap energy of 2.825 eV. The donor-to-acceptor-pair (DAP) transition peak at 2.698 eV and the associated LO phonon replicas were also detected. Further increases in the nitrogen flow resulted in a spectrum dominated by the FA and DAP transitions (and their LO phonon replicas), as shown in figure 5c. At the highest nitrogen gas flows investigated (figure 5d), the FA transition disappeared, while the DAP and phonon replicas merged into a single broad feature. Similar PL spectra for variations in the nitrogen incorporation for ZnSe:N grown by MBE were observed by Ohkawa et al. In addition, for a given doping condition, a Zn-rich surface stoichiometry appeared to

incorporate the nitrogen species more effectively, since the PL spectrum from a Zn-rich surface stoichiometry exhibits a single broad DAP transition, while the spectrum from a Se-rich surface stoichiometry consisted of multiple, easily distinguished DAP (and LO phonon replicas) transitions. The PL spectra from 2 μm Cl-doped ZnSe epilayers consisted of a single donor-bound exciton transition at 2.795 eV. By comparing various samples, the intensity of the donor-bound exciton transition was found to increase linearly as the Cl concentration was increased from 2 x 10^{16} to 2 x 10^{17} cm^{-3}. The free electron carrier concentration was determined by room temperature Hall measurements and correlated to atomic incorporation levels obtained by SIMS measurements.\(^7\)

The incorporation level of nitrogen in the ZnSe:N samples was also determined by secondary ion mass spectroscopy. A nitrogen-implanted ZnSe standard was used to calibrate the nitrogen concentration, resulting in a detection limit that was estimated to be approximately 10^{17} cm^{-3}. For growth under Se-rich surface stoichiometry, the measured nitrogen concentration ranged from below the detection limit of the SIMS apparatus (< 10^{17} cm^{-3}) to 4.0 x 10^{18} cm^{-3}. In comparison, samples grown under Zn-rich conditions had nitrogen concentrations ranging from 2 x 10^{17} to 5.0 x 10^{18} cm^{-3}. These results agreed with the qualitative comparisons of the PL features obtained from the samples that were grown under Zn-rich and Se-rich surface stoichiometries.

The ZnSe:N films were electrically characterized by C-V measurements due to the difficulty in forming an ohmic contact to p-type ZnSe. The measurements were made using Cr/Au Schottky contacts in a ring-dot electrode configuration at a measurement frequency of 1 MHz. The devices were typically 200 μm dots with a 20 μm wide separation between the center dot and the outer ground plane. Thus far, net acceptor concentrations (N_A-N_D) of 1.0 x 10^{17} cm^{-3} have been measured and correspond to a nitrogen concentration of 1.5 x 10^{16} cm^{-3} (as measured by SIMS).

### 3.2.1 Publications


### Thesis

Chapter 3. Molecular Beam Epitaxy of Compound Semiconductors

3.3 Novel Epitaxial III-V Buffer Layers for Wide Bandgap II-VI Visible Sources

Sponsors
Advanced Research Projects Agency
Subcontract 284-25041
National Center for Integrated Photonic Technology
Contract 542-381
National Science Foundation
Grant DMR 92-02957

Project Staff
Professor Leslie A. Kolodziejski, Dr. Gale S. Petrich, Christopher A. Coronado, Philip A. Fisher, Easen Ho, Jody L. House, Kan Lu, Kuo-Yi Lim

The ZnSe-based material system, in which (Zn,Mg)(S,Se) barrier layers and (Zn,Cd)Se well layers are used, offers the ability to fabricate a multitude of heterostructure devices capable of emitting light throughout the entire visible region. A significant device bottleneck, however, is the lack of large area, high quality bulk II-VI substrates. As a result, the II-VI devices are grown on III-V substrates which can potentially lead to the presence of appreciable lattice-mismatch strain within the II-VI layer. The detrimental effects of lattice-mismatch constrain the present state-of-the-art blue-green semiconductor lasers to be designed to be fully strained, such that the structures are pseudomorphic to bulk GaAs. The use of (In,Ga)P epitaxial buffer layers as an "alternate substrate" enables the (In,Ga)P lattice constant to be continuously varied from that of ZnS (GaP) to roughly that of Zn0.5Cd0.5Se (InP) (see figure 1). By matching the lattice constant of the III-V "alternative substrate" to the II-VI structure, the strain is minimized in the II-VI device layers, and the misfit dislocations can be engineered to remain primarily near the (In,Ga)P/GaAs interface, as opposed to the ZnSe/(In,Ga)P interface. The use of these novel...
buffer layers is anticipated to lead to an improvement in the device's performance, as well as to provide greater flexibility in II-VI laser design to achieve shorter wavelengths of blue emission. Additional benefits are expected to be derived (1) from the chemical differences which exist between the ZnSe and (In,Ga)P, as compared to GaAs, for future device processing, and (2) by utilizing various band structure engineering techniques to provide a graded transition in the valence band discontinuity present between GaAs and ZnSe which is necessary for optimized electronic transport.

The III-V layers were grown using GSMBE in a dedicated reactor. The (In,Ga)P buffer layers were grown at 450°C (on a 0.5 μm homoepitaxial GaAs buffer layer), where the substrate temperature was calibrated using the melting point of InSb (525°C). Elemental effusions cells containing In and Ga, in addition to PH₃, which is cracked at 900°C, were used as sources. The (In,Ga)P surface exhibited a sharp (2x1) surface reconstruction during growth, as observed by RHEED. During the initial minutes following nucleation, RHEED intensity oscillations were observed and were used to estimate the alloy composition in real time. The precise alloy composition was determined by (400) and (511) reflections of high resolution x-ray diffraction rocking curves. The thicknesses of the (In,Ga)P buffer layers were (1) estimated from the Group III fluxes, (2) obtained from RHEED oscillations which indicate the growth rate, and (3) subsequently measured by a profilometer following the selective removal of the ZnSe and (In,Ga)P by chemical etching. At the completion of the (In,Ga)P buffer layer, the surface was coated in situ with amorphous As by lowering the substrate temperature to below 100°C in an arsenic-rich environment (obtained by using a cracked AsH₃ flux). The samples were then transferred ex situ to the II-VI GSMBE reactor. The amorphous As was thermally desorbed from the (In,Ga)P buffer layers at approximately 270°C. During the nucleation of ZnSe on the (In,Ga)P epitaxial surface, a streaky RHEED pattern was observed in most cases, and persisted throughout the entire growth of the ZnSe layer. However, a bulk diffraction pattern was observed from the ZnSe during the initial stages of growth, if the temperature of the (In,Ga)P buffer layer was increased to temperatures nearing 400°C during the As desorption process. Excessive phosphorus evaporation is speculated to have occurred at the higher As desorption temperatures and contributed to the three dimensional nucleation behavior of ZnSe on (In,Ga)P.

The ZnSe/(In,Ga)P/GaAs heterostructures were examined by high resolution (four and double crystal) x-ray diffraction. The (In,Ga)P compositions were calculated using both the (400) and the (511) rocking curves and correlated with the RHEED intensity oscillations. The information derived from the (511) rocking curves indicated that the (In,Ga)P films having thicknesses of roughly 1 μm were pseudomorphic to the GaAs substrate, and that even films having thicknesses greater than 4 μm were only partially relaxed. Figure 6a shows the (400) x-ray rocking curve obtained from the ZnSe/(In,Ga)P/GaAs heterostructure containing a 4.3 μm In₀.₅₅Ga₀.₄₅P buffer layer. Figure 6b shows the identical structure following the removal of the 1.1 μm ZnSe layer by selective etching. The FWHM for each of the peaks from the III-V layers were 30 arc seconds for GaAs and 40 arc seconds for the (In,Ga)P. The (511) rocking curves for this structure indicated that the (In,Ga)P was not fully relaxed, but still contained approximately 0.11 percent lattice mismatch between the ZnSe and the (In,Ga)P. According to the (511) rocking curves, the composition of the (In,Ga)P film was such that when fully relaxed, the in-plane lattice constant would be the same as that of ZnSe. The width of the (400) ZnSe peak reflects the presence of dislocations at the ZnSe/III-V heterointerface due to the presence of the lattice-mismatch which still remains.

Comparisons of the surface morphology using optical Nomarski microscopy and scanning electron microscopy (SEM), indicated that the ZnSe films grown on the (In,Ga)P buffer layers (and on the bulk GaAs substrates) were featureless at SEM magnifications as high as 12,000 times. For epitaxial buffer layers of (In,Ga)As lattice-matched to ZnSe, the III-V surface was reported to have significant cross hatching due to lattice relaxation. Similar surface features have been observed for InGa₁₋ₓP films which have been grown at In compositions (x = 0.55), which when fully relaxed, would have a lattice constant greater than the in-plane lattice constant of ZnSe. For In compositions of x = 0.55, the layers were only partially relaxed, but exhibited in-plane lattice constants lattice-matched to ZnSe; in this case, the (In,Ga)P surface appeared cross-hatched. In addition, for a given In composition, growth conditions using a higher growth temperature (500°C) resulted in a cross-hatched surface.

The optical properties of the ZnSe films were measured using PL spectroscopy. The 10K PL spectra

---

9 Private communication with R.L. Gunshor, Purdue University, LaFayette, Indiana.
of the ZnSe on the pseudomorphic (In,Ga)P buffer layers, or on bulk GaAs substrates, were dominated by the luminescence attributed to transitions associated with donor-bound excitons. Defect-related luminescence originating from mid-bandgap deep levels was not observed. However, the photoluminescence from ZnSe grown on the partially-relaxed (In,Ga)P buffer layers contained features identified as those due to both donor-bound excitons (2.795 eV) and free excitons (2.803 eV) having roughly the same intensity (as can be seen in figure 7). For the samples examined, which have a very high degree of purity, the $Y_0$ (2.603 eV) and $I_v$ (2.775 eV) transitions were also detected. Both the $Y_0$ and $I_v$ transitions have been reported to be due to extended defects\textsuperscript{10} which suggested that the ZnSe layer still contained defects due to lattice relaxation. The observation of the $Y_0$ and $I_v$ transitions served as additional evidence that these (In,Ga)P buffer layers were only partially relaxed, whereas the ZnSe was completely relaxed (and confirmed by cross-sectional transmission electron microscopy). The transition occurring at 1.938 eV is attributed to the PL originating from the (In,Ga)P buffer layer.

![Figure 7](image-url) 10 K photoluminescence spectrum of the 1.1 μm ZnSe film grown on a partially relaxed 4.3 μm (In,Ga)P film. The feature at 1.938 eV is attributed to the photoluminescence from the (In,Ga)P buffer layer. Features at 2.803, 2.795, 2.776 and 2.603 are identified as transitions due to free excitons, donor-bound excitons, $I_v$, and $Y_0$, respectively. The origin of the feature at 2.717 is unknown and under investigation. (The feature at 1.907 eV is the second harmonic of the 3250 Å line from the HeCd laser.)

Characterization by photoluminescence indicates that high quality ZnSe epilayers have been grown on (In,Ga)P buffer layers by GSMBE. However, the critical thickness of the (In,Ga)P epilayers greatly exceeded the value predicted by the Matthews and Blakeslee force balancing model (47 nm), as well as by a model which uses an assumption of energy balance (3.2 μm). Various techniques that induce lattice relaxation are under investigation so that we can control the selection of a specific in-plane lattice constant for heteroepitaxy. Thus far, (In,Ga)P is a promising epitaxial substrate material for the realization of devices based on ZnSe and related II-VI compounds.

### 3.3.1 Publications


### Thesis


### 3.4 Integrated Photonic Devices: The Channel-Dropping Filter

**Sponsor**

National Center for Integrated Photonic Technology Contract 542-381

**Project Staff**

Professor Hermann A. Haus, Professor Leslie A. Kolodziejski, Professor Henry I. Smith, Dr. Gale S. Petrich, Jay N. Damask

The channel-dropping filter, first proposed by Professor Hermann Haus, is a member of a family of integrated photonic wavelength-division multiplexers (WDM) and demultiplexers. Unlike other multiplexers which terminate and resolve an entire bit stream within one device, the channel-dropping filter can selectively add (or remove) only one wavelength channel, while the other remaining channels are left undisturbed. The property of single channel filtering offers a new degree of freedom to the WDM network architecture. Our objective is to fabricate the integrated resonant channel-dropping filter using the (In,Ga)(As,P) compound semiconductor material system.

The (In,Ga)(As,P) semiconductor system meets several of the material requirements for the fabrication and eventual implementation of the channel-dropping filter. The (In,Ga)(As,P) heterostructure system is compatible with optical communication systems operating at 1.55 μm and is lattice-matched to InP over a wide range of direct energy bandgaps ($\lambda_2 = 0.87 \mu m$ for InP through $\lambda_2 = 1.67 \mu m$ for In$_{0.5}$Ga$_{0.5}$As). The quaternary material system also possesses sufficiently large differences in the index of refraction for various alloy compositions, which is necessary to confine an optical mode, and the material has been reported to exhibit waveguide losses of $\leq 1$ dB/cm.

Our current research has focused on the device design, epitaxial growth, and fabrication of passive low-loss waveguides. Vertical optical mode confinement in the waveguides is achieved by controlling the index of refraction, via the correct selection of the energy bandgap for each epitaxial layer. The energy bandgap is determined by the mole fraction of the constituent alloy elements: In, Ga, As, and P.

---


Chapter 3. Molecular Beam Epitaxy of Compound Semiconductors

The III-V-dedicated gas source molecular beam epitaxy reactor uses elemental sources for In and Ga, and gaseous hydride sources for As and P. The ratio of Group III and Group V elements determines the energy bandgap and lattice constant for a particular quaternary alloy.

Control of the alloy composition for \( \text{In}_{1-x}\text{Ga}_x\text{As}_{y}\text{P}_{1-y} \) is difficult due to the vastly different sticking coefficients of the Group III and Group V species. For a narrow range of growth temperature, the sticking coefficients for the indium and gallium atoms are near unity, whereas the sticking coefficients for the arsenic dimers and phosphorus dimers are less than unity, varying strongly with substrate temperature. The In/Ga flux ratio, and hence composition, is selected by the appropriate choice of the two elemental source temperatures. However, the As/P composition ratio cannot be similarly established. In competition for surface sites, As will displace P by a 50 to 1 ratio. Therefore, the epitaxial surface must have an arsenic surface coverage depleted of As, so as to allow for P incorporation. Thus, the As/(In+Ga) ratio is controlled, and phosphorus is the remaining constituent species which is in abundance and primarily sets the Group V to Group III overpressure.

The control algorithm for the alloy composition is initiated by establishing the In/Ga ratio which will determine the bandgap of the alloy. By modifying the As content, the \( \text{In}_{1-x}\text{Ga}_x\text{As}_{y}\text{P}_{1-y} \) alloy is lattice-matched to the InP substrate. This control algorithm allows for variation of the energy bandgap, while maintaining the lattice match to InP.

Figure 8a shows an \( \text{(In,Ga)}\text{(As,P)}/\text{InP} \) rib waveguide design that has been optimized for both maximum coupling to a Bragg grating (not shown) and for minimum fabrication-induced waveguide loss. (The Bragg grating will be etched on the top of the rib.) The 200 \( \text{Å} \) quaternary layer provides an etch stop for the fabrication process, to enable the height of the waveguide rib to be well defined and repeatable from sample to sample. Figure 8b is a SEM micrograph of a waveguide facet that was stained to enhance the contrast between epitaxial layers. In a parallel effort, the masks necessary to characterize the waveguides have been designed and manufactured with emphasis on the study of waveguide loss, as measured by cutback and Fabry-Perot techniques, waveguide-waveguide coupling, and facet reflectivity of waveguides angled with respect to the cleaved end.

**Thesis**


**3.5 Heterovalent Interfaces Composed of II-V/III-V Heterostructures**

**Sponsor**

National Science Foundation
Contract DMR 92-02957
Grant DMR 90-22933

**Project Staff**

Professor Leslie A. Kolodziejski, Dr. Gale S. Petruch, Jody L. House, Kuo-Yi Lim

Epitaxial growth of the wide bandgap semiconductor ZnSe onto GaAs buffer layers is complicated by the disparity between the electronic configuration which exists during the formation of the heterointerface. Due to the change in valency occurring from one monolayer to the next as the II-VII/III-V interface is formed, interface states are generated at the heterointerface. The interface states affect the efficiency of the carriers being injected from the GaAs to the ZnSe (for example in a p-n diode configuration). One method for investigating the nature of the interface states in the ZnSe-on-GaAs "noninverted" heterostructure is through the examination of the electrical characteristics of depletion-mode MISFET devices, where ZnSe is the insulator for the GaAs channel. Modification of the density of interface states will be studied by engineering the stoichiometry of each layer during the interface formation. The fabrication of the ZnSe/GaAs MISFET devices involves the development of the wet etch and ohmic metallization techniques, in addition to the creation of an appropriately designed device mask set. The masks have been designed so that the gate lengths range from 2 to 50 microns. Currently, the growth parameters and the processing steps for the ZnSe/GaAs heterostructure are under development.

The fabrication of the inverted interface, i.e., GaAs-on-ZnSe, for the fabrication of multilayered structures composed of ZnSe wide bandgap barriers and GaAs quantum well layers, represents a challenge in materials science. The growth approach must address issues regarding materials which are fabricated at vastly different growth temperatures (the optimum growth temperatures differ by almost 300°C). The first effort toward fabrication of the inverted structure has been to examine the
Figure 8. A cross sectional view of the (In,Ga)(As,P)/InP rib waveguide. (a) The designed structure showing the individual layers and thicknesses. The 200 Å quaternary layer provides an etch stop for the waveguide. (b) A SEM micrograph of the InP/(In,Ga)(As,P)/InP waveguide structure.
effects of very low growth temperatures on the GaAs layers, as characterized by high resolution x-ray diffraction and photoluminescence. Only at the very lowest growth temperature (~270°C) is excess As incorporation observed. In addition, ZnSe layers (grown on GaAs) have been subjected to various temperature extremes, as would be likely to occur during the growth of the GaAs well layer. As determined by photoluminescence, the optical properties of the ZnSe layers appear to remain unaffected by rapid thermal anneals, for times typically encountered during growth of the well material. With these two encouraging results, experiments are underway to begin the growth of the inverted GaAs-on-ZnSe interface.

3.6 Optoelectronic Very Large Scale Integrated Circuits

Sponsor
National Center for Integrated Photonic Technology
Contract 542-381

Project Staff
Professor Clifton G. Fonstad, Jr., Professor Leslie A. Kolodziejski, Dr. Gale S. Petrich, Joseph F. Ahadian, Kan Lu

The project described herein has as an objective the development of Al-free compound semiconductor lasers for commercially viable OptoElectronic VLSI circuit (OE-VLSI) technology. The lasers which are being fabricated will eventually be regrown onto commercially fabricated GaAs MESFET VLSI circuits. These fully metallized VLSI circuits have previously been found to withstand temperatures of up to 500°C for several hours without degradation. Conventionally-grown (Al,Ga)As-based lasers require growth temperatures greater than 600°C to achieve high quality material, and therefore will not allow for regrowth onto the completed VLSI circuitry. In contrast, however, the (In,Ga)(As,P) heterostructure system is normally grown at substrate temperatures near or below 500°C. At present, multiple quantum well (In,Ga)As/GaAs/(In,Ga)P laser structures have been designed. High quality (In,Ga)P epi layers on GaAs substrates have been deposited, and both n- and p-type doping studies are underway to calibrate the dopant sources and growth parameters.

3.7 Heteroepitaxy of GaAs onto Corrugated Surfaces of Si

Project Staff
Professor Leslie A. Kolodziejski, Professor Henry I. Smith, Professor Carl V. Thompson, Dr. Gale S. Petrich, Sean M. Donovan, Kan Lu

GaAs-based devices provide unique optoelectronic capabilities which are not available with Si-based devices. However, the processing technology for GaAs is not as mature and well characterized as that for Si VLSI processing technology. A natural step would be to combine the two device systems via monolithic integration, which requires the heteroepitaxial growth of GaAs onto Si. A significant obstacle, however, occurs due to the 4.1 percent lattice mismatch, and the presence of large differences in thermal expansion coefficients between the two materials giving rise to unacceptably high dislocation densities in the GaAs overlayer. In earlier work by Professor Smith and coworkers,14 it was observed that by etching a silicon substrate to contain a saw-tooth-like pattern that the threading dislocation density in the surface regions of a heteroepitaxially-grown GaAs film was greatly reduced. A detailed investigation is underway in order to understand the mechanism responsible for the reduction in the dislocation density, which will result in the successful integration of GaAs and Si functionalities. The study will also aid in the understanding of polar on non-polar epitaxy using GaAs-on-Si as the test structure. Our work thus far has primarily emphasized refinement of the process for producing the saw-tooth patterned Si substrates and preparing the surface for GaAs epitaxy.

Section 2  Quantum-Effect Devices

Chapter 1  Statistical Mechanics of Quantum Dots

Chapter 2  Single-Electron Transistors

Chapter 3  Edge Structure of a Quantum Dot in a Large Magnetic Field

Chapter 4  Nanostructures Technology, Research, and Applications

Chapter 5  Single-Electron Spectroscopy
Chapter 1. Statistical Mechanics of Quantum Dots

Academic and Research Staff
Professor Boris L. Altshuler, Dr. Benjamin D. Simons, Dr. Nobuhiko Taniguchi

Graduate Students
Anton Andreev, Eduardo R. Mucciolo

Technical and Support Staff
Imadiel Ariel, Kristine M. Bowring

1.1 Project Description

Sponsor
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Quantum dots—isolated systems of electrons whose dimensions are smaller than typical electron dephasing length—are being investigated as potential elements of novel electronic devices. In addition to their interest for such applications their fundamental physical properties are of a substantial interest. Combination of quantum interference of electronic waves with geometrical restrictions in all directions and with the interaction between electrons yields a rich variety of behavior.

To develop a statistical approach to the description of chaotic energy spectra in the presence of external fields, we have started with a disordered mesoscopic ring with Aharonov-Bohm flux \( \Phi \). The energy of a given electron state \( E_n \) is a periodic function of \( \Phi \) with a period \( \Phi_0 = \hbar c/e \). At the same time on the interval \( \Phi < \Phi_0 \) this function \( E_n(\Phi) \) is random. We have found that the autocorrelation function of the derivatives \( dE_n/d\Phi \) (i.e., of the currents carried by the level at different fluxes) determined as

\[
C(\Phi) = \Delta^{-2} \langle \frac{dE_n(\Phi_1)}{d\Phi_1} \frac{dE_n(\Phi_1 + \Phi)}{d\Phi} \rangle
\]  

(1)

is universal, i.e., independent on the disorder and on particular characteristics of the quantum dot:

\[
x = X [C(0)]^{1/2}, \quad \varepsilon_\alpha(X) = E_\alpha(X)/\Delta
\]

(3)

in the interval \( \Phi_0 < \Phi < \Phi_c = g^{1/2} \Phi_0 \), where \( g \) is the conductance of the ring.

As soon as \( \Phi \) becomes of the order of or bigger than \( \Phi_0 \), the universality disappears, and \( C(\Phi) \) turns out to be dependent on sample geometry, degree of disorder, etc. However, for a disordered quantum dot, the \( C(\Phi) \)-function at \( \Phi = \Phi_c \), given the geometry, can be evaluated within the same perturbative approach as equation 2. At the same time, \( C(\Phi) \) at \( \Phi < \Phi_0 \) cannot be evaluated perturbatively: equation 2 apparently diverges when \( \Phi \rightarrow 0 \). We have developed a nonperturbative formalism to treat the region of very small fluxes. This allowed us to discover an even higher level of universality than is demonstrated by equation 2.

Suppose the Hamiltonian of our quantum chaotic system depends on an external perturbation through some parameter \( X \) (in our previous example \( X = \Phi \)), having eigenvalues given by the random functions, \( E_\alpha(X) \). Such a hamiltonian can describe an atom, a molecule, or a whole variety of mesoscopic quantum systems and devices in the presence of external field. It can include disorder and interaction between electrons. With no loss of generality, we assume that \( \langle \delta E_\alpha(X)/\delta X \rangle = 0 \), where \( \langle ... \rangle \) denote an average over \( X \) and over a typical range of levels. We have shown that the rescaling

\[
x = X [C(0)]^{1/2}, \quad \varepsilon_\alpha(X) = E_\alpha(X)/\Delta
\]

(3)

\[\]
where $C(0)$ is determined equation 1 (with the substitution of $X$ for $\phi$) makes the statistics of $\epsilon_\alpha$ universal, i.e., independent on particular quantum system and on the nature of the parameter $X$.

To demonstrate this remarkable universality together with Professor Daniel Kleppner's group, we compared the correlation functions $c(x) = C(X)/C(0)$ for a disordered quantum dot and for the hydrogen atom in a magnetic field and found very good agreement for the two completely different systems.3

We have calculated exactly a number of correlation functions such as the correlator of densities of states at different energies and values of the dimensionless parameter $X$.4 This allowed us to discover another remarkable fact: the universal correlations that describe the response of energy levels of, for example, a disordered metallic grain to a magnetic field or any other perturbation are equivalent to time-dependent correlations with inverse square interaction.5

Also, the methods that we have developed allowed us to study dielectric and transport properties of periodic systems (periodic arrays of quantum dots or antidots) when electron motion within a primitive cell is chaotic.6 For a large conductance $g$ of a single primitive cell, integration over the Fermi surface provides an effective averaging that leads to the universal answers. We have determined the dynamic dielectric function and ac conductivity exactly. It is possible to also describe exactly the transition between the diffusive motion of electrons and ballistic motion of Bloch waves.

In connection with the experiments on the current switching in a mesoscopic conductor, we have studied the effect of an external magnetic field on a defect that tunnels incoherently between two locations in a mesoscopic metal.7 The magnetic field influences the energy asymmetry (bias) and the polaron effect due to the conduction electrons. Both effects contribute to the magnetic field dependence of the tunneling rate. For a two-level system that has no bias without magnetic field, the modification of the polaron effect dominates at low fields and increases slightly the tunneling rate, while at large fields the bias increase can significantly suppress the average tunneling rate.

1.2 Publications


2.1 Project Description

Sponsors

Joint Services Electronics Program
Contract DAAL03-92-C-0001
National Science Foundation
Grant ECS-88-13250

Our JSEP-supported research on the behavior of electrons in nanometer-sized semiconductor structures has led to an unexpected and remarkable discovery. We found that, when a field effect transistor (FET) is made very small and separated from its source and drain by tunnel junctions, it behaves in a way completely different from a conventional FET. A conventional device turns on only once when electrons are added. For example, to turn on a Si transistor that is 0.5 \( \mu \text{m} \) square, a few thousand electrons must be added. However, if a device of similar size is separated from its leads by tunnel junctions, it turns on and off again each time an electron is added.

In collaboration with IBM, we fabricated a device structure using MBE GaAs as shown in figure 1 (inset). This device is similar to a conventional FET because heavily doped GaAs serves as the substrate and functions as the gate. To accomplish this, first a layer of AlGaAs is grown to serve as the insulator. On top of this is a layer of GaAs in which the electrons accumulate when a positive voltage is applied to the gate. To confine the electrons to a small region and isolate them from source and drain, gold electrodes are patterned on top of the GaAs layer. The protrusions of these electrodes create potential barriers through which electrons must tunnel when passing through the device. The constrictions also create a pool of electrons between the barriers.

Figure 1 also shows the behavior of the device as the gate voltage is increased. The conductance from source to drain shows sharp peaks turning on and off with ratios as high as \( 10^4 \). A calculation shows that the capacitance of the device is only \( 10^{-16} \text{F} \), so that the voltage between peaks or valleys in figure 1 is exactly that necessary to add a single electron to the transistor. Thus, the device is called a single-electron transistor.

For our purposes, it is more accurate to think about this and other structures in which electrons are confined in all three dimensions as artificial atoms. When electrons are confined to a small region of space their charge and energy become quantized. The quantization of energy is something we are familiar with: for instance, solving the Schrodinger equation in a small volume gives discrete energy levels. As the typical dimension \( r \) of the region is decreased, the spacing of the levels increases, typically as \( r^2 \).

However, quantization of the charge is less familiar to us. Although the charge of an electron in free space is quantized, the charge in any small region of a conductor is not, because the wavefunction of the electrons extends over the entire volume of conductor. Charge quantization is induced by iso-
Chapter 2. Single-Electron Transistors

10-1

Figure 1. Current through the controlled-barrier atom as a function of the voltage on the gate at a temperature of 60 mK. The inset shows a schematic of the device structure.

Figure 1. Current through the controlled-barrier atom as a function of the voltage on the gate at a temperature of 60 mK. The inset shows a schematic of the device structure.

Iating a small part of the conductor. From the scaling theory of conductance developed in the 1980s, we know that this isolation occurs when the conductance of the insulator is less than e²/h. Thus, when the tunnel barriers conduct less than e²/h, both the charge and energy of the isolated region are quantized.

A simple way to explain the periodic conductance resonances shown in figure 1 follows. An electron tunnels from one lead onto the pool of electrons between the barriers and then onto the other lead. First we assume that the pool is neutral, compensated by charge on the gate. To add a charge Q to the pool requires energy Q²/2C, where C is the total capacitance between the particle and the rest of the system. Since one can add no less than one electron, the flow of current requires a Coulomb energy e²/2C. Therefore, because of charge quantization there is an energy gap in the spectrum of states for tunneling. For an electron to tunnel onto the artificial atom, its energy must exceed the Fermi energy of the contact by e²/2C, and for a hole to tunnel, its energy must be below the Fermi energy by the same amount. Consequently, the energy gap has the width e²/C. If the temperature is low enough so that kT < e²/2C, neither electrons nor holes can flow from one lead to the other.

With the gate voltage V₉, we can alter the energy required to add charge to the particle. V₉ is applied between the gate and source, but if the drain-source voltage is very small, the source, drain and particle are at almost the same potential. With V₉ applied, the electrostatic energy of charge Q on the particle is

\[ E = QV₉ + Q²/2C. \]  

(1)

The first term is the attractive interaction between the positively charged gate electrode and the (negative) charge Q, and the second term is the repulsive interaction among the bits of charge on the particle. Equation 1 shows that the energy as a function of Q is a parabola with minimum at Q₀ = -CV₉. For simplicity, it is assumed that the gate is the only electrode that contributes to C; in
realities, there are additional factors. By varying $V_g$, we can choose any value of $Q_0$, which is the charge that would minimize the energy in equation 1, if the charge were not quantized. However, because the real charge is quantized, only discrete values of the energy $E$ are possible. When $Q_0 = -(N+1/2)e$, there is a smaller, but non-zero, energy for adding or subtracting an electron. Under these circumstances, no current can flow at low temperature. However, if $Q_0 = -(N+1/2)e$, the state with $Q = -(N+1)e$ and that with $Q = -(N+1)e$ are degenerate; the charge fluctuates between the two values even at zero temperature. Consequently, the energy gap in the tunneling spectrum disappears, and current can flow. Therefore, the peaks in conductance are periodic, occurring whenever $CV_g = Q_0 = -(N+1/2)e$, spaced in gate voltage by $e/C$. This explanation is called the Coulomb blockade model.\textsuperscript{5}

Artificial atoms can be made in various ways. One way uses only metals.\textsuperscript{5} For this "all-metal atom", the energy separation of the quantized levels is too small to measure, but the effects of charge quantization are clear. For this case, the Coulomb blockade works very well. Many researchers\textsuperscript{5} have made "quantum dots", which are analogous to the single-electron transistor with no gate. By varying the voltage between source and drain, one can measure the spectrum of energy levels for adding an extra electron to the atom.\textsuperscript{4} Figure 2 shows the differential conductance of our device as a function of source-drain voltage. Near zero bias there is a gap, in which the conductance is zero, caused by the Coulomb blockade. It takes an energy $e^2/2C$, where $C$ is the device capacitance, to add an extra electron to the artificial atom. Thus, no current flows until the source-drain voltage shifts the energy by this amount. Once the Coulomb blockade gap is overcome, a spectrum of peaks is seen in the differential conductance. These correspond to excited states of an artificial atom with an extra electron or hole. If the gate voltage is increased, energy levels shift to lower energy because the electrons are attracted by the positively charged gate.

In figure 2, we show the positions of peaks similar to those in the top panel plotted as a function of the gate voltage. The Coulomb blockade gap is clearly seen opening and then closing for each electron added to the atom. In addition, the quantized excited states are evident, separated by a smaller energy $AE$. The level at the top of the gap corresponds to the minimum energy required to add an extra electron to the atom. The next level is the amount of energy required to add one electron in the first excited state, and so on. At higher voltages (not shown) are levels required for adding two and more electrons.

From figure 1, one sees that the voltage necessary to add one electron is about 1 mV. That means that the Coulomb energy gap for our device is about 1 meV. As expected, when $kT$ is comparable to this, the single-electron behavior is washed out by thermal fluctuations. In fact, because other electrodes contribute to the capacitance of our small transistor, the energy scale is about 0.5 meV. Furthermore, our analysis shows that the conductance peaks disappear when $kT$ is about one-quarter of the gap, that is, at temperatures above a few K. Since any practical application will require operating temperatures that are much higher, a primary goal of our research is the exploration of what needs to be done to increase the energy scale. Reaching this goal will have a major impact on basic physics studies as well. The resolution of single-electron spectroscopies is $kT$. Therefore,


Chapter 2. Single-Electron Transistors

of the device. However, as the size of the device is reduced, the capacitance will eventually be dominated by the edges and will scale only with the linear dimension of the isolated region. Fortunately, the quantum mechanical energy spacing $\Delta E$ continues to scale with the area. Therefore, we believe that, with careful design of the MBE layer structure and the gate configuration, we should be able to increase the energy scale by a factor $\sim 100$ so that liquid nitrogen temperature operation is probable and room temperature operation might be possible. This will require close collaboration with researchers growing MBE crystals. We have established this kind of collaboration with Michael R. Melloch of Purdue University.

Another way to reduce capacitance is using Si instead of GaAs. In Si MOSFETs, the lower dielectric constants of both the semiconductor and the insulator (SiON) are an advantage for single-electron devices. Although the charges at the Si-SiO$_2$ interface cause unwanted potential fluctuations interfering with the device operation, their density is low and may not cause serious problems in a very small device. In our early experiments, we found that the energy scale introduced by these surface charges was about 1 meV. Therefore, if we increase the energy scale, the surface charges will become less important.

Making devices in Si offers an additional advantage because they can be easily integrated with conventional circuits. This advantage will be crucial for useful applications to become a reality.

2.2 Publications


Chapter 3. Edge Structure Of A Quantum Dot In A Large Magnetic Field

Academic and Research Staff
Professor Patrick A. Lee, Dr. Konstantin Matveev

Graduate Students
Dmitri Chklovskii

Technical and Support Staff
Imadiel Ariel

3.1 Project Description

Sponsor
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Two years ago, Professor Raymond Ashoori demonstrated that it is possible to tunnel a single electron on and off a quantum dot and map out the energy level structures of the first five or ten electrons. This experiment was also done in a magnetic field so that the spectrum can be monitored as the field is varied. In the future, it is expected that one can enter the fractional quantum Hall effect (FQHE) regime, i.e., when only the lowest Landau level is occupied. It is then important to theoretically study what is the expected behavior for the tunneling onto a quantum dot in the FQHE limit.

We addressed this problem two years ago, assuming that the electrons are confined by a rapidly varying potential. We found a strong suppression of the tunneling amplitude due to many-body effects, and novel features such as alternating strong and weak tunneling amplitude with increasing electron numbers in the case of the 2/3 filled Landau level. However, in practice the confining potential is produced by gates relatively far away, so that the confining potential is slowly varying. In the past year, we initiated a project to study the behavior of the edge of the quantum dot in the case of a slowly varying potential. It is important to understand the edge structure because tunneling occurs predominantly at the edge.

Traditionally, the FQHE is understood in terms of the Laughlin wavefunction, which describes an incompressible electron state at special filling factors. Recently, it has been suggested that a partially filled Landau level may be described by making a transformation to composite fermions. This is achieved by attaching two magnetic field flux quanta to each electron. The theory explains several experimental observations, and its predictions were recently experimentally confirmed.

We use the composite fermion approach to study an experimentally relevant system where electrons are confined to a quantum wire or a quantum dot. We solve for the electron wavefunction by multiplying it with a Jastrow-like function. Figure 1 demonstrates the good agreement between our calculations and the results obtained from the Laughlin wavefunction by Rezayi and Haldane. However, their study was limited to simple fractions and few electron systems.

---

Chapter 3. Edge Structure Of A Quantum Dot

Figure 1. Density profile for the 1/3 quantum Hall state. Dashed line: Rezayi-Haldane calculations for the Laughlin wavefunction. Full line: Our composite fermion Hartree calculation. Length is measured in magnetic length units. Unit corresponds to a completely filled Landau level.

We are able to extend our calculations to other fractional states and greater number of electrons. This allows us to perform the first systematic study of the fractional edge states. We find that the fractional edge states can be described very naturally in the composite fermion representation as the integer edge states and the snake states. We can predict the number of fractional edge states formed in a given confinement scheme as well as the density distribution (see figure 2). These results are crucial for the explanation edge magnetoplasmon experiments in mesoscopic devices. The results are also the starting point for a theory of tunneling onto the quantum dot in the FQHE regime.

3.2 Publications


Figure 2. Density profile for the 2/5 state (composite fermion Hartree calculation). Despite the ripples reflecting the density-density correlation function one can see that 2/5 state goes through a 1/3 state before entering the vacuum.
Chapter 4. Nanostructures Technology, Research, and Applications

Academic and Research Staff
Professor Henry I. Smith, Dr. Mark L. Schattenburg, Richard J. Aucoin, James M. Carter, Robert C. Fleming, Euclid E. Moon, Irving Plotnik, Scott E. Silverman

Visiting Scientists and Research Affiliates
Nubuyoshi Koshida

Graduate Students

Undergraduate Students
Julie C. Lew, Brian M. Smith

Technical and Support Staff
Donna R. Martinez, Mark K. Mondol, Jeanne M. Porter, Robert D. Sisson

4.1 NanoStructures Laboratory
The NanoStructures Laboratory (NSL) at MIT (formerly the Submicron Structures Laboratory) develops techniques for fabricating surface structures with feature sizes in the range from nanometers to micrometers and uses these structures in a variety of research projects. The NSL includes facilities for lithography (photo, holographic electron beam, ion beam, and x-ray), etching (chemical, plasma and reactive-ion), liftoff, electroplating, sputter deposition, and e-beam evaporation. Much of the equipment and nearly all of the methods utilized in the NSL are developed in-house. Generally, commercial integrated circuit (IC) processing equipment cannot achieve the resolution needed for nanofabrication, and it lacks the required flexibility. Research projects, described briefly below, fall into four major categories: (1) development of submicron and nanometer fabrication technology; (2) nanometer and quantum-effect electronics; (3) periodic structures for x-ray optics; spectroscopy and atomic interferometry; and (4) crystalline films on non-lattice-matching substrates.

4.2 Scanning Electron-Beam Lithography Facility

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001
Semiconductor Research Corporation
Contract 94-MJ-550

Project Staff
Scott E. Silverman, Professor Henry I. Smith, Juan Ferrera

Since its inception in 1978, the Submicron Structures Laboratory (recently renamed the NanoStructures Laboratory) has relied on outside facilities, most notably the Naval Research Laboratory (NRL) in Washington, D.C., and IBM in Yorktown Heights, New York, for scanning-electron-beam lithography (SEBL). In November 1993 we received, as a donation from IBM, an SEBL system, designated the VS-2A, shown in figure 1.

---

1 Tokyo University of Agriculture and Technology, Tokyo, Japan.
Chapter 4. Nanostructures Research

4.3 Spatial-Phase-Locked Electron-Beam Lithography

Sponsors

Joint Services Electronics Program
Contract DAAL03-92-C-0001
Semiconductor Research Corporation
Contract 94-MJ-550
U.S. Army Research Office
Grant DAAL03-92-G-0291

Project Staff

James M. Carter, Juan Ferrera, Scott E. Silverman,
Professor Henry I. Smith, Vincent V. Wong

It is well known that scanning-electron beam lithography can write extremely fine lines, ~10 nm in thin PMMA and ~1 nm in AlF₃. However, because writing fields in electron-beam lithography are quite small (10³ to 10⁴ beam steps), large-area patterns must be created by stitching together the small fields, using a laser interferometer to provide X-Y positioning information. However, it is often overlooked that, due to instability and drift, the precision with which this can be done is much poorer than the resolution. Typically, stitching errors of 20 to 50 nm (often larger) are observed at field boundaries.

Figure 1. Photograph of VS-2A scanning-electron-beam lithography system donated to MIT by IBM.

This is an experimental system based on many years of IBM technology development in SEBL. The VS-2A system will be the cornerstone of a new SEBL facility located in Building 38. The new facility should be in full operation by May 1994.

The goals of the new facility are to (1) provide the MIT research community with an in-house SEBL capability for writing directly on experimental device substrates; (2) advance the state-of-the-art in SEBL, particularly with regard to pattern placement accuracy and long-range spatial phase coherence; and (3) pattern x-ray nanolithography masks. Our approach to improved pattern placement accuracy is based on a new technique, recently invented at MIT, called spatial-phase locking. It is expected that the new SEBL facility will concentrate on sub-100 nm electronic and quantum-effect devices and optoelectronic devices such as DFB lasers and channel dropping filters for wavelength-division multiplexing in optical communication systems.

We have proposed to solve this problem by developing a technology we call spatial-phase-locked electron-beam lithography (SPLEBL), which will provide pattern placement accuracy and precision finer than the resolution. A global fiducial grid is placed on the substrate. The grid is created by holographic lithography, ensuring long-range spatial-phase coherence. It is transparent to the beam, but enables a control computer to keep track of the beam location by means of phase-locking techniques and thus correct for any drift in the system.

In order to demonstrate the efficacy of SPLEBL, x-ray masks for distributed-feedback lasers and channel-dropping filters were fabricated in collaboration with IBM's Thomas J. Watson Research Center. These classes of devices have linewidths of ~115 nm and span many fields of the e-beam lithography system. They require field stitching errors of ~1 nm for good performance. We have achieved stitching errors below 2 nm and have demonstrated sub-nanometer repeatability for spatial-phase locking. Initial results are shown in figures 2 through 4.
Chapter 4. Nanostructures Research

Stitching Error (Phase Discontinuity) at E-Beam Field Boundary

Figure 2. Measurement of the spatial-phase of two overlapped gratings as a function of distance along a direction parallel to the grating lines, as depicted in inset. Set 1 was written using a fiducial grating as a guide. Then the stage was moved and set 2 written, again using the fiducial grating as a guide. Stitching error is below the noise level (i.e., <2 nm).

Repeatability of One-Dimensional Spatial-Phase Locking

Figure 3. Histogram of the phase difference between two successive frame acquisitions on a reference fiducial grating, corrected for stage drift. The standard deviation value of 0.3 nm represents the repeatability or precision of spatial-phase locking in 1-D for our experimental conditions: 8 periods of the 200 nm gold reference grating per viewing window; 100 pA beam current at 50 kV; a p-n diode backscatter detector.
Figure 4. Scanning-electron micrograph of a channel-dropping filter (CDF) grating (spatial period 230 nm) written across four scan fields of the e-beam lithography system. The gratings, which are too fine to be seen directly, are made visible via a moiré between them and the SEM scan raster. The fiducial reference (spatial period = 380 nm) was put in the substrate in advance. The quarter-wave phase step is clearly visible in the CDF grating, but no phase errors are visible at the boundaries of the four scan fields.

4.4 X-Ray Nanolithography

Sponsors

Advanced Research Projects Agency/
Naval Air Systems Command
Contract N00019-92-K-0021
Joint Services Electronics Program
Contract DAAL03-92-C-0001
National Science Foundation
Grant ECS 90-16437

Project Staff

James M. Carter, Nitin Gupta, Scott D. Hector, Gabrielle M. Owen, Dr. Mark L. Schattenburg, Professor Henry I. Smith, Vincent V. Wong, Isabel Y. Yang

For several years, we have been developing the tools and methods of x-ray nanolithography (i.e., sub-100 nm features). We have explored its theoretical and practical limitations, and endeavored to make its various components (e.g., mask making, resists, electroplating, sources, alignment, etc.) reliable and "user friendly". Because of the critical importance of the x-ray mask technology, we discuss this in a separate section (4.5).

Our sources for x-ray nanolithography are simple, low-cost electron-bombardment targets, typically CuL (λ = 1.32 nm), separated by a 1.4 μm-thick SiNx vacuum windows from helium-filled exposure chambers. In the future we hope to replace the CuL sources with higher flux plasma-based x-ray sources.

For most applications that require multiple mask alignment, we currently use a simple microscope-based system which provides about 0.3 μm superposition precision. We are also developing (see section 4.6) a high precision mask alignment system that should provide overlay approaching 1 nm.

In earlier research, we showed that for x-ray wavelengths longer than ~0.8 nm, the range of the photoelectron emitted when an x-ray photon is absorbed in resist does not limit the resolution. Down to feature sizes ~20 nm, diffraction is the major concern. By means of accurate electromagnetic calculations, taking into account the vectorial character of the electromagnetic field and the dielectric properties of the absorber, we have shown that when source spatial coherence is optimized, diffraction does not limit resolution as
severely as had been predicted by simple Fresnel diffraction calculations.

![Proximity Gap vs. Linewidth](image)

Figure 5. Plot of maximum mask-sample gap, G, versus minimum feature size, W, for two values of the parameter α.

Figure 5 plots the maximum mask-to-sample gap, G, versus minimum feature size, W, for two values of the parameter α which connects gap and feature size: $G = \frac{\alpha W^2}{\lambda}$. Modeling and experiment verify that α can be as large as 1.5 while retaining good process latitude.

For the linewidth range from 70 to 20 nm, mask-substrate gaps must be below 5 μm. This is not a problem in research but for manufacturing may be unacceptable. For this reason, we investigated the feasibility of using arrays of zone plates for projection imaging with x-rays of either 4.5 nm or ~1.0 nm wavelength.

### 4.5 Improved Mask Technology for X-Ray Lithography

#### Sponsors

Advanced Research Projects Agency/Naval Air Systems Command
Contract N00019-92-K-0021
Joint Services Electronics Program

Contract DAAL03-92-C-0001
National Science Foundation
Grant ECS 90-16437

#### Project Staff

Martin Burkhardt, James M. Carter, Juan Ferrera, Scott D. Hector, Huiying Li, Mark K. Mondol, Gabrielle M. Owen, Dr. Mark L. Schattenburg, Robert D. Sisson, Professor Henry I. Smith, Vincent V. Wong, Isabel Y. Yang

At feature sizes of 100 nm and below, the mask-to-sample gap, G, must be less than ~10 μm. We have developed a mask configuration compatible with this requirement in which the mask membrane is flat to < 250 nm. We continue to make incremental improvements in this technology aiming at a process that will enable x-ray mask blanks to be fabricated almost entirely with automated IC processing equipment.

Our mask technology is based on low-stress, Si-rich silicon nitride, SiNX. This material is produced in the IC Laboratory at MIT in a vertical LPCVD reactor. The resulting films are clean and uniform, and x-ray mask membranes made from them are extremely robust. They can be cleaned and processed in conventional stations. Radiation hardness remains a problem at dose levels corresponding to production (i.e., millions of exposures) but for research the material is quite suitable.

For absorber patterns we use both gold, Au, and tungsten, W. Both can be obtained with near-zero stress (i.e., < 10^7 MPa) which implies that pattern distortion should be negligible (i.e., < 1 nm). The gold is electroplated onto the membrane after resist exposure and development using a specially designed apparatus. The W is sputter deposited and patterned by reactive-ion etching. During sputter deposition, the plasma environment leads to a nonuniform temperature distribution, which, in turn, causes nonuniform stress. In order to ensure uniform W stress over an entire membrane, a He-backside temperature homogenization apparatus was constructed. Heat is transferred to the membrane from a heated aluminum block separated from the membrane by a 1 mm gap. Results are shown in figure 6. We verify the achievement of stress < 10^7 MPa using a Linnik interferometer equipped with a CCD, a frame grabber and special purpose software that averages out-of-plane deflection over an entire frame.
Chapter 4. Nanostructures Research

Uniform, Zero-stress W via He-backside Temperature Homogenization

Figure 6. Plot of out-of-plane deflection (left ordinate) and stress (right ordinate) versus radial position for 300 nm-thick W, sputter deposited at 200°C onto a SiN, x-ray mask membrane, with and without He backside temperature homogenization. A uniform zero stress is achieved with He, whereas stress is nonuniform without the He.

We have also constructed an apparatus that makes precision measurements of membrane deflection in response to gas pressure differential. This "bulge tester" will enable us to make non-destructive measurements of absorber stress in the low 10^6 MPa range.

Patterning of x-ray masks is done by holographic lithography for periodic structures but, for patterns of arbitrary geometry, it is done by e-beam lithography, either in the MIT e-beam facility or in collaboration with NRL or IBM. We use CAD tools at MIT and convert the data into formats compatible with the e-beam exposure systems. Data is shipped to NRL or IBM by electronic mail. After e-beam exposure, masks are shipped back to MIT by express mail where development and Au electroplating are carried out. This collaboration has already demonstrated that patterning x-ray masks by e-beam can be done remotely, and, by implication, that university researchers with limited facilities can have access to nanolithography via x-ray alone and do not need to own or even visit an e-beam lithography facility. Figure 7 shows an x-ray mask of a coulomb-blockade device in which the finest line-width in the 200 nm-thick electroplated gold is 40 nm.

For patterning x-ray masks with W absorber, a reactive-ion-etching process is required which puts considerable power into the membrane substrate. Since membranes have very low thermal mass and conductivity, we use He-backside cooling in a reactive ion etcher. Membranes can be cooled to below -20°C. At such low temperatures the isotropic etching component is suppressed leading to highly directional etching.

4.6 A High Precision Mask Alignment System

Sponsors
Advanced Research Projects Agency/
Naval Air Systems Command
Contract N00019-92-K-0021
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff
Euclid E. Moon, Professor Henry I. Smith
In order for any lithographic technique to be of general utility, a compatible alignment technique must also be provided. The technique must be capable of superposition precision that is a small fraction of the minimum feature size. Previously, we demonstrated a new scheme for aligning x-ray masks to substrates that was based on the principle of on-axis interference from gratings on the mask and substrate that differ in spatial period (e.g., \( P_1 \) on mask, \( P_2 \) on the substrate). The interference pattern was compared in spatial phase to a fixed fiducial on the mask using a CCD camera and frequency domain image analysis. An alignment standard deviation of 6 nm was achieved. The goal of the present work is threefold: first, to improve the superposition precision by an order of magnitude; second, to fully automate the alignment process; and third, to make the system compatible with industrial x-ray lithography sources and steppers.

In order to improve the superposition precision, a new alignment mark was designed that consists of two adjacent gratings on the mask (\( P_1 \) and \( P_2 \)) facing two complementary gratings on the substrate (\( P_2 \) and \( \pi P_1 \)). The resulting on-axis interference pattern consists of two sets of fringes of the same spatial period which move in opposite directions.
By matching them in spatial phase, one achieves alignment. We believe that in conjunction with a CCD and appropriate image processing, a superposition precision below 1 nm can be achieved. Figure 8 illustrates the basic principle.

Figure 8. Alignment marks containing a \((p_1, p_2)\) pair for X alignment and a \((p_1, p_3)\) pair for Y alignment. (a) Marks for mask and substrate. (b) Depiction of interference pattern that would be obtained when mask and substrate are aligned in both X and Y.

A new X, Y, Z, \(\theta\) stage is being constructed which incorporates high-resolution closed-loop drives and piezos with sub-nanometer repeatability. A wide variety of masks and wafers can be accommodated, including the NIST and MIT mask standards, as well as small substrates and eight inch wafers. An automatic image acquisition and alignment algorithm has been written which compares the spatial-phases of an on-axis interference pattern in about 200 msec. These improvements will be incorporated in a system configuration, shown in figure 9, which is consistent with those under development in industry.
4.7 Optimization of Synchrotron-Based X-ray Lithography

Sponsors
Advanced Research Projects Agency/
    Naval Air Systems Command
    Contract N00019-92-K-0021
National Science Foundation
    Grant ECS 90-16737

Project Staff
Richard J. Aucoin, Scott D. Hector, Dr. Mark L. Schattenburg, Professor Henry I. Smith

There are two approaches to high flux x-ray sources for commercial x-ray lithography: hot, dense plasmas, generated either by lasers or electrical discharges; and synchrotrons. Traditionally, the former have operated around $\lambda = 1.4$ nm and the latter below $\lambda = 1$ nm, with a substantial fraction of the absorbed power below the Si K edge at 0.7 nm. Operation at the shorter wavelengths dictates thicker absorber (~0.6 μm) which becomes a difficult task at 100 nm linewidths. Our analysis indicates that the peak of the synchrotron spectrum absorbed by resist can be shifted to 1 nm or even longer without loss of exposure efficiency. This is illustrated in figure 10. Such a shift would greatly simplify the mask-making task. However, to accomplish this the 25 μm-thick Be vacuum window currently used in synchrotrons must be replaced by a ~2 μm-thick Si window. We have demonstrated that SiN$_x$ membranes this thick can withstand a pressure differential of one atmosphere. We are currently testing various versions of elliptical (2.5 x 55 nm) windows as possible vacuum windows for synchrotrons.
must be used. All such sources have limited temporal coherence, and thus one is forced to employ an achromatic scheme, such as shown in figure 11. In order to make the apparatus more reliable, the depth-of-focus had to be increased. To this end, we introduced a collimating lens and a slit scanning system. We have also incorporated a white-light interferometer that utilizes optical paths through the quartz plates that closely approximate the paths of the exposing beams. A photodetector and lock-in amplifier (figure 12) are used to find the optimum sample position. With the substrate to be exposed in place, the white light interferometer ensures maximum fringe contrast during exposure with the 193 nm deep-UV radiation. Using these improvements, we can reliably expose 100 nm-period gratings over areas of more than 2 cm². We plan to use this system to fabricate 100 nm-period x-ray masks and free-standing gratings for He atom interferometry. We also plan to take the technique one step further to 50 nm periods (25 nm lines and spaces) using a 13 nm undulator as a source.

4.9 Ion Beam Lithography

Sponsor
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff
Scott E. Silverman, Professor Henry I. Smith, Anto Yasaka

Focused-ion-beam (FIB) systems are potentially useful as nanolithography tools because of the near absence of backscattering, which is a significant problem in electron-beam lithography, especially on high-atomic-number substrates. We currently work with Be⁺ ions and have demonstrated 50 nm lines and spaces. The FIB system is being modified to accept an improved Be alloy source. We will explore the feasibility of combining spatial-phase-locking with FIB lithography. This could prove to be the superior technology for making x-ray nanolithography masks based on W absorbers.
Chapter 4. Nanostructures Research

**Figure 11.** Achromatic-holographic lithography (AHL) configuration in which the separation between the substrate and the second phase grating is controlled by a piezoelectric micrometer.

\[
sin \theta = \frac{\lambda}{nP},
\]
\[
\Delta = \frac{\lambda}{2n \sin \theta}
\]
\[
\Delta = \frac{P}{2}
\]

**Detection Scheme for White-light Interferometer**

**Figure 12.** Schematic of white-light interferometer used to match optical path lengths in the AHL configuration. A collimated white-light beam is split into two paths and recombined to interfere at a photodiode. An interference peak, detected via a lock in amplifier, indicates a matching of path lengths and hence maximum image contrast in the AHL configuration.
4.10 High Performance Self-aligned Sub-100 nm MOSFETs Using X-ray Lithography

We have fabricated sub-0.1 \( \mu m \) N-channel MOSFET devices using x-ray lithography. Extremely well-controlled short-channel effects were achieved through appropriate channel and source/drain engineering.

Figure 13 shows the device schematic. The retrograde channel doping profile was accomplished using shallow indium and deep boron implants. The source/drain halo extensions were formed using a low energy arsenic implant with indium pre-amorphization and counter-doping. The self-aligned polysilicon gates were fabricated using x-ray nanolithography and an anisotropic etching process. The aligned, microgap x-ray exposures were carried out using mesa-etched SiNx x-ray masks with Au absorber pattern. The original (mother) mask patterns were written at the NRL by e-beam lithog-

---

2 Naval Research Laboratory, Washington, DC.
Chapter 4. Nanostructures Research

raphy, and then several duplicates (daughters) were made using x-ray lithography. We have characterized the chemically amplified resist SAL-601 and developed a process which allowed us to achieve the minimum linewidths with control and a large process latitude. This involved optimization of the post exposure bake time and temperature. We found that with a post-exposure temperature of 110°C and a post-exposure baking time of 40 seconds, minimum linewidths of approximately 100 nm were obtained. To eliminate "skin" formation on top of the SAL-601, a thin (200-250 nm) layer of polyvinyl alcohol (PVA, which is water soluble) was spun on top of the resist. The resist served as a mask for etching a thin LTO layer which, in turn, acted as a hard mask for the selective anisotropic etching of polysilicon using C12 plasma. Figure 14 shows the self-aligned SAL-601/LTO. Figure 15 shows a polysilicon gate defined by x-ray lithography and C12 plasma etching.

Figure 16 shows the I-V characteristics of a sub-0.1μm nMOSFET with L₉₉ of 850 nm and 5.3 nm gate oxide. This device has a current drive of 0.74 nA/μm and a saturated transconductance of 500 nS/mm at 2.0 V power supply with a threshold voltage of 0.36 V and a sub-threshold slope of 88.1 mV/dec. This indicates that these 0.1μm MOSFET devices fabricated by x-ray lithography exceed the performance of the current state-of-the-art 0.1μm MOSFET devices produced by e-beam lithography.

4.11 Fabrication of T-gate Devices Using X-ray Lithography

Sponsor
Advanced Research Projects Agency/
Naval Air Systems Command
Contract N00019-92-K-0021

Project Staff
James M. Carter, Nitin Gupta, Professor Henry I. Smith

Monolithic microwave integrated circuits (MMICs) have potential applications in automobile navigation, collision-avoidance, and personal wireless communication systems. High-speed MODFET devices require very short gate lengths, while preserving low resistance. Large gate widths are required for high current drive. To meet these conflicting demands, researchers have developed so-called "T-gate" and "gamma-gate" processes in which the base of the gate is very short (~100 nm) while the upper part is large and overlaps the short base, similar to a mushroom, or the letters T or G. Such structures are readily achieved using direct-write electron-beam lithography (EBL). However, this technology is expensive, slow, and unlikely to meet future production volume needs. For these reasons, we developed a process for fabricating T-gates using x-ray lithography. Initial results were obtained using tri-layer PMMA. The three layers consisted of a bottom one of 950K PMMA, a middle layer of PMMA/MAA (the co-polymer of PMMA), and an upper layer of 496K PMMA. These three resists have different x-ray sensitivities. Experimental results are shown in figures 17 and 18.

Figure 14. Patterning for MOSFET gate with x-ray lithography exposure of chemically-amplified resist SAL-601 and RIE of low-temperature oxide using CHF₃.
Chapter 4. Nanostructures Research

Only one alignment, one exposure, and one development are required. However, the process latitude is narrow, and the process may not be suitable for manufacturing.

In order to enhance the process latitude, we are pursuing a bi-layer process that utilizes high-sensitivity chemically amplified resists. This process has the flexibility that arises from two separate alignments, exposures, and developments.

There are a number of interesting directions that may be followed once a high latitude process is established. For manufacturable MMIC systems, both MESFETs and HEMTs are required for low-noise and power applications. Initially, relatively simple GaAs MESFETs will be studied, followed by GaAs and InP HEMTs of varying degrees of complexity. Further studies could include low-temperature-grown GaAs MESFETs for high breakdown voltages, self-aligned devices, or gate materials other than Au, such as W.
Figure 17. Scanning electron micrograph showing the resist profile cross-section for a 120 nm T-gate.
Chapter 4. Nanostructures Research

Figure 18. Scanning electron micrograph of a 250 nm T-gate (Ti/Au) after lift-off:

4.12 Studies of Coulomb Charging Effects and Tunneling in Semiconductor Nanostructures

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001
U.S. Air Force - Office of Scientific Research
Grant F-49-62-92-J-0064

Project Staff
Professor Dimitri A. Antoniadis, Martin Burkhardt, David J. Carter, Arvind Kumar, Professor Michael R. Melloch, Professor Terry P. Orlando, Professor Henry I. Smith

Quantum-effect devices, whose minimum feature sizes are comparable to the Fermi wavelength (about 50 nm in a typical inversion layer), have promising potential in novel electronics applications. Quantum-dot devices have drawn particular attention. In such devices an electron gas is confined electrostatically in all three dimensions, forming a

---

76 RLE Progress Report Number 136
A small "island" of electrons ~100 nm, bounded on all sides by potential walls. This small electron "island" resembles an atom in that there can be only an integer number of electrons, and these electrons can occupy only certain discrete energy levels. The conductance of the dot, when connected to leads through tunneling barriers, exhibits strong oscillations as the voltage of the gate is varied. Each successive conductance maximum corresponds to the discrete addition of a single electron to the dot. At temperature in the mK range, the conductance decreases by orders of magnitude in between adjacent conductance maxima because there is a large energy cost for an electron in the lead to enter the dot. This energy cost can be removed by changing the gate voltage, resulting in the observed periodic dependence of the conductance on gate voltage.

A variety of tunneling devices were fabricated using x-ray nanolithography and measured at cryogenic temperatures. The devices were fabricated on modulation doped AlGaAs/GaAs substrates. After mesa etch and deposition of ohmic contacts, x-ray nanolithography was used to put down the gate patterns. A typical fabrication sequence for the gate level of a device design is shown in figure 19. After the fabrication of the mother mask using e-beam lithography and electroplating, a daughter mask is made using x-ray lithography at a gap between mother and daughter mask of about 1.5 μm (figure 19b). As an x-ray source we use an electron bombarded copper target emitting CuKα x-rays (λ = 1.3 nm). The daughter mask is then used to expose device patterns on appropriate semiconductor substrates, as illustrated. For the gates, aluminum was evaporated to a thickness of 120 nm and lifted off. The ability of x-ray lithography to expose vertical profiles in thick resists ensures easy liftoff and simplifies device processing because only one gate metalization is needed.

A novel type of quantum dot device, featuring control over the shape of the quantum dot, was also fabricated using x-ray nanolithography. The device consists of four consecutive quantum point contacts which can be controlled independently. A quantum dot is produced by biasing two pairs of quantum point contacts below the first subband. This allows only tunneling into the out of the dot to take place. The remaining gates can then be used to sweep the electrochemical potential within the dot. The size of the quantum dot can be adjusted up to a size of 600 nm by biasing the two outermost quantum point contacts below the first subband. A micrograph of the fabricated device is shown in figure 20. The device was measured in an Oxford Instruments 300 mK probe. Strong Coulomb blockade oscillations were observed for several Coulomb dot sizes at a temperature of 300 mK. Figure 21 shows a typical plot of conductance versus gate voltage for a quantum dot defined by the two outermost pairs of quantum point contacts.

![Fabrication sequence of tunnel devices using x-ray nanolithography](image)

**Figure 19.** Fabrication sequence of tunnel devices using x-ray nanolithography: (a) mother mask; (b) daughter mask; (c) finished device.
Figure 20. SEM micrograph of the variable sized quantum dot.

Figure 21. Plot of conductance versus gate voltage for a quantum dot defined by the two outermost pairs of quantum point contacts.
4.13 Dual Electron Waveguide Device Fabricated Using X-ray Lithography

Sponsors

Advanced Research Projects Agency/
Naval Air Systems Command
Contract N00019-92-K-0021

National Science Foundation
Grant DMR 87-19217
Grant DMR 90-22933

Project Staff

Professor Jesús A. del Alamo, William Chu,
Christopher C. Eugster, Professor Henry I. Smith

Dual-electron-waveguide devices were fabricated using x-ray nanolithography. An electron waveguide is essentially a one-dimensional channel in which electrons travel without scattering. In our device, two such channels are electrostatically formed in close proximity to one another by depleting those electrons in an AlGaAs/GaAs modulation-doped heterostructure which reside underneath the gates. The x-ray mask used in the fabrication process was patterned using 50 keV e-beam lithography at the Naval Research Laboratory. After replicating this mask to change its polarity (using x-ray lithography), we align the daughter mask to the AlGaAs/GaAs sample and expose using x-ray nanolithography.

In an electron waveguide, much like in an optical waveguide, discrete transverse modes arise due to lateral confinement. The conductance of each waveguide mode is equal to a fundamental constant $2e^2/h$. This results from the cancellation of the energy dependence in the product of the 1D density of states and the electron velocity. Figures 22 and 23 show the results of testing the dual electron waveguides at 1.6K.

4.14 Novel Mesoscopic Superconducting Devices

Sponsor

Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff

James M. Carter, Jay N. Damask, Juan Ferrera,
Professor Hermann A. Haus, Professor Leslie A.
Kolodziejski, Euclid E. Moon, Professor Henry I.
Smith, Vincent V. Wong

4.15 Channel-Dropping Filters Fabricated Using X-ray Lithography

Sponsors

Joint Services Electronics Program
Contract DAAL03-92-C-0001
U.S. Army Research Office
Grant DAAL93-92-G-0291

Project Staff

Professor Dimitri A. Antoniadis, David J. Carter,
James J. Hugunin, Professor Terry P. Orlando, Professor Henry I. Smith

Superconducting integrated circuits have demonstrated fast switching speeds, low power dissipation, and low-loss wiring to distant devices. One of the hurdles between these devices and a practical technology is the lack of a useful three-terminal device. We are currently addressing that problem with the examination of several forms of superconducting FETs on both Si/SiO$_2$ and GaAs/AlGaAs. By exploiting the quantum interference effects found in very short channel devices (<100 nm), we believe it will be possible to produce three-terminal superconducting devices with gains greater than unity.

In addition to the promise of a practical technology, these devices can be used to investigate several interesting questions in mesoscopic physics. In particular, we are interested in the interaction between the macroscopic quantum states in the superconductor and the mesoscopic quantum states available in traditional quantum-effect devices. By exploiting our experience with quantum dots and quantum point contacts on both Si/SiO$_2$ and GaAs/AlGaAs, we hope to study the unique effects predicted for these structures with superconducting contacts. These include quantization of the superconducting critical current, supercurrent carried by resonance states, and others.
Figure 22. Plot of conductance versus split-gate voltage (proportional to waveguide width) for each of two waveguides, as illustrated. Note that conductance is approximately quantized in steps of $2e^2/h$. 

Dual Electron-Waveguide Device

$W = 0.3 \, \mu m$

$T = 1.6 \, K$

$L = 0.15 \, \mu m$

$V_{GS}$ (V)
Chapter 4. Nanostructures Research

Family of Quantized Conductance Steps, Revealing Location of Impurity

Figure 23. Plot of conductance versus bottom gate voltage for several values of the middle gate bias, which has the effect of shifting the location of the conducting channel. Perturbations of the conductance steps reveal the location of an impurity in the gap between gates.
Channel-dropping filters (CDFs) are novel optoelectronic devices which are promising candidates for future wavelength-division-multiplexed optical communication systems. Figure 24 shows a schematic of a CDF. The device operation relies on the frequency-selective nature of the quarterwave-shifted (QWS) grating structure, commonly used to improve longitudinal mode control in distributed feedback (DFB) lasers. Such QWS-DFB structures can interact with optical signals in such a way as to selectively transmit, reflect or detect the signals (i.e., wavelengths) that are resonant with the structure. We work with two material systems for the optical waveguides: InGaAsP/InP and silica-on-silicon. For such material systems the necessary grating lengths can span over 1 mm and the grating periods are ~240 nm and ~500 nm for InGaAsP/InP and silica-on-silicon, respectively. Furthermore, the grating lines must be spatially-coherent to within $\lambda/150$ over the entire span to reduce linewidth broadening and spectral offsets from the desired resonance, and to avoid spurious responses in the filter spectrum. Recently, we have developed a new technique called spatial-phase-locked e-beam lithography (SPLEBL) which has demonstrated an interfield phase-locking precision standard deviation of $\sigma = 0.35$ nm. Figure 4 of section 4.3 shows an e-beam-written QWS-DFB grating with a period of $\Lambda = 230$ nm. The continuity of the moiré pattern across the field boundaries indicates no discernible stitching error, while the $\pi$-phase shift at the QWS boundary (field 3) indicates a true $\lambda/4$ phase step. Using SPLEBL x-ray masks have been written with several QWS-DFB gratings of various lengths and periods.

**Figure 24.** Schematic of a quarter-wave-shifted channel-dropping filter. The gratings, which span several fields of an e-beam lithography system, must be precisely tuned and spatially coherent in order to selectively tap off the narrow channel $f_2$.

We are also utilizing a novel on-axis interferometric alignment scheme to obtain sub-0.1 $\mu$m alignment of the QWS-DFB grating (x-ray level) to the rib waveguides (optical level). Such accuracy is necessary to minimize the resonant effect of the QWS-DFB grating on the unloaded waveguide, or optical bus. X-ray lithography and reactive-ion-etching in $\text{CHF}_3$ is then used to transfer the QWS-DFB gratings into the rib waveguide.
4.16 Ridge-Grating Distributed-Feedback Lasers Fabricated by X-ray Lithography

Distributed-feedback lasers are essential components in optical communications systems because they operate in a single-longitudinal-mode and are easily integrable with electronic drive circuitry. In the fabrication of a typical DFB laser, an epitaxial regrowth step is carried out on top of the grating, which can be yield-limiting. In particular, because of etch back the coupling constant after epi-growth can be difficult to predict. We are developing a novel DFB laser structure in which gratings are etched on either side of a ridge waveguide, as shown in figure 25. This structure eliminates the need for regrowth and maximizes the utility of ridge-waveguide devices, which are relatively simple to fabricate. The structure also decouples materials growth, waveguide fabrication and grating fabrication, which should increase device yield.

Figure 25. Schematic of novel DFB laser structure.

To realize such a device structure, we have developed a fabrication process that utilizes x-ray lithography for the grating fabrication because it offers good process latitude, high throughput, long-range spatial-phase fidelity in the gratings, and large depth-of-focus. After the epitaxial growth of a complete InGaAlAs graded-index, separate-confinement, multi-quantum-well-layer diode structure, by molecular-beam epitaxy on an InP substrate, a 3-4 μm-wide ridge waveguide is formed by CH₄/H₂ reactive-ion-etching (RIE). A first-order grating with a period of 230 nm for the target lasing wavelength of 1.55 μm is then patterned in PMMA using x-ray nanolithography, as shown in figure 26. The x-ray mask is patterned using holographic lithography, which guarantees excellent uniformity and long-
range spatial-phase coherence in the grating, followed by gold electroplating.

Figure 26. Pattern for DFB laser grating exposed using x-ray lithography (λ = 230 nm). Ridge height is 1.0 μm.

Reactive-ion etching (RIE) in CHF₃ is used to transfer the grating pattern from PMMA into an oxide hard mask, and subsequent RIE in CH₂/H₂ is used to transfer the grating into both sides of the ridge waveguide. Figure 27 shows an etched InP grating of 250 nm depth next to a ridge waveguide of 1 μm height. As shown, the etched gratings are immediately adjacent to the ridge sidewall, which maximizes the coupling. Figure 3b shows the same InP grating in cross-section. Optical feedback occurs via interaction of the lateral fields with the grating. Holographic lithography could not be used directly for the grating exposure because of troublesome coherent reflections from the ridge topography.
Figure 27. DFB laser grating etched into either side of ridge waveguide in InP using CH$_4$/H$_2$ RIE. ($\lambda = 230$ nm. Etch depth = 250 nm).

For advanced optical communication systems (e.g., those implementing wavelength-division multiplexing) quarter-wave-shifted DFB lasers, hundreds of microns in length, are necessary. This will require x-ray masks made by e-beam lithography that are free of distortion and stitching errors. As described in section 4.3 we are pursuing a new technique called spatial-phase-locked e-beam lithography that we believe will solve the problems of distortion and stitching errors.

4.17 Fabrication of Sub-micron MSM Photodiodes by X-ray Lithography

Sponsor
Advanced Research Projects Agency/
Naval Air Systems Command
Contract N00019-92-K-0021

Project Staff
Professor Jesús A. del Alamo, Akbar A. Moolji, Professor Henry I. Smith

Over the last few years, planar InGaAs-InP-based metal-semiconductor-metal (MSM) photodetectors have attracted considerable interest. Among other advantages, their ease of fabrication, high speeds and their compatibility with field-effect transistors are key features that make these devices good candidates for long wavelength, integrated optical communication systems.

Figure 28. Schematic cross section of a standard MSM photodetector.

Figure 28 is a schematic cross section of the generic MSM photodetector. Work done in the past on these devices has shown this particular combination of layers to be the one that optimizes the operation of the detectors for long wavelengths. The electrode fingers (marked alternately + and -) act as collectors for carriers that are generated when infra-red radiation impinges on the active areas between them. The barrier enhancement layer (50 nm InAlAs), directly below these fingers, serves to enhance the Schottky barrier height of the contact. Since InGaAs itself fails to form a good Schottky contact, the insertion of this thin, high bandgap layer is a necessity that helps suppress the deleterious dark current in these devices. The
300 nm InAlAs buffer layer that lies between the active InGaAs and the semi-insulating InP layers, is the most recent addition to the MSM structure. Its presence has been shown to reduce the parasitic capacitance of the structure over the 100 MHz - 1 GHz range, hence improving the RC-limited bandwidth of the device substantially.

The main bottleneck for the detector bandwidth, however, arises from the transit time of the carriers in the region between the electrode fingers. It is natural to expect that this limitation can be overcome by reducing finger spacing. Detectors with finger spacings down to 0.5 μm have already been reported in the literature. Our efforts are now geared towards pushing this limit even further. The current work is an extension of an earlier effort, made about two years ago at MIT, which yielded good working devices with a bandwidth of about 6 GHz. The smallest finger spacing that was achieved then was about 1 μm. Using otherwise the same fabrication process, we will now pattern the fingers by x-ray nanolithography. This will allow us to scale the finger spacings down to almost 50 nm. Furthermore, the broad process latitude that x-ray nanolithography provides should result in a higher yield of good working devices. More importantly, from a long-range perspective, the large depth-of-focus associated with this technology will permit us to simultaneously fabricate sub-micron size gates for FETs that may then be integrated with these detectors.

As long as the device parasitics are not allowed to dominate, and the detectors can operate in the transit-time limited mode, our efforts should result in structures with greatly improved frequency responses.

4.18 Submicrometer-Period Transmission Gratings for X-ray and Atom-Beam Spectroscopy and Interferometry

Sponsor
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff
James M. Carter, Julie C. Lew, Jeanne M. Porter, Dr. Mark L. Schattenburg, Satyen Shah, Professor Henry I. Smith

4.19 High-Dispersion, High Efficiency Transmission Gratings for Astrophysical X-ray Spectroscopy

Sponsor
National Aeronautics and Space Administration
Contract NAS8-36748

Project Staff
Richard J. Aucoin, Professor Claude R. Canizares, Robert C. Fleming, Jeanne Porter, Dr. Mark L. Schattenburg, Professor Henry I. Smith
Figure 29. Schematic of the free-standing SiN, gratings of 200 nm period. In order to verify that these gratings are free of distortion, chromium atoms were evaporated through the free-standing grating onto a holographically exposed grating of the same period. The resulting moiré fringes reveal the absence of any systematic distortion.
Chapter 4. Nanostructures Research

Through a collaboration between the Center for Space Research and the NanoStructures Laboratory (NSL), transmission gratings are provided for the Advanced X-ray Astrophysics Facility (AXAF) x-ray telescope, currently scheduled for launch in 1998. Many hundreds of low-distortion, large-area, gold transmission gratings of 200 nm period and 400 nm period are required. These will provide high resolution x-ray spectroscopy of astrophysical sources in the 100 eV to 10 keV band.

Because of the requirements of low distortion, high yield, and manufacturability, a fabrication procedure involving holographic lithography has been selected. In order to ensure spatial-period fidelity, master reference gratings are used to periodically recalibrate the holography system. The grating patterns are transferred into the substrate using a tri-level resist scheme and reactive-ion etching, followed by gold electroplating. An etching step then yields membrane-supported gratings suitable for space use. Flight prototype gratings have been fabricated and continue to undergo space-worthiness tests. Progress in this area focuses on increasing the yield and flexibility of the fabrication procedures and perfection of grating evaluation tests.

**4.20 GaAs Epitaxy on Sawtooth-Patterned Silicon**

**Sponsor**

Joint Services Electronics Program
Contract DAAL03-92-C-0001

**Project Staff**

Sean M. Donovan, Professor Khalid Ismail, Professor Leslie Kolodziejski, Professor Henry I. Smith, Professor Carl V. Thompson

The growth of GaAs on Si offers the possibility of combining high-speed and optoelectronic GaAs devices with Si integrated-circuit technology. Ordinarily, the 4.1 percent mismatch between the two materials leads to high dislocation densities. However, it has been shown here that when GaAs is grown on sawtooth-patterned Si substrates, the dislocation density is less than \(10^5/\text{cm}^2\), orders of magnitude lower than in GaAs films grown on planar Si substrates. We investigate this effect both for its potential of greatly improving the quality of GaAs on Si and as a model for understanding the mechanism of dislocation reduction. Oriented gratings of 200 nm period are fabricated in Si$_3$N$_4$ on (100) Si substrates using holographic lithography. Anisotropic etching in KOH is then used to produce sawtooth-profile gratings in the Si. Then these serve as substrates for GaAs growth by gas source MBE.

**4.21 Publications**

**Journal Articles**


Li, H., M. Mondol, G. Owen, and H.I. Smith. "Uniform, Zero-Stress W via He-Backside Tem-

---

4 Cairo University.
Chapter 4. Nanostructures Research

perature Stabilization." Submitted to J. Vac. Sci. Technol. B.


Meeting Papers


Chapter 4. Nanostructures Research

at the Symposium on VLSI Technology, Honolulu, Hawaii, June 7-9, 1994.


Theses


MIT Patents

Chapter 5. Single-Electron Spectroscopy

Academic and Research Staff
Professor Raymond C. Ashoori

Graduate Students
David Berman, Mikhail G. Brodsky, Ho-Bun Chan, Sven Heemeyer

5.1 Project Description

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001
National Science Foundation
Young Investigator Award

Two years ago, we developed single-electron capacitance spectroscopy (SECS). This is a means for measuring the quantum energy levels of nanoscale objects such as quantum dots, single impurity atoms, and localized electronic states in a semiconductor. The resolution of the method is limited only by the sample temperature, and objects containing as few as one electron may be probed. We intend to develop this type of single-electron spectroscopy into a general tool for studying nanostructured materials. Since my arrival at MIT in 1993, the focus of our work has been largely in the setup of our low temperature laboratory. Here I will discuss the SECS technique and our use of it to study a single quantum dot or artificial atom.

Our artificial atoms are semiconductor structures so small that they can contain as few as one electron. These can be considered as small boxes containing a number of electrons which can be varied. Like real atoms, electrons are attracted to a central location. In a real atom, this central location is a positively charged nucleus; in an artificial atom, electrons are typically trapped in a bowl-like parabolic potential well in which electrons tend to fall in toward the center of the bowl. One can consider the artificial atom as a tiny laboratory in which quantum mechanics and the effects of the electron-electron interaction can be studied. Because of its large size, the artificial atom is in a different physical regime from real atoms; as such we can expect a lot of new physics in the electronic energy level spectra of artificial atoms.

Atomic spectra of real atoms of a particular species are identical. This is not true of artificial atoms. It is not yet possible to manufacture many identical artificial atoms. For this reason, we have developed SECS, a method for observing spectra in a single artificial atom.

Figure 1a shows the new scheme. The artificial atom is placed between the plates of a tunnel capacitor. It is close enough to one of the plates so that single electrons can quantum mechanically tunnel between the artificial atom and the nearby plate. The artificial atom is far enough from the other capacitor plate to prohibit any tunneling to this plate occurring. Electric fields can be created by applying a voltage between the plates of the capacitor. If the top plate is made positive compared to the bottom one, electrons from the bottom plate will be attracted in the direction of the top plate, toward the artificial atom. Single electrons can thus be coaxed to tunnel into the artificial atom. Similarly, they can be expelled from the artificial atom using negative top plate voltages.

The motion of the single electrons can be detected using a simple physical principle. When a single electron tunnels into the artificial atom, it moves closer to the top plate of the capacitor. Electrons in the top plate tend to be pushed away from the plate; i.e., some charge is induced on the top plate. In our samples, the amount of charge induced is about half of an electron's charge.

The induced charge on the capacitor plates gives rise to a detectable voltage across the tunnel capacitor. The voltage measured by a detector will be the amount of the charge divided by the input capacitance of the detector. Given that our artificial atom samples are in a cryostat (typically at 0.3 K), wires leading to a room temperature detector will have on order 1000 pF of shunt capacitance. This large shunt capacitance would make measurement of the induced charge impossible.

In order to greatly reduce the shunt capacitance, we build an on-chip capacitance bridge with a high electron mobility transistor (HEMT) detector located about 1 mm away from the tunnel capacitor. In this arrangement, the tunneling of a single electron induces a voltage of about 10 nV at the input of the HEMT. The artificial atom is contained in a GaAs/AlGaAs quantum well produced by molecular
Figure 1. (a) A scheme for making precision measurements on an artificial atom. The artificial atom may be thought of as a small bowl in which electrons can be collected one at a time. It is positioned between two plates of a capacitor. Single electrons can be induced to tunnel into the artificial atom by application of a voltage on the top plate. Detectable charge is induced on the top plate as a result of this tunneling. (b) Schematic cross section through the structure containing the artificial atom.

beam epitaxy. Figure 1b is a schematic diagram. The bottom n⁺ GaAs layer serves as the bottom tunnel capacitor plate. A thin (150 Angstrom) AlₓGaᵧAs acts as the tunnel barrier. The thick AlₓGaᵧAs layer above the GaAs electrically insulates the electrons in the artificial atom from the top Cr electrode. Patterning, by electron beam lithography and reactive ion etching, of the top layer of the structure provides lateral confinement of electrons in the quantum well.

Figure 2 displays the output of the HEMT, corresponding to the capacitance of the small capacitor, as the voltage on the capacitor plates is swept. Notice the peaks. The first of these peaks, moving from left to right, is caused by the first electron to move into the previously empty artificial atom. Electrons can then be counted one by one as they move into the artificial atom. Shown in figure 2 are signals from the first 25 electrons to enter the artificial atom. An obvious question is: what is the meaning of the spacings between the peaks? It turns out that the horizontal scale in figure 2 can actually be read as an energy scale, with energy increasing to the right. It takes a certain amount of energy to add each subsequent electron to the arti-
Figure 2. The results of an ultrasensitive capacitance vs. top plate voltage measurement on the sample pictured in figure 1. Moving from left to right, the first peak corresponds to the first electron entering the previously empty artificial atom. Each successive peak indicates a subsequent electron admission. The spacings between the peaks indicate the energy required to add a successive electron to the artificial atom, and the widths of the peaks are determined by the sample temperature (0.35 Kelvins).

Why does it take more energy to add each subsequent electron to the atom? There are two reasons. One is that the electrons already in the atom repel subsequent electrons from being admitted; it takes a certain amount of energy to overcome this repulsion. Additionally, newly admitted electrons cannot enter quantum states that are already occupied (Pauli Principle); they must enter higher quantum energy levels.

If a magnetic field is turned on, the energies of the quantum energy levels change. In the simplest sense, this happens because the magnetic field places a force on moving electrons. The orbits of electrons in the artificial atom are thus changed by the magnetic field, causing the peaks seen in figure 2 to move. This ability to follow quantum energy levels in magnetic field is a powerful probe of the way electrons behave within the atom. To observe it better, plots such as the one in figure 3 are created. It is a compendium of many data sets such as the one shown in figure 2, now taken at varying values of the magnetic field strength. In figure 3, the horizontal axis represents magnetic field strength and the vertical axis represents voltage across the capacitor (or energy). The capacitance is plotted on a gray-scale, with white representing a capacitance peak.

Let us first examine the bottom trace of figure 3. It plots the energy of the first electron as a function of magnetic field. Notice that this energy increases as the magnetic field strength is increased. This behavior of the one electron energy turns out to be easily predictable. A magnetic field will tend to make a moving electron move in circles instead of in a straight line. The sizes of these circles become smaller as the magnetic field strength is increased. Effectively, the magnetic field confines the electrons. According to quantum theory, the more tightly
one confines an electron, the higher its energy. A fit to this trace determines that the diameter of the first electron’s wavefunction is about 400 Angstroms.

The next trace seen in figure 3 is for two electrons in the artificial atom. Notice that this trace appears qualitatively different from that of the first electron. Rather than smoothly moving up in energy, the two-electron energy shows a very clear kink at a magnetic field strength of 1.5 Tesla.

Why does the energy to add the second electron energy differ so much from the one electron energy? The reason is that the two electrons in the artificial atom interact with one another. Indeed this interaction produces the kink at 1.5 Tesla. Aside from carrying electric charge, electrons also carry magnetism. If electrons are thought of as small bar magnets, then the spin of an electron points between the poles of the magnet. At zero magnetic field, the two electrons in their lowest energy configuration have their spins pointing in opposite directions. At 1.5 Tesla, the spin of one of the electrons flips, so that the magnet moments of both electrons line up with the external magnetic field. This spin-flipping has never been observed before.
tron, this spin-flipping is predicted to take place at the astronomic magnetic field of 400,000 Tesla! It turns out that the large size of the artificial atom brings the magnetic field required for the spin-flipping into the observable range.

At higher electron numbers, the physics of the artificial atom becomes difficult to solve and sophisticated computer modeling is required to understand the energy levels. For low magnetic fields, an model of electronic "shell" structure analogous to the s, p, d, and f shells in real atoms is appropriate. However, for more than about 10 electrons in the artificial atom and at magnetic fields above a few Tesla, features appear in the data that are understandable in terms of a magnetic shell structure in which confinement by the applied magnetic field dominates the confinement of the artificial atom's parabolic potential.

With the creation of the few electron artificial atom, the ultimate limit of small sized electronics is being achieved. There remains much physics to explore, with a vast amount of information in the details of the spectra. Basic ideas about the effects of interactions between electrons can be tested in an unprecedented way.

Several projects are presently underway to enhance both the utility and sensitivity of SECS. We have commissioned a top loading dilution refrigerator which will allow samples to be cooled from room temperature to 20 mK in about three hours. This will allow us to measure a large quantity and variety of samples. Also, a scanning electron microscope has been converted for use in electron beam lithography. Initial experiments will focus on spectra of artificial atoms and spectra of individual silicon impurity atoms in GaAs. Finally, we expect that the use of single-electron transistors (see chapter 2 of this section) as charge sensors may significantly enhance the sensitivity of our SECS measurement. We are commencing work on integrating metallic single-electron transistors onto our artificial atom samples.

### 5.2 Publications


Section 3  Optics and Devices

Chapter 1  Optics and Quantum Electronics

Chapter 2  Optical Propagation and Communication

Chapter 3  High-Frequency (> 100 GHz) and High-Speed (< 1 ps) Electronic Devices
Chapter 1. Optics and Quantum Electronics

Academic and Research Staff
Professor Hermann A. Haus, Professor Erich P. Ippen, Professor James G. Fujimoto, Professor Peter L. Hagelstein, Professor Louis D. Smullin, Dr. Santanu Basu Dr. Brett E. Bouma, Dr. Joseph A. Izatt

Visiting Scientists and Research Affiliates
Dr. Lucio H. Acilioli, Dr. Artur D. Gouveia-Neto, Dr. Franz X. Kärntner, Dr. Charles P. Lin, Dr. Jerôme M. Paye, Dr. Matasaka Shirasaki, Yuzo Hirayama

Graduate Students

Undergraduate Students
Fernando A. Broner, Darlene J. Ford, Mohammed J. Khan, Charles Yu

Technical and Support Staff
Mary C. Aldridge, Donna L. Gale, Cynthia Y. Kopf, Deborah S. Manning

1.1 Additive Pulse Mode-locking

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001
U.S. Air Force - Office of Scientific Research
Contract F49620-91-C-0091

Project Staff
Professor James G. Fujimoto, Professor Hermann A. Haus, Professor Erich P. Ippen, Dr. Jerôme M. Paye, Farzana I. Khatri, Gadi Lenz, John D. Moores, Lynn E. Nelson

Additive pulse modelocking (APM) is a technique of modelocking that simulates fast saturable absorber action by interferometrically transforming phase modulation caused by the Kerr effect into amplitude modulation. In its first realization with coupled cavities, where the gain was supplied by an F-center crystal, this technique was called the "soliton laser." When it was found that the fiber in the auxiliary resonator providing the phase modulation did not have to possess negative dispersion, we discovered the principle of APM.

The analytic theory of APM\(^1\) has been developed under the assumption that a pulse experiences only small changes per pass through the laser system due to each of the phenomena participating in the pulse shaping process. This assumption covers a wide range of operating regimes, since large changes per pass tend to lead to undesirable spectral characteristics of the pulse and/or to pulse instability. Yet, there are at least two conditions under which the change per pass can be large without attendant pulse deterioration:

1. The case of large positive and negative linear dispersion in the resonator, the pulse changing in width and height due to the dispersion, the net change in one roundtrip remaining small.

2. The other condition may occur in coupled cavity systems, with large pulse changes in the nonlinear "auxiliary" resonator, yet small changes in the pulse in the main resonator, due to the fact that the pulse in the auxiliary resonator

Chapter 1. Optics and Quantum Electronics

The resonator is of relatively small energy compared with the pulse in the main resonator.

Condition (1) is realized in some fiber lasers and is described in section 1.2. An analytic theory has been developed for this case which agrees with the experimental observations.2 Contrary to the usual sech-like shape of the pulses observed in most passively modelocked systems, the pulses are approximately gaussian, a fact verified experimentally.

Recent experiments3 showed a very stable performance with large phase shifts and large spectral broadening of the pulse in the auxiliary resonator. The spectrum of the pulse in the main resonator is quite clean and transform limited. This is condition (2) described above. We have developed an analytic theory for this "overdriven case" coupled with computer simulations of this mode of operation.4 This approach is made possible by the recognition that the greatly broadened spectrum of the pulse fed back from the auxiliary resonator is filtered out by the regenerative gain of the main resonator. This makes possible an approximate analytic treatment of the action of the Kerr medium (fiber). Because the system performs particularly well in this regime, an understanding of its operation helps us apply the principle to other coupled cavity systems.

1.2 Fiber Ring Laser

Sponsors

Charles S. Draper Laboratories
Contract DL-H-441692
Joint Services Electronics Program
Contract DAAL03-92-C-0001
U.S. Air Force - Office of Scientific Research
Contract F49620-91-C-0091

Project Staff

Professor Hermann A. Haus, Professor Erich P. Ippen, Christopher R. Doerr, Lynn E. Nelson, Kohichi R. Tamura

Significant progress has been made in the modelocking of fiber ring lasers. Last year pulses 452 fs long were reported. These pulses were passively modelocked using the polarization-APM principle, the interference between two polarizations causing the phase-to-amplitude conversion. This past year, graduate student Kohichi Tamura was able to generate 75 fs pulses from a fiber ring laser using a new Stretched Pulse APM principle. This new method is important for the simple reason that erbium doped fiber lasers delivering reasonable average power with short pulses have a Kerr effect that is too large. The active fiber cannot be made short (<1.5m). The mode diameter is small throughout the entire length of the resonator. When the nonlinear phase shift in the laser resonator approaches 2π, the pulses tend to become unstable. This prevents the generation of ultrashort pulses of reasonable energies, unless ways are found to reduce the nonlinearity. The Stretched Pulse APM works with pulses that "breathe" as they circulate through the ring, remaining short only within two small sections of the overall length of the ring. The fiber ring is composed of two pieces of fiber of opposite dispersion. The gain section of erbium doped fiber is positively dispersive, the passive section closing the ring is negatively dispersive. The two dispersions approximately cancel. The pulse can expand to as much as 20 times its shortest width. By tapping off the pulse through a coupler and fiber of appropriate length, the width of the output pulse can be made to be equal to the minimum width. The Stretched Pulse principle could be applied to any laser in which the Kerr nonlinearity is excessive.

Harmonic modelocking of all-fiber lasers is essential if one aims at GHz repetition rates, since the fundamental repetition rate of a fiber laser is in the 40-100 MHz regime. Because the relaxation time of the gain medium (Er or Nd) is long compared with 1 ns, the gain medium does not stabilize the pulse energies of a harmonically modelocked fiber laser. At AT&T Bell Laboratories, harmonic modelocking was achieved using a Fabry-Perot resonator internal to the fiber ring,5 so that each pulse seeded

---


its successor. The side-coupled resonator used at NTT accomplished the same purpose. Both schemes need stabilization of the optical lengths of the two resonators involved. Graduate student Christopher Doerr achieved harmonic modelocking at 1 GHz by reversing Additive Pulse Modelocking into Additive Pulse Limiting. He simply changed the polarization transformers in the fiber ring from a configuration that provided reduced loss for increasing intensity to one that provided increased loss for increasing intensity. His system is self-stabilized.

We have begun work on thulium doped fiber lasers emitting in the 1.9 \mu m regime. Thulium is eye-safe, has a much wider bandwidth than erbium, and a very smooth bell-shaped gain curve. APM of thulium fiber lasers will lead to pulses much shorter than those obtained with erbium. Preliminary studies on cw Fabry-Perot type thulium doped lasers showed pump induced "darkening" of the fiber that fully disappears after the pump is turned off. The process is well understood, not requiring further investigation, since the presence of lasing reduces the effect and permits pumping at a level which produces satisfactory levels of gain.

1.3 Long Distance Fiber Communications

Sponsors
MIT Lincoln Laboratory
National Science Foundation
Grant ECS 90-12787

Project Staff
Professor Hermann A. Haus, Christopher R. Doerr, Farzana I. Khatri, John D. Moores, William Wong

We have extended our simulation studies to explore high bit-rate TDM on optical fibers. Repeaterless terrestrial fiber communications operate under different constraints from those of transoceanic fiber communications. Also, the millions of miles of fiber already laid with zero dispersion at 1.3 micron wavelength require new methods of upgrading to higher bit-rates and 1.5 micron operation, since erbium amplifiers are superior to any other fiber communications amplifier. Solitons are not likely to play a role in the upgraded existing fiber network, but they do offer interesting possibilities for properly chosen fiber dispersion. We have run extensive analytic studies and computer simulations to study the effect of Raman noise, Raman self-frequency shift, soliton interaction and the Gordon-Haus effect to explore limits of achievable performance.

Another important aspect of long distance propagation is optical storage. High bitrates can be stored in a fiber ring if one can maintain the zeros and ones injected into the ring indefinitely. We have developed the theory for one of these storage rings. Remarkably, Christopher Doerr in his experiments on harmonic modelocking, demonstrated the possibility of storing a random sequence of zeros and ones indefinitely by switching his harmonically modelocked fiber ring from the APL state into the APM state. Under these conditions, he was able to show that random sequences of zeros and ones can maintain themselves forever, consistent with theoretical predictions.

1.4 Squeezing

Sponsors
Charles S. Draper Laboratories
Contract DL-H-441692
Fujitsu Laboratories
U.S. Navy - Office of Naval Research
Grant N00014-92-J-1302

Project Staff
Professor Hermann A. Haus, Professor Erich P. Ippen, Keren Bergman, Luc Boivin, Jeffrey K. Bonds, Christopher R. Doerr, David J. Dougherty, Ilya Lyubomirsky, Dr. Matasaka Shirasaki, Dr. Franz X. Kärtner

The MIT effort in squeezing concentrates on the pulse excited Sagnac fiber loop reflector first proposed in 1989. This scheme has the advantage of

---


reusing the pump as the local oscillator. The first results on squeezing\textsuperscript{10} were achieved in a fiber with exceptionally low noise caused by Guided Acoustic Wave Brillouin Scattering (GAWBS); i.e., index fluctuations caused by thermally excited acoustic waves which change the phase of the wave propagating through the fiber. For some very narrowly defined fiber parameters, it is possible to distribute the acoustic resonances so that they do not converge into the (low-) frequency range of the measurement window (40-90 kHz) for pulses of 100 MHz repetition rate (the rate generally produced by modelocked laser sources). The fiber used met these exceptional conditions. Since these conditions are not easily met, it is important to develop techniques that suppress the GAWBS noise. The GAWBS spectrum "rolls off" before it reaches 1 GHz. Thus, if the pulse repetition rate is 1 GHz or higher, the convolution of the GAWBS noise into the low frequency measurement window can be avoided entirely.

Graduate student Keren Bergman demonstrated the suppression of Guided Acoustic Wave Brillouin Scattering (GAWBS) in her squeezing experiments in fiber Sagnac rings by developing a low noise Nd:YLF modelocked laser source modelocked at 1 GHz. Ms. Bergman was able to obtain noise levels 5.1 dB below shot noise. The power of this method is that it can be used in any pulsed system subject to GAWBS, as long as the detection window is limited to low frequencies.\textsuperscript{11}

Careful investigation of the parameters of the system has demonstrated that the amount of shot noise reduction is consistent with the system parameters. In fact, the peak of the gaussian pulse experiences 13 dB of squeezing, but due to the averaging of the squeezing upon detection, and the imperfect quantum efficiency of the detector, the shot noise reduction is limited to 5.1 dB. Theoretical studies have shown that with pulse preshaping\textsuperscript{12} the shot noise reduction can be improved.

We have carried out a careful investigation of the Raman noise effect on squeezing. Raman noise is somewhat analogous to GAWBS, except for the fact that the index fluctuations are induced by optical phonons rather than acoustic ones. Thus, it is a broad-band noise that does permit reduction or elimination by the schemes proven successful in the case of GAWBS. Previous investigators have concluded\textsuperscript{13} that the noise does not affect squeezing appreciably at presently achieved levels (5 dB shot noise reduction) and that it is pulse-width independent. Our analysis is based on new careful experiments of the silica fiber Raman gain near zero frequency offset. One finds that the gain does not rise linearly as a function of frequency offset. The consequence is that the noise is not pulse-width independent, but rather increases with decreasing pulse-width, suggesting that for pulses shorter than 1 ps the Raman noise may become important.

Since squeezed radiation quickly degenerates through loss, it is important to use it in systems that, in principle, are loss-free for the squeezed radiation. Conventional fiber gyro designs do not have this property and hence modified designs must be used. In the search for such a modified design, Mr. Doerr and Dr. Shirasaki, a visitor from Fujitsu Laboratories, have arrived at a design which not only satisfies this criterion, but has several other unique properties that could lead to an improvement in conventional fiber gyro performance. Its advantages are:

1. Conventional fiber gyros bias the Sagnac fiber loop with a nonreciprocal phase shifter, or through modulation of the input excitation. The new design biases via a simple reciprocal wave plate.

2. The design is insensitive to "drifts" in all elements to first order.

The gyro has been built and tested, and all indications are that it will meet our expectations.

An analysis of squeezing calls for an accurate quantum theory of nonlinear optical systems. The excellent agreement between theory and our experiments suggest that the theoretical analysis of single-mode quantum processes developed by us and others is a reliable description of a class of such processes. Every measurement of an optical observable involves a nonlinear process. Hence the theory developed in connection with squeezing


can be applied to long-standing questions about the theory of quantum measurement. One of these questions is related to the evolution of the wave function of a system subjected to a measurement. The postulate of van Neumann concerning the "collapse of the wave function" upon a measurement can be explored in detail. The collapse is an extreme form of a measurement in which perfect accuracy is achieved. Most measurements are not of this character and must be treated differently. The quantum formalism developed for nonlinear optics enables one to characterize such more general, and more realistic, measurements.\textsuperscript{14}

1.5 Integrated Photonic Components: The Channel Dropping Filter

Sponsors

National Center for Integrated Photonic Technology
National Science Foundation
Grant ECS 90-12787

Project Staff

Jay N. Damask, Mohammed J. Khan, Professor Hermann A. Haus, Professor Leslie A. Kolodziejski, Juan Ferrera, Vincent V. Wong, Professor Henry I. Smith

The channel-dropping filter uses a combination of side-coupled quarter-wave shifted distributed Bragg reflectors to selectively tap a single wavelength-division multiplexed (WDM) channel from a central optical bus waveguide.\textsuperscript{15} Design studies have shown that five 10 GHz and 2 GHz channels at -20 dB crosstalk can be supported in the silica and InP semiconductor material systems, respectively. Moreover, a design modification may allow over twenty 10 GHz channels to be dropped with the aforementioned crosstalk figure.

Unlike current concepts in WDM receivers, such as the Integrated $N \times N$ Multiplexer/Router, the channel-dropping filter does not terminate the entire WDM bit stream to resolve one channel. Instead, each filter taps only one channel from the optical bus, leaving the remaining channels undisturbed and adding a new degree of design freedom for the network architect. Applications for the filters range from tapping a synchronizing or triggering bit stream to building a full WDM system to optically controlling a phased-array radar. The high-Q integrated optical resonators, the requisite technology for the channel-dropping filter, can itself be used for optical digital signal processing applications.

There are four components to the current project: (1) theory and design, (2) material development, (3) fabrication of waveguides and Bragg gratings, and (4) measurement and characterization. Theoretical tools have been developed to assist device design and to assess the fabrication tolerances that must be met to repeatably manufacture these grating-based filters.

Material development is following a dual course. The InGaAsP/InP semiconductor material system is an attractive platform because MSM detectors and other components can be directly integrated with the filters. The films are being grown in Professor Kolodziejski's GSMBE system and characterized in a number of support facilities. The silica-on-silicon system is a low-loss material system that can support a larger number of WDM channels than an equivalent semiconductor-based filter system can. The silica films are deposited by diffusion and LPCVD methods in the Integrated Circuits Laboratory at MIT.

Optical and x-ray lithography is required to pattern the waveguides and Bragg gratings, respectively. A new technique called Spatial Phase-Locked E-Beam Lithography, developed jointly between the Nanostructures Laboratory, directed by Professor Smith, and IBM, has demonstrated e-beam field-to-field stitching errors of $\pm$ 1 nm. Only with such small stitching errors can the requisite spatially-coherent gratings be written. The first set of straight and quarter-wave shifted Bragg patterns has been written for subsequent transfer to silica-based waveguides.

Careful measurement and characterization of all devices, from straight waveguides to Bragg reflectors to optical resonators to channel-dropping filters, is necessary to compile a complete understanding of the behavior and operation of the WDM filters. Both spectral and temporal measurements are required. A collaboration with Drs. J. Donnelly and S. Chinn of MIT Lincoln Laboratory has recently been formed to further the device fabrication and optical measurement components of the research.


We have started to investigate realizations of the channel dropping filter other than those based on the quarter wave shifted DFB resonator (QWS-DFBR). Gain gratings offer interesting new possibilities. The theoretically predicted performances are superior to QWS-DFB resonators. One remarkable property of the gain grating system is that one single side-coupled resonator can both couple out the signal from the "bus," as well as reinject it; this should be contrasted with a QWS-DFB grating structure that requires three resonators to accomplish this same function.

1.6 Ultrafast Nonlinearities in Active Semiconductors

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001
National Center for Integrated Photonic Technology
U.S. Air Force - Office of Scientific Research
Contract F49620-91-C-0091

Project Staff
Ali M. Darwish, David J. Dougherty, Katherine L. Hall, Charles T. Hultgren, Gadi Lenz, Professor Erich P. Ippen

Future high bit-rate communication systems may employ devices that contain active optical waveguides. Examples of these systems include all-optical switches, wavelength shifters, and optical amplifiers. The realization and reliable performance of such devices require knowledge of the ultrafast optical nonlinearities exhibited by these active waveguide structures. For some devices, such as optical amplifiers, highly linear behavior is essential. With other devices, such as switches and wavelength shifters, large optical nonlinearities are desirable and translate into good performance. Thus, the ability to design active waveguide structures and accurately predict their nonlinear and linear characteristics is important.

For long-haul fiber communications, InGaAsP waveguides that operate around 1.5 μm wavelength are the devices of choice. For shorter distance transmission, where pulse dispersion is not detrimental, AlGaAs waveguides operating around 0.8 μm could be used as well. We have two parallel research efforts underway for the study of both of these materials systems. Studying AlGaAs and InGaAsP waveguides together gives us the unique opportunity to compare the ultrafast behavior of devices from these two materials systems. The similar dynamics that they exhibit help to reveal fundamental processes that are common to all semiconductors.

To measure ultrafast nonlinearities in an active waveguide we employ short optical pulses in a pump-probe configuration. For InGaAsP studies around 1.5 μm, we use a color center laser that generates pulses around 150 femtoseconds in duration. The AlGaAs measurements near 0.8 μm are performed with a modelocked Ti:sapphire laser that generates pulses between 80 and 120 femtoseconds. The AlGaAs experiments employ a time-division interferometer (TDI) arrangement. The TDI is a pump-probe measurement technique that gives us the ability to measure both the gain and refractive index nonlinearities induced in a waveguide by a short optical pulse. The InGaAsP experimental setup employs a heterodyne pump-probe technique that also allows measurement of both gain and refractive index dynamics.16 The pump-probe measurements performed with these systems on InGaAsP and AlGaAs devices have enhanced our understanding of optical nonlinearities in semiconductors on a picosecond and faster time scale.

The dominant dynamics we have observed are attributed to carrier heating, spectral hole burning, and two-photon absorption.17 In addition, the refractive index exhibits an instantaneous response which is caused by the optical Stark effect. Below-band refractive index measurements performed in AlGaAs have allowed us to characterize the resonant response exhibited by the optical Stark effect. These below-band studies have also demonstrated that free carrier absorption is an important mechanism for carrier heating and, further, that there is a delay of about 100 femtoseconds in the onset of this carrier heating response. Other above-band measurements have shown that stimulated transitions can also give rise to carrier temperature changes. In particular, carrier cooling through stim-


ulated absorption has been observed. Other studies have revealed that chirp on the pulses can lead to artifacts in the refractive index measurements. Thus, careful control of pulse chirp is essential for making accurate index measurements.

In addition to being interesting from a fundamental standpoint, these pump-probe studies of active AlGaAs and InGaAsP waveguides are of practical interest. In the future, the index dynamics we have measured and characterized could become the basis for an all-optical switch. By studying them, we can determine ways in which these nonlinearities can be enhanced or suppressed, as desired, by proper design of the waveguide structure. In some cases, the processes may not be designed away, resulting in fundamental limitations to device performance.

1.7 Femtosecond Raman Measurements in Optical Fibers

Sponsor
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff
David J. Dougherty, Professor Hermann A. Haus, Professor Erich P. Ippen, Dr. Franz X. Kärtner

The magnitude and spectral dependence of the Raman gain of silica glass were studied by a number of researchers twenty years ago. The results of these investigations have subsequently been used in investigations of the effect of Raman gain on pulse compression using optical fibers, and in designing fiber Raman lasers and amplifiers. Intra-pulse Raman effects, such as recently proposed theories of noise generation in squeezing experiments using solitons and the soliton self-frequency shift, depend on the low frequency region of the gain spectrum. To assist further investigation of these effects, we have made new measurements of the Raman gain spectrum from 6 cm\(^{-1}\) to 900 cm\(^{-1}\) with a novel femtosecond probing technique.

The experiment used the beams of two Ti:sapphire lasers arranged to counter-propagate through 700 meters of single mode polarization preserving fiber. The pump beam was CW with a spectral width of approximately 3 GHz and a typical average power of 20 mW. The probe laser was mode-locked and provided 100 fs pulses which were coupled into a short length of fiber to provide an 80 nm broadband probe beam. The large bandwidth of the probe spectrum enabled us to obtain the Raman gain spectrum over large Stokes and anti-Stokes detunings simultaneously. In principle, the resolution of this technique is limited only by the resolution of the spectrometer and by pump back-scatter at zero detuning.

The peak gain at 440 cm\(^{-1}\) was measured to be 1.6 \( \times 10^{-11} \) cm/W, approximately 10 percent greater than previously measured. For applications such as Raman amplifiers or lasers which use the peak of the spectrum, the Raman gain for cross-polarized pump and Stokes beams may be taken to be zero. This is true only at large Stokes detunings. Our method has enabled us to present the first direct measurements of the cross-polarized Raman gain spectrum in glass fibers. The gain peaks at around 100 cm\(^{-1}\) and rapidly decreases for larger detunings. Below 50 cm\(^{-1}\), however, the gain has the same shape as the co-polarized spectrum and is only a factor of 0.45 smaller. Cross-polarized gain should therefore be included in calculations of effects such as the soliton self-frequency shift which are due to the low frequency Raman modes of the glass.

In addition to characterizing the Raman gain, our experiment also uncovered a new coherent coupling mechanism between counter-propagating pulses and CW beams in fibers. The effect is observed through oscillations in the signal spectrum occurring on the anti-Stokes side of the pump wavelength. The size of the signal depends linearly on the pump power, and on the square of the probe power. For probe pulses polarized parallel to the pump, the spectral oscillations always occur on the Anti-Stokes side of the pump wavelength and move with the center wavelength reference as the pump is tuned. If the probe is present in both polarization modes of the fiber, oscillations are seen on the cross polarized component of the probe, but shifted to the wavelength which has the same group velocity as the pump light.

This behavior is explained by transient Brillouin scattering of the pump followed by coherent cross-phase modulation. Sound waves launched by the

---

beating of the pump and probe electric fields scatter pump light backwards into the probe direction with an intensity envelope described by the impulse response of the elastic waves of the medium. This scattered electric field then heterodynes with the probe field by means of the instantaneous Kerr nonlinearity to create an oscillating, time dependent phase shift (and thus frequency shift) of the highly chirped probe pulse. Because glass has large positive dispersion at 800 nm and the scattered pump light trails the zero detuning wavelength, the oscillations appear on the anti-Stokes side. So far, they have limited our resolution of the Raman gain at detunings less than 6 cm⁻¹. However, the understanding we have obtained of such dynamic frequency modulation in chirped-probe, two-beam coupling experiments should make it possible to apply this simple technique to studies of fast nonlinearities in other waveguide devices.

1.8 Coherent Phonons in Solid-State Materials

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001
U.S. Air Force - Office of Scientific Research
Contract F49620-91-C-0091

Project Staff
Tak K. Cheng, Professor Erich P. Ippen

Pulses of laser light, which are short compared to the optical phonon period, can be used to impulsively generate the coherent lattice vibrations in solid state materials. Our current interest in coherent phonons is two-fold: (1) using time-resolved femtosecond spectroscopy to further our understanding of lattice and electron dynamics, and (2) exploiting the coherent phonon phenomenon to modulate the physical characteristics of the material at THz frequencies. For semimetals and narrow-gap semiconductors, the reflectivity modulations due to coherent ion motion have been particularly large (on the order of 10 percent in some materials) and have suggested significant modulation of the material's electronic structure. In these cases, we have determined that the coherent phonon excitation is caused by a mechanism we call Displacive Excitation of Coherent Phonons (DECP).²⁰

The DECP mechanism uses the pump pulse to promote electrons to the bonding orbitals of the solid, thereby changing the ionic potential's coordinate of minimum energy. Because the lattice cannot respond adiabatically to this rapid displacement of the ionic equilibrium coordinate, the lattice is set into oscillation about the new equilibrium. With the DECP mechanism, the generated coherent phonon can be large enough to modulate significantly the energy bands of a solid through a deformation potential coupling.²¹ Energy band modulation is manifested through transient reflectivity modulations as large as 12 percent in semiconducting TiO₂ at the A₁g optical phonon frequency (~ 7 THz). In experiments involving high excitation intensities, it is observed that the optical phonon is screened by the photoexcited carriers, resulting in an initial down-shift of the optical phonon frequency relative to the value found through spontaneous Raman scattering.

We have recently performed coherent phonon measurements on single crystal Sb, as a function of pump intensity, to study systematically the effect of the phonon frequency shifts. These shifts are directly related to the time evolving curvature of the ionic potential caused by the pump-induced plasma. The plasma screens the bare-ion potential of the solid yielding a "dressed" phonon with a modified phonon frequency, analogous to the case of acoustic phonons in metals. By monitoring the phonon frequency, we directly measure the relaxation of the carrier-ion screening process.

In most femtosecond studies of carrier dynamics, the induced optical changes are the result of carrier distributional changes that evolve with energy and momentum loss to the lattice. In the present experiment, the induced optical changes are dominated by the lattice and we measure changes in the phonon frequency as affected by the ions' interaction with the excited carriers. In the most direct way possible, we observe electron-phonon interactions as experienced from the point of view of the lattice itself.

1.9 Publications and Meeting Papers


Chapter 1. Optics and Quantum Electronics


1.10 Ultrashort Pulse Generation in Solid State Lasers

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001
National Science Foundation
Grant ECS 85-52701
U.S. Air Force - Office of Scientific Research
Contract F49620-91-C-0091
U.S. Navy - Office of Naval Research (MFEL)
Grant N00014-91-C-0084

Project Staff
Laura E. Adams, Dr. Brett E. Bouma, Dr. Artur D. Gouveia-Neto, David Huang, Malini Ramaswamy, Morrison Ulman, Dr. Lucio H. Acioill, Dr. Joseph A. Izatt, Professor James G. Fujimoto

1.10.1 Ultrashort Pulse Generation in Solid State Lasers

The development of new laser technology for ultrashort pulse generation is directly relevant to the investigation of ultrafast phenomena in materials and devices as well as the development of high speed communications and signal processing systems. In recent years there has been a rapid growth of solid-state laser technology which has revolutionized the field of short-pulse generation. Wavelength tunable pulses with durations of less than 100 fs can now be generated using a variety of new solid state laser systems. Solid state lasers have high energy storage, high reliability, and low cost. Over the last several years, our group has been active in the development of new technologies for ultrashort pulse generation in solid state lasers. Working in collaboration with Professors Erich P. Ippen and Hermann A. Haus, we have developed a new modelocking technique called additive pulse modelocking (APM) and demonstrated it in a wide variety of systems including the Ti:Al2O3 laser as well as the diode pumped Nd:YAG and Nd:YLF lasers.22 We have also developed theoretical descriptions of short pulse generation to explain the fundamental mechanisms for pulse shaping in lasers as well as to provide design criteria for new laser systems.23

Over the last year, one of the major areas of investigation has been to explore the factors which limit the shortest pulse durations that can be generated from lasers. We have explored the role of higher order intracavity dispersion in limiting the pulse duration from Ti:Al2O3 lasers and developed a novel technique for controlling third order intracavity dispersion using a thin film Gires-Tournois Interferometer. This allowed us to build a laser with independently adjustable second and third order dispersion which generated a record pulsed width of 28 fs.24 Subsequently, other techniques to minimize intracavity dispersion have been developed and currently 11 fs pulses can be generated directly from a Kerr-Lens-Mode locked Ti:Al2O3 laser.25

1.10.2 Kerr Lens Modelocking for High Power Pulse Generation in Nd:YLF

Recently a new short pulse generation technique has been developed based on the use of self-focusing (Kerr lens) effects which can produce a fast saturable absorber like action.26 This technique is called Kerr lens modelocking (KLM) and has had a major impact on the development of third generation solid state lasers. Our group has been investigating new techniques for KLM including the development of monolithically integrated nonlinear mirrors for KLM as well as theoretical descriptions


for self-focusing nonlinearities to facilitate the design of KLM laser systems.\textsuperscript{27}

To date, most current femtosecond laser geometries use the Kerr nonlinearity in the gain medium to obtain saturable absorber action. Our group has recently developed a new nonlinear intracavity device for passive modelocking which is called the microdot mirror.\textsuperscript{28} The microdot mirror uses nonlinear self-focusing that occurs in a bulk intracavity element to produce fast saturable absorber modelocking. The advantages of the microdot modelocker are that it is modular, compact and independent of pump geometry and cavity design.

We have demonstrated the microdot mirror to modelock an Argon laser pumped Ti:Al\textsubscript{2}O\textsubscript{3} laser.\textsuperscript{28} Recently, we have extended the use of this novel modelocker to demonstrate KLM in a lamp-pumped Nd:YLF laser\textsuperscript{29} to achieve pulse durations of 2.3 ps. These pulses are the shortest pulse durations that have been generated from a lamp pumped Nd:YLF laser and are close to the bandwidth limit of the Nd:YLF gain medium. According to theory, stable chirp-free generation of pulses requires both saturable absorber action and net negative dispersion. Our laser incorporated a microdot mirror modelocker for saturable-absorber action and an air-space Gires-Tournois interferometer for compact dispersion compensation. The amount of dispersion could be varied by adjusting the spacing between the mirrors in the GTI. This is, to our knowledge, the first demonstration of KLM in a lamp-pumped laser geometry. We believe that it proves the utility of the microdot-mirror technique to modelock lamp-pumped lasers and thus has important implications for the development of high power modelocked lasers.

We are currently investigating the scalability of KLM into another regime—that of compact, high rep-rate laser geometries. This effort is motivated by the need for compact, reliable and diode-pumpable femtosecond laser geometries. In current KLM laser designs, the primary hindrance to the achievement of compactness is operation in the negative dispersion regime and the consequent use of a prism pair. Another possible regime of operation which we are exploring is the positive dispersion regime, where by enhancing the saturable absorber action to self-phase-modulation ratio one can in principle achieve stable short pulse generation. Theory has shown that this requires working close to the geometrical stability limit of the laser, and we are examining the experimental validity of this concept for femtosecond laser design. A concurrent approach to the problem is the development of novel and compact ways of providing dispersion compensation using novel laser geometries.

### 1.11 Cavity Dumping Techniques for Modelocked Ti:Al\textsubscript{2}O\textsubscript{3} Lasers

For many applications, it is desirable to generate high intensity ultrashort pulses for studies of nonlinear phenomena. Several amplification techniques have recently been developed to extend the range of available pulse energies from Ti:Al\textsubscript{2}O\textsubscript{3} laser sources. However, the requirement of multiple stages and/or multiple pump lasers makes these oscillator-plus-amplifier systems complex and relatively expensive. Further, the repetition rate of many of these sources is in the kHz range and is low enough to limit detection sensitivity for ultrafast measurements.

To address these issues we have studied cavity dumping as an alternative approach to amplification for generating increased pulse energies. Cavity dumping maintains the performance advantages of modelocked Ti:Al\textsubscript{2}O\textsubscript{3} lasers in terms of stability, tunability and pulse duration, while providing a significant increase in pulse peak power at variable repetition rates. We have performed what is to our knowledge the first demonstration of cavity dumping in a femtosecond Ti:Al\textsubscript{2}O\textsubscript{3} laser.\textsuperscript{30} Because solid state lasers have relatively small gain cross sections and high saturation energies, a significant increase in intracavity pulse energy can be achieved using high Q cavities. Our cavity dumper was an acousto-optic Bragg cell placed in an intra-


cavity focus. Pulse durations as short as 50 fs with pulse energies as high as 100 nJ have been achieved at repetition rates as high as 950 kHz. The dumping process is non-resonant and does not limit the bandwidth of the laser down to the 50 fs level. The resultant peak powers of 1 MW are sufficient to permit the study of a wide range of nonlinear optical effects.

This technique should prove to be an attractive and relatively simple approach for high intensity short pulse generation for a wide range of ultrafast studies. The limitation on the available pulse energies was imposed in our system by the multipulse instabilities that tend to arise at high pulse energies. Electro-optic dumping techniques, which we plan to investigate in future experiments, may be able to achieve higher pulse energies.

1.1.1.1 Modelocking Techniques for Transient Flashlamp Pumped Ti:Al$_2$O$_3$ Lasers

Nearly all short pulse Ti:Al$_2$O$_3$ lasers to date have been cw modelocked systems with relatively low pulse energies. Short pulse energies in the range of microjoules to millijoules can be attained using regenerative and multipass amplifier systems. Multistage chirped pulse amplification techniques can produce short pulse energies up to the joule range. While these techniques achieve excellent performance, they require multiple pump lasers for the oscillator and amplifier.

In order to develop a simple economical high peak power laser source, we have investigated mode-locking in a pulsed flashlamp pumped Ti:Al$_2$O$_3$ laser. Modelocking transient systems are extremely challenging because, unlike cw lasers, the pulse energy evolves rapidly. Pioneering studies on transient modelocking were performed in flashlamp pumped Ruby and Nd:Glass lasers almost three decades ago.$^{31}$ Performances of 50 mJ, 1 nsec and 1 mJ, 50 psec, respectively, were achieved in these lasers using saturable absorber dyes whose excited state lifetimes were fast compared to the pulse durations obtained. The next generation of mode-locked sources obtained shorter pulse durations in quasi cw operation with "slow" saturable absorbers using a combination of saturable absorption and saturable gain. These techniques were used in both pulsed and cw pumped configurations, eventually reaching pulse durations as short as 30 fs in the colliding pulse modelocked laser. Relatively few investigations of modelocking have been performed in flashlamp pumped Ti:Al$_2$O$_3$. Mode-locking with saturable absorber dyes$^{32}$ and intracavity second harmonic generation$^{33}$ have been demonstrated and resulted in relatively long pulse durations of ~100 ps. With the development of new modelocking techniques based on solid-state, intracavity nonlinear modulators, it is now possible to develop high performance modelocked pulsed Ti:Al$_2$O$_3$ laser technology.

In the transient Ti:Al$_2$O$_3$ system, pre-saturation exponential growth of the pulse belies traditional modelocking techniques, since pulse energy changes dynamically in time. In addition, relaxation oscillations arise from the high gain and limit the time available for pulse shortening to just a few μs, on the order of the excited state lifetime in Ti:Al$_2$O$_3$. We have modelocked flashlamp pumped Ti:Al$_2$O$_3$ using hybrid active/passive modelocking that combines different intracavity pulse shaping mechanisms in order to obtain optimal pulse shortening during the buildup time of the transient laser pulse. Active modelocking is effective for rapidly shortening pulses which are long compared to the modulation period, while in fast saturable absorber modelocking, pulse shortening increases with increasing intensity. We have employed an acousto-optic modulator (AOM) to provide active modelocking. When the AOM is the only modelocking element pulses with durations of 100 ps can be generated. In order to further reduce the pulse duration we have employed an antiresonant semiconductor Fabry-Perot saturable absorber (AFPSA) as a cavity end mirror for fast saturable absorber pulse shaping.$^{34}$ The combined active and passive modelocking allows the production of 5 ps pulses at


a single pulse energy of ~ 20 μJ. Extension of this work to higher pulse energies is underway. By scaling the mode size on the saturable absorber and by cavity dumping, single pulse energies in the millijoule range should be obtainable.

The hybrid modelocking technique demonstrated is not constrained to any one particular modelocking mechanism. Other fast saturable absorbers in combination with optical limiters and energy integrating nonlinearities may be used to produce picosecond pulses in the millijoule range throughout the tuning bandwidth of TiAl₂O₃. In addition, the excellent mode quality of this laser should allow the production of powerful blue-green and ultraviolet pulses through efficient harmonic generation.

1.11.2 Modelocking Techniques for High Power Diode Arrays

High power, picosecond pulses are of interest for optical communications and signal processing applications. The highly efficient and compact nature of semiconductor diode lasers makes these sources attractive for generating short optical pulses. Active modelocking of single-stripe diodes at high modulation frequencies can produce subpicosecond pulses. However, the power available from mode-locked single-stripe diodes is limited. More power can be made available for modelocking by use of large-area diode arrays or broad-area lasers. However, the large capacitance of broad-area lasers and diode arrays makes high frequency modulation difficult, limiting the pulse widths that can be achieved from active mode locking. Another method for generating short pulses is passive mode locking with a saturable absorber created by ion implantation of one diode facet. Passively mode-locked ion-implanted diode arrays have been demonstrated. If, instead, the saturable absorber is independent of the laser structure, it can be custom designed for optimal saturation and recovery characteristics. Single-stripe devices have been mode-locked in an external cavity using a multiple-quantum-well (MQW) saturable absorber. Without external amplifiers, the maximum average power was limited to 1 mW.

It is desirable to have a source for generating high-power picosecond pulses that offers both simplicity and flexibility of design. We have developed a simple, compact, high power picosecond pulse source consisting of a 400 μm wide MQW broad-area laser and a patterned MQW sample providing both saturable absorption and spatial mode control. To our knowledge, this is the first demonstration of a hybrid mode-locked broad-area semiconductor laser with an MQW saturable absorber in an external cavity. The absorber that we have used is custom tailored for our desired absorption and is grown under conditions that ensure a fast recovery time. Our external cavity is limited to only a few optical elements including an output coupler, saturable absorber, and lenses.

Our saturable absorber is a low-temperature (LT) (~300°C) molecular-beam-epitaxy-grown AlGaAs-GaAs MQW structure. The LT MQW structure is grown on top of a dielectric mirror stack grown on a GaAs substrate. In other diode mode-locking experiments that used a MQW saturable absorber, the GaAs substrate was removed and the absorber was epoxied to a high reflection mirror. Here we have the ability to design the dielectric mirror and integrate it monolithically with the MQW saturable absorber. This approach was demonstrated previously for saturable absorber mode locking of solid state lasers. Stable modelocking requires that the loss both saturate and recover faster than the gain. Previously, proton bombardment and tight focusing methods were used to speed up the recovery time of the saturable absorber. We rely on the low-temperature MBE growth to provide a fast recovery time.

Replacing a single-stripe diode laser with a broad-area semiconductor laser introduces the issue of mode control in order for light to be well focused onto the saturable absorber. We have developed a novel method for achieving this mode control. The MQW sample is patterned into microdots using a selective chemical etch. By translating the sample, we lase on dots of increasingly smaller diameter, which act as apertures, forcing the laser to operate in a few of the lowest order modes.


Chapter 1. Optics and Quantum Electronics

The MQW sample, which is divided into a unpatterned region and a microdot region, is mounted onto an aluminum mirror. The effects of saturable absorption and mode control are evidenced by mode-locking with the Al mirror, the unpatterned MQW, and the microdot yielding pulse durations of 51, 37, and 15 ps, respectively. The 15 ps result was obtained by hybrid mode-locking at 593 MHz, with average power of 9 mW, and peak power of 1 W. The largest dots provided insufficient mode control, resulting in pulses comparable to those obtained with the unpatterned MQW. The smaller-diameter dots on our sample were damaged by the etching process. Recently, we have improved our processing by moving to a microstripe geometry patterned using ion-assisted etching.

Pump-probe measurements of our MQW sample revealed a recovery time of ~10 ps. We believe additional pulse shortening may be realized by using a faster recovery time MQW saturable absorber produced by lower temperature growth. MQW samples grown at 200°C and 250°C are currently being investigated.

1.12 Spectrally Resolved Autocorrelation for Femtosecond Diagnostics

While the field of ultrashort pulse generation has progressed rapidly, measurement techniques for femtosecond pulses are still relatively inadequate. Currently, temporal measurement techniques, such as intensity autocorrelation, determine pulse duration by making assumptions about the pulse shape. Spectral domain measurements do not provide adequate information to retrieve the complete temporal profile. We have recently addressed this issue by developing and experimentally demonstrating a new method for the complete determination of the amplitude and phase of femtosecond pulses.

Our method, called spectrally resolved autocorrelation (SRA), is both convenient and accurate and relies on a simple measurement of the spectrum of the second-harmonic generated in a conventional autocorrelation setup. A robust iterative algorithm is then used to retrieve the complex temporal amplitude of the pulse from the SRA. In contrast to other approaches, the SRA can be used to directly measure nonamplified pulses at the output of a femtosecond laser. We have demonstrated the SRA technique using our regeneratively mode-locked KLM Ti:Al₂O₃ laser as a source of pulses. We believe that this diagnostic technique will be useful in applications such as the generation of very short pulses or pulse compression, where accurate information about the amplitude and phase of pulses is essential.

1.13 Ultrafast Phenomena in Materials and Devices

Sponsors

Joint Services Electronics Program
Contract DAAL03-92-C-0001
National Science Foundation
Grant ECS 85-52701
U.S. Air Force - Office of Scientific Research
Contract F49620-91-C-0091
U.S. Navy - Office of Naval Research (MFEL)
Grant N00014-91-J-1956

Project Staff

Igor P. Bilinsky, Malini Ramaswamy, Chi-Kuang Sun, Morrison Ulman, Professor Erich P. Ippen, Professor James G. Fujimoto

1.13.1 Femtosecond Carrier Dynamics in GaAs - AlGaAs

We have established a multidisciplinary collaborative program to investigate ultrafast dynamics in semiconductor materials and devices. Our approach incorporates state-of-the-art experimental femtosecond laser measurements and theoretical ensemble Monte Carlo simulation. The combination of experimental and theoretical studies at an advanced level has allowed us to form an accurate and detailed description of ultrafast carrier dynamics in semiconductors.

Our collaboration is based on femtosecond laser experiments performed at MIT and Monte Carlo simulations and other analysis performed at the University of Florida in Professor C.J. Stanton's solid state theory group. We expect that this

Chapter 1. Optics and Quantum Electronics

project will ultimately enable us to predict the ultrafast dynamics of carriers in a wide range of semiconductors, engineered materials, and devices. This knowledge has broad reaching significance beyond fundamental interest because the carrier dynamics in a given semiconductor determine the physical limits of device speed. Since carriers scatter as fast as once every few femtoseconds, accurate models and relevant experiments must explore phenomena on a femtosecond time scale. Experimentally femtosecond pulsed lasers are required, and Monte Carlo simulation has proven to be the most effective analytical technique.

Working within this collaborative framework, we performed a study on AlGaAs using a tunable femtosecond source.40 In this work, we developed a source of optical pulses as short as 40 femtoseconds tunable in wavelength from the red to the near infrared. We used the tunable source to excite carriers to a sequence of different initial distributions in the conduction band. Then we probed the evolution of these carrier distributions. The carrier distribution changes rapidly via carrier-carrier scattering and relaxes in energy by phonon emission. We used the mole fraction of aluminum in AlGaAs as an additional parameter to control the possible mechanisms by which the carrier distribution evolves. For example, we picked Al$_{0.1}$Ga$_{0.9}$As in an effort to isolate the effects of electron intervalley scattering. Interv valley scattering is a fundamental process that is responsible for many high field transport properties in GaAs and AlGaAs. Professor Stanton’s group at the University of Florida simulated our results and created a model that predicts the behavior we observed over a wide range of initial carrier energies.

A colliding pulse modelocked ring dye laser is at the heart of our apparatus. It generates optical pulses as short as 35 femtoseconds. These pulses are amplified to microjoule energies in a two stage dye amplifier that is pumped by a copper vapor laser. The amplified pulses are focused into a flowing jet of ethylene glycol to generate a femtosecond white light continuum. The continuum is then filtered in a Fourier grating filter which selects the desired wavelength and bandwidth pulse corresponding to the energy and spread in energy of photoexcited carriers. We have recently improved our system by replacing the copper vapor laser with a diode pumped, Q-switched, frequency doubled Nd:YLF laser.41 Its all solid state design offers improved reliability and ease of use compared to the copper vapor laser.

Recently, we performed an experiment to examine the dynamics of carriers in the presence of a background cold carrier plasma.42 This experiment approximates the conditions found in diode lasers and is complementary to the diode laser experiments described in the next section. Our experiments use 0.2 micron thick bulk samples so it is possible for us to use shorter pulses than are used in standard diode experiments where the sample is several hundred microns long and dispersion limits resolution. We use a three pulse pump-probe technique in which the third pulse, the “prepulse,” is focused on the sample a few hundred picoseconds before the pump-probe measurement is performed. The prepulse is far enough ahead of the other pulses that the carriers it generates have time to cool to the lattice temperature, but not so advanced that those carriers have time to recombine. By using samples with different mole fractions of aluminum, we were able to “tune” the semiconductor instead of varying the laser wavelength. We investigated dynamics above and below the Fermi level established by the cold plasma and the effect of intervalley scattering. We found that the cold plasma increases the hot carrier thermalization rate by only a small amount and that this rate is slower than the intervalley scattering rate. We are continuing these experiments with our tunable laser system to better resolve the dynamics in energy.

The theoretical modeling of our results is undertaken in three steps. First, the semiconductor band structure is calculated. Next, the evolution of the carrier distribution function is tracked by a Monte Carlo simulation solution to the Boltzmann Transport Equation. Finally, the differential change in transmission versus pump-probe delay—the experimental results—are predicted. It is necessary to undertake this procedure in a large computer simulation because there are many optical transitions and scattering channels that must be considered. Monte Carlo simulation is the only way to extract

---

information about the carrier distributions from our experimental data. The simulation also permits us to artificially vary the strength of various scattering mechanisms to determine the sensitivity of the data to various processes. We have found that our models accurately predict the results of our tunable experiments over the full range of initial carrier energies that we investigated.

1.13.2 Carrier Dynamics in InGaAs Strained Layer Diodes

Nonlinear gain and transient carrier dynamics in diode lasers play important roles in laser linewidth, modulation bandwidth, amplification, and short pulse generation. Picosecond and femtosecond pump probe measurements of nonlinear gain dynamics were performed in bulk GaAs, InGaAsP MQW, and InGaAs/InGaAsP strained-layer MQW amplifiers. These studies have shown that non-equilibrium carrier temperature effects play a dominant role in nonlinear gain and carrier dynamics. Carrier temperature changes are caused by a number of processes including state filling produced by stimulated transitions, free carrier absorption, and two photon absorption. Previous studies have shown that free carrier absorption dominates carrier heating processes. As ~ 150 fs relaxation time had been observed which was attributed to either spectral hole burning or turn on delay of free carrier heating. In addition to nonlinear gain, carrier dynamics also produce nonlinear index effects such as self phase modulation. Recently, several new experimental techniques such as bias lead monitoring, time domain interferometry, and heterodyne detection have been developed to study carrier dynamics in waveguide device. These techniques are based on conventional single-wavelength pump-probe geometry. No multiple-wavelength pump-probe measurements have been reported.

Recently, we invented a new technique for performing independent multiple wavelength pump-probe measurement in waveguide devices. The output of a Kerr lens mode-locked Ti:Al₂O₃ laser was coupled to an optical fiber for self-phase-modulation spectral broadening. Two spectral windowing assemblies, which can select frequency and pulsewidth, were used after the fiber to produce synchronous different-frequency pump and probe pulses. The pump and probe wavelength can be independently varied and thus energy relaxation dynamics can be explicitly measured. This method represents a new and powerful approach for performing ultrastable experimental studies.

Working in collaboration with investigators at MIT Lincoln Laboratory, we have performed the first carrier dynamics studies in InGaAs/AlGaAs strained-layer quantum well diode lasers. Strained-layer quantum well devices represent one of the most active and technologically promising areas of current optoelectronics device research. With epitaxial growth of nonlattice-matched materials, strained-layer materials provide an added degree of freedom to band structure engineering and device fabrication and design. High power, high efficiency, long lifetime, and low threshold current density semiconductor lasers have been achieved in the laboratory by InGaAs strained layer devices in the past few years.

The sample used for the studies is a broad-area InGaAs/AlGaAs graded-index separate-confinement heterostructure single quantum well (GRIN-SCH


SQW) ridge waveguide diode laser with a bandgap near 960 nm. Femtosecond gain dynamics was investigated using a multiple-wavelength pump probe technique. Studies were performed by fixing the pump and probe wavelength and varying the bias current so that gain, transparency, and absorption were produced at the pump wavelength. Studies demonstrate that carrier temperature changes mediated by both free-carrier absorption and stimulated transitions strongly govern transient gain dynamics. The energy of the pump wavelength relative to the transparency point determines which process dominates the transient response. In contrast to previous studies, our investigations show that free carrier heating plays a much weaker role in determining carrier temperature dynamics than stimulated transition effects. Stimulated transition induced carrier cooling was observed in the absorption region which represents the first observation of carrier cooling in GaAs based devices.

In collaboration with theoretical physicists at University of Florida, Gainesville, a detailed theoretical model for gain dynamics was developed to aid in the interpretation of these results. In the model, transient gain and differential transmission are computed in a multiband effective mass model for the quantum well bandstructure and optical matrix elements including biaxial strain, valence subband mixing, and polar optical phonon scattering both within and between subbands. Transient photogeneration of electron hole pairs by the pump pulses and subsequent relaxation of carriers by polar optical phonon scattering are calculated in a Boltzmann equation framework. Good agreement between the experiments and theoretical calculations were obtained, with the exception of the features seen in the experimental data from two photon absorption and spectral hole burning. We find that the dominant scattering mechanism is longitudinal polar optical phonon scattering, and stimulated transitions play an important role in carrier dynamics.

An increased understanding of the physical mechanisms of gain dynamics from carrier temperature changes is important for the design of new devices. In particular, the reduction of carrier heating has important implications for reducing parasitic gain saturation effects in short pulse modelocked laser diodes and amplifiers.

1.13.3 Nonequilibrium Electron Dynamics in Metals

Studies of interactions of free carriers between themselves and with their environment constitute one of the major problems of solid state physics. This has been addressed directly in the time domain employing femtosecond techniques both in semiconductors and metals. Of particular interest is the case of metals which display a very high electron density whose behavior can be modeled on relatively simple basis. The possibility of creating and probing a transient nonequilibrium electron population in metal with ultrashort laser pulses has been demonstrated by different groups. In previous experiments, electron-electron interactions were assumed to be sufficiently fast to thermalize the electron gas on a time scale of the order or shorter than the laser pulse duration, although some deviations from an instantaneous response were observed. Recent investigations using transient photoemission have demonstrated the existence of non-Fermi electron distribution with thermalization

times as long as 600 fs.\textsuperscript{55} These results were observed in gold film for large changes of the electron temperature (of the order of 400°K) with a limited time resolution. Similar conclusions were drawn at lower laser fluence by analyzing the temperature dependence of the optically measured electron-phonon interaction time in gold and silver.\textsuperscript{56}

In order to analyze the effect of the non-instantaneous electron-electron interaction on the optical response of a metal film, recently we have performed transient reflectivity and transmissivity measurements in the very low perturbation regime ($\Delta T \sim 3°K$).\textsuperscript{57} In contrast to previous experiments, we develop a high sensitivity multiple wavelength femtosecond pump-probe technique. A nonequilibrium electron distribution was excited by free carrier absorption with an infrared pulse from a mode-locked Ti:Al$_2$O$_3$ laser with 120 fs pulsewidth and 880 nm - 1065 nm wavelength. The relaxation dynamics were followed by measuring the transient reflectivity and transmissivity using frequency doubled probe pulses in the visible, 440 nm - 532.5 nm, corresponding to the energy between the d-band and Fermi surface. This technique permits a more definitive measurement of electron dynamics by separating the effects of pump induced transitions from those monitored by the probe. In addition, by tuning of the probe wavelength, the transient electron distribution can be monitored close to Fermi surface where the optical property changes are maximum. This renders the measurement very sensitive and permits weak perturbations of the electron distribution. In this condition, the system response is linear, and the measured changes in reflectivity and transmissivity can be directly related to the electron distribution. This is in contrast to previous transient reflectivity and transmissivity measurements which probe the distribution far away from the Fermi surface and thus necessitate large perturbations of the electron gas. Experiments were performed in optically thin (200 Å) gold films, because the band structure of gold is relatively well known and electron diffusion effects can be neglected. The high stability, high repetition rate and tunability of the Ti:Al$_2$O$_3$ laser allows measurements at very low pump fluence inducing electron temperature changes only of the order of a few Kelvins.

Measurements show evidence for non-Fermi electron distribution with an electron thermalization time of $\sim 500$ fs and an electron-lattice cooling time of 1 ps, independent of the laser fluence in the range of 2.5 mJ/cm$^2$ - 200 mJ/cm$^2$ (corresponding to estimated peak temperature change 3-200°K). Measurements were also performed in an optically thick sample (1200 degrees). The apparent thermalization dynamics of the electron gas is much faster and is shown to be dominated by transport effects.\textsuperscript{58}

Recently, we developed a three-coupled rate equation model\textsuperscript{59} based on an extension of the classical model of Anisimov.\textsuperscript{60} The electron gas distribution changes are separated into thermalized and non-thermalized parts that are both coupled to the lattice. The diffusion in an optically thin sample is neglected. Combined with spectral lineshape analysis, the experimental data indicates the existence of a spectrally broad initial non-Fermi distribution relaxed around $\sim 300$ fs through the interaction with both electron gas and lattice system. The thermalization time through interaction with electron gas is on the $\sim 500$ fs time scale and the relaxation time through the interaction with lattice is on the $\sim 1$ ps time scale, separately. At energy close to the Fermi surface, long thermalization times on the order of 1 to 2 ps are observed from the reduction of the electron-electron scattering.


rates. The different line-shapes at different time delays reflect the lattice contribution.

The estimated electron thermalization time is comparable to the previous estimations. On the basis of the very high electron density of metals, one would expect very fast electron-electron interactions. However, the interaction efficiency is considerably reduced by phase space filling which blocks most of the energetically possible interaction channels and by screening which considerably reduces the efficiency of the Coulombic interaction. An overall thermalization time of few hundred femtoseconds can be estimated on the basis of the Fermi liquid theory. Our results are significant because they provide some of the first detailed information on non-Fermi electron dynamics in metal and demonstrate new experimental techniques for their investigations.

1.13.4 Time Gated Scanning Tunneling Microscopy

Femtosecond laser technology enables the highest time resolution measurements of solid state phenomena possible today. In the past decade femtosecond lasers have progressed from specialized systems to a commercially available technology that is now widely applied in physics and chemistry. Similar development has occurred in scanning probe microscopy. Since the invention of the scanning tunneling microscope in 1982 atomic resolution images of surfaces have become more and more routine every year. In our program on time resolved scanning tunneling microscopy we intend to combine the ultrafast time resolution of femtosecond lasers with the atomic spatial resolution of scanning probe microscopes. The goal is to develop an instrument that can perform highly localized measurements in four dimensions simultaneously.

In a scanning tunneling microscope (STM) a sharp metal tip is positioned a few angstroms from the surface under investigation. An electrical bias is applied between the tip and the surface. When the tip is close to the surface a small current will tunnel from the tip to the sample. The size of this current depends exponentially on the distance between the tip and the surface. In the constant current mode of operation, a feedback loop controls the height of the tip above the surface to maintain a constant tunneling current. This height is recorded as the tip is scanned across the sample and a topographic map of the surface results. In fact, the map shows the surface of constant electronic density of states, however the corrugations do not correspond directly to atomic positions.

If the tunneling current in an STM could be gated on a femtosecond time scale, it might be possible to observe nonequilibrium surface effects or perform ultrafast spectroscopy on single atoms or quantum dots. Other investigators have tried to gate the tunneling bias with photoconductive switches. Some studies have focused on the application of high speed SPM to circuit sampling. Preliminary experiments of this type have reported picosecond time resolution. In these experiments factors that limit the time resolution include spreading of the electrical pulse on its transmission line and the capacitance of the tip sample junction.

We are pursuing a different method of gating the tip which is based on creating a nonequilibrium carrier distribution in the tip, the sample or both. We know from our studies of hot electrons in metals and semiconductors that a photoexcited hot carrier distribution relaxes significantly within 100 femtoseconds and thermalizes in approximately 1 picosecond. We expect that tunneling will be significantly different from a hot electron distribution than from an equilibrium distribution. We are testing this concept with an equal pulse correlation technique using an amplified femtosecond laser system.

1.14 Laser Medicine

Sponsors

National Institutes of Health
Grant R01-GM35459-08
U.S. Air Force - Office of Scientific Research
Grant F49620-93-1-0301


1.14.1 Optical Coherence Tomography Technology

The development of optical coherence tomography (OCT) has led to a high-resolution imaging technology for biological tissues. An incident optical beam, which is reflected or backscattered from tissue structures of varying optical density, is used to construct an image. Time delays associated with the returned light yield useful information about both the depth and reflectance of the structures. OCT is analogous to ultrasound imaging, but offers a non-invasive, non-contact system. With 10 μm resolution, OCT provides better than 10 times the resolution of current tomographic imaging technologies such as CT, MRI, and ultrasound B mode imaging.

OCT technology utilizes low coherence interferometry to obtain the depth information of the tissue. Two arms comprise the interferometer. A sample arm represents the distance to the biological tissue while the reference arm includes a translating mirror. When the optical path lengths between the two arms are matched to within the coherence length of the light source, fringes appear as an output. The translation of the reference mirror produces a 1-D depth reflectivity profile of the tissue. This interferometric ranging technique is called optical coherence domain reflectometry (OCDR).

Using optical heterodyne detection with the application of noise-reduction techniques originally developed for optical communication, a sensitivity of 100 dB is possible for detecting the light reflected from the tissue.

A scanning mechanism using mirrors mounted on computer-controlled galvonometers enables precise lateral positioning of the beam. Two-dimensional cross-sectional maps of the tissue backscattering magnitude are obtained when the beam is scanned in a predetermined pattern. Software routines selectively scan the beam in either a linear mode for slice images or a circular mode for circumscripting about selected tissue structures such as the optic nerve head of the eye. The data can be digitally filtered and later presented with false-color or gray scale as an optical coherence tomograph.

Our development of the OCT technology and its integration into the clinical setting has been accelerated by the strong collaborative effort between the research facilities at MIT, the Optical Communications Group at MIT Lincoln Laboratories, and the clinical facilities at the New England Eye Center of Tufts University School of Medicine. Recent OCT advancements in the area of ophthalmology have improved the system's capabilities and ease of use. The system utilizes compact and inexpensive fiber-optic technology, including a fiber-optic interferometer, to deliver and receive the superluminescent diode laser light to and from the subject. The scanning and imaging optics have been retrofitted onto an ophthalmic slit-lamp biomicroscope to not only simplify the delivery system, but also make the technology portable and compatible with existing ophthalmic instrumentation. The 200 μW incident optical power on the retina is consistent with a conservative interpretation of the American National Standard Institute (ANSI) guidelines for safe laser retinal exposure. These guidelines have been followed throughout the development process to ensure future patient safety.

Hardware improvements have reduced the amount of time necessary for acquiring a scan. By improving the scan mechanism with the addition of a lightweight feedback controlled retroreflector and modifying the data processing circuitry and controlling software, only 2.5 s are required to obtain a 15 mm linear cross-sectional scan. This short scan


time reduces the likelihood of eye motion artifacts in the data as well as the patient’s discomfort. Additional software improvements have provided real-time updating color displays of the acquired data. This feedback is critical for alignment procedures and necessary if specific ocular structures must be scanned.

We have also developed a high-resolution OCT system which uses the broad bandwidth of argon-pumped Ti:Al2O3 fluorescence. The 130 nm bandwidth of the fluorescence, centered at 790 nm, provides a coherence length down to 3 μm. This shorter coherence length should provide a threefold increase in depth resolution. Presently, we are in the process of acquiring these higher-resolution images. Further work on this new laser source will increase the power of the fluorescence thereby reducing the acquisition time and increasing the signal-to-noise ratio of the data.

Development of the OCT system will progress toward a more clinically and patient friendly medical instrument that will not only prove to be a vital diagnosis tool for the physician, but also a multifaceted research device for the scientist. Success with this technology in both the laboratory and clinical environment has placed OCT on the verge of becoming a viable new biomedical imaging modality.

1.14.2 Optical Coherence Tomography in Ophthalmic Diagnosis

In collaboration with investigators at the New England Eye Center of the Tufts University Medical School, we have been investigating OCT as a new ophthalmic imaging instrument in both the anterior and posterior eye. Several features make OCT particularly attractive for imaging ocular tissue in comparison to existing ophthalmic examination techniques. OCT is non-contact and has superior resolution to conventional clinical ultrasound. In contrast to both scanning laser tomography and scanning laser ophthalmoscopy, the axial resolution of OCT only depends on the temporal coherence properties of the source and not on the pupil-limited numerical aperture of the eye or ocular aberrations. Micron scale axial resolution is particularly important for the early diagnosis and monitoring of degenerative retinal diseases. We have developed an OCT system which is integrated with a standard slit-lamp biomicroscope for in vivo tomography of the anterior and posterior eye. OCT probe light is directed into the eye with a pair of orthogonally mounted, computer controlled galvanometric scanning mirrors, enabling arbitrary transverse scanning patterns. The probe beam focus is coincident with the slit-lamp image plane which allows simultaneous scanning and operator visualization of the eye through either the slit-lamp ocular or an attached CCD camera. For tomography of the anterior eye, the structure of interest is placed directly at the probe beam focus. Images of the posterior eye are obtained in a manner similar to indirect ophthalmoscopy, whereby a 78 diopter Volk lens placed in front of the eye is used to relay an image of the retina to the slit-lamp image plane. A computer monitor provides a real-time display of the tomograph in progress, which for retinal images is fully updated every 2.5 seconds.

In the anterior eye, the micron scale lateral and longitudinal resolution of OCT permits highly accurate biometry of large scale ocular structures, as well as the evaluation of changes in cellular morphology associated with pathologies of the cornea, iris, and lens. High sensitivity is important to detect weak backscattering from within nominally transparent structures such as the cornea and lens. We have obtained in vivo OCT tomographs of the full anterior chamber from normal human subjects. These tomographs accurately profile the gross morphology of large-scale structures such as the cornea, iris, lens anterior capsule, and lens. Clinically relevant measurements may be obtained directly from these images, and include anterior chamber depth and angle, corneal thickness and curvature, and refractive power. These measurements have potential applications in contact lens fitting, post-cataract surgery intraocular lens implant power calculations, real-time monitoring of keratorefractive surgery, and the monitoring and quantification of keratitis, corneal edema, intraocular tumors, and angle-closure glaucoma.

Full-scale tomographs provide information on the large scale morphology of anterior segment structures. By narrowing the field of view, we can obtain high resolution OCT images of ocular microstruc-

---


ture which may provide important histopathological information concerning disease progression and the healing process in vivo. For example, in initial studies we have used OCT to evaluate the damage and healing due to laser thermerkeratoplasty (LTK) of the cornea, a relatively new therapy currently under study for the treatment refractive errors of the eye.\textsuperscript{66} An in vivo OCT tomograph taken one month post-operatively of an LTK treatment in a rabbit cornea shows an area of increased reflectivity clearly demarcating the radial extent and penetration depth of the thermal damage. Several important features of the healing process are also documented, including visualization of the single-cell layer endothelium and regrowth and thickening of the epithelium above the lesion.

The high resolution and high sensitivity of OCT in the posterior eye makes it uniquely suited for clinically relevant tomography of the human retina. We have produced high-resolution retinal tomographs of the foveal and optic disk regions of human subjects in vivo in under 2.5 s. In addition to providing important topographical information, such as optic disk cupping, these tomographs can discriminate the layered structure of the retina and normal anatomic variations in retinal nerve fiber layer thickness and retinal thickness.\textsuperscript{67} These measurements demonstrate higher resolution in the posterior eye than can be obtained with any other noninvasive imaging technique, and may be directly applicable to the early, objective diagnosis and monitoring of a variety of degenerative retinal diseases, including glaucoma, macular degeneration, macular hole, and macular edema.

In collaboration with physicians at the New England Eye Center, we have recently initiated studies of macular and optic nerve head pathology with OCT in human patients. These tomographs provide the first high-resolution cross-sectional images of retinal pathology in living human patients. OCT imaging of the foveal region appears promising for the diagnosis and monitoring of macular disease. We have obtained tomographs from patients suffering from retinal detachment, retinoschisis, and macular hole. In contrast to fundus photography or fluorescein angiography which both require relatively skilled interpretation, in each of these cases the cross-sectional view of the OCT tomograph permitted direct visualization of the detachment, loss of retina, or serous cavity, and allowed facile diagnosis of the pathology. OCT may also be useful for clinical imaging of the optic nerve head and peripapillary region, where quantitation of retinal and RNFL thickness is directly relevant to the early diagnosis and treatment of glaucoma. Unlike tonometry, visual field testing, or optic nerve head cupping, OCT measurements of nerve fiber layer thickness and loss may be able to provide an objective, early indicator of glaucoma onset.\textsuperscript{70} In preliminary studies, we have profiled retinal thickness in both normal and glaucomatous eyes with tomographs taken in circular scans around the optic disk. These images show nerve fiber layer loss in patients with severe glaucoma. Additional study is required to assess whether OCT has the measurement reproducibility to quantify the small reductions in nerve fiber layer thickness that occurs early in disease progression.

We plan to continue our investigations of clinically relevant tomography with OCT in both the anterior and posterior eye. In particular, OCT is promising in the early diagnosis of macular edema and degeneration. Macular edema manifests as increased retinal thickness, which is difficult to assess with slit-lamp observation or fundus photography, and is poorly correlated with fluorescein leakage.\textsuperscript{71} OCT can provide images of retinal thickness with micron-scale axial resolution, potentially permitting increased diagnostic sensitivity and quantitative assessment of the degree and localization of retinal thickening. The early clinical diagnosis of age-related macular degeneration, the leading cause of new blindness in the elderly in industrialized countries, depends on the detection of choroidal neovascular membranes which often form before the onset of vision loss. Although fluorescein angiography is highly sensitive to


leakage through these permeable membranes, angiography may not always accurately localize the leakage source, thus preventing the accurate delivery of photoagulation therapy. The high transverse resolution of OCT may provide a useful adjunct to angiography by pinpointing the location of the anatomic defect and aiding in the delivery of efficient therapy. These studies suggest that with further development, OCT has the potential to become a standard ophthalmic diagnostic tool for the examination of both the anterior and posterior eye.

1.14.3 Optical Coherence Microscopy

Optical coherence microscopy (OCM) is a new method for coherent confocal microscopy which uses low-coherence interferometry to enhance optical sectioning in highly scattering media. In OCM, improved confocal imaging is achieved through high detection sensitivity and high contrast rejection of out-of-focus light. OCM is an adaptation of optical coherence tomography (OCT), a technique developed in our laboratory which has recently been demonstrated for noninvasive, micron scale resolution cross-sectional imaging in the eye and other biological tissues. The extension of this technique to confocal microscopy offers the potential for micron-resolution imaging deep into highly scattering media such as skin and endoscopically accessible tissues, and may make possible the development of a new form of noninvasive "optical biopsy."

The implementation of OCM combines aspects of confocal microscopy and low-coherence interferometry. In our setup, a scanning confocal microscope is implemented using a single mode optical fiber as the confocal aperture for both illumination and light collection. The confocal microscope comprises the sample arm of a single mode fiber optic Michelson interferometer, illuminated with broadband light from a superluminescent diode source. In the interferometer, light collected from the sample is combined in a fiber-optic beamsplitter with light returning from a calibrated length reference arm, generating an interferometric signal only when the distance to a scattering site in the sample matches the reference arm length to within the source coherence length (18 μm FWHM). By matching the reference arm length to the sample arm focal plane distance, image light from the focal plane is strongly selected against scattered light which has traversed other path lengths. The operation of OCM is analogous to optical heterodyne detection providing high sensitivity and high dynamic range (> 100 dB) detection of image light.

The sensitivity of the OCM to light from the focal region falls off exponentially with axial distance, rather than geometrically as in a conventional confocal microscope. This rapid fall off is produced by the roughly Gaussian shape of the superluminescent diode source spectrum, which yields an exponential decay of the coherence function. This high rejection of light from outside the focal volume provides the potential for confocal imaging deep into scattering media with exponential attenuation such as biological tissues.

We have demonstrated the advantage of coherence gating for imaging absorbing structures embedded in highly scattering media in scattering phantoms as well as preliminary studies in biological tissues in vitro. High contrast two-dimensional images of absorbers placed in a highly scattering tissue phantom have been obtained, even though the backscatter signal at the absorber depth was overwhelmed by light scattered from other planes. Purely confocal images of the same sample contained no image information. Preliminary images of internal tissue microstructure have been obtained in mammalian gastro-intestinal tissues in vitro. These images illustrate delineation of epithelial and subepithelial layers in soft tissues and differentiation of collagenous and cartilaginous materials in heterogeneous tissue samples.

We have developed a single-backscatter theoretical model to describe coherent confocal imaging in highly scattering media both with and without coherence gating. This model has been used to identify and predict fundamental limits to the imaging capa-


bilities of both confocal microscopy and OCM in highly scattering media. Assuming typical microscope optical parameters, predictions suggest that OCM can extend the range of confocal imaging from approximately 5 to approximately 15 scattering mean-free-paths of sample depth. This range of penetration is sufficient to permit imaging of many microstructural features near the surfaces of tissues which would be of clinical and histopathological interest.

These preliminary studies show promise for obtaining acceptable contrast confocal images up to several hundred microns deep (or up to three times the depth of confocal microscopy) for high resolution imaging in medical applications. OCM may thus provide a substantial advantage for imaging without the need for biopsy in biological tissues.

1.14.4 The Ultrashort Pulse Laser Scalpel

Successful application of short optical pulses for noninvasive cutting of intraocular structures is well established. Laser induced breakdown allows photodisruption or cutting of intraocular structures without intervening surgical incisions. To date, the majority of clinical applications of laser induced optical breakdown have utilized nanosecond pulses in the millijoule energy range and single pulse exposures. However, mechanical side effects of laser induced breakdown with nanosecond sources pose potential hazards to adjacent ocular structures and tissues. Our investigations have demonstrated that significant reduction of collateral tissue damage may be achieved through the use of ultrashort pulses. These studies are part of an ongoing collaboration between investigators at MIT, the New England Eye Center of New England Medical Center Hospitals, and the Wellman Laboratories of Photomedicine at Massachusetts General Hospital.

We have studied and compared the mechanisms, scaling behavior, and tissue effects of single pulses in the nanosecond and picosecond ranges. At comparable deposited energies, the damage zones resulting from optical breakdown induced by nanosecond and picosecond pulses are comparable; however, the threshold energy for breakdown is much lower for picosecond pulses, and near-threshold picosecond pulses produce greatly reduced collateral damage zones. Ultrashort pulses with high peak intensities can produce plasma-mediated ablation of transparent tissues, such as the cornea. Picosecond and femtosecond pulse durations have been demonstrated to produce much smoother excision edges and less damage to the adjacent tissue than nanosecond pulses.

Following these initial studies, we have developed a clinically viable picosecond laser scalpel based on a modelocked, Q-switched Nd:YAG laser with external pulse selection. This laser delivers single 100 picosecond duration pulses at a repetition rate variable from 3 to 1000 Hz; each pulse produces minimal collateral damage, while multiple pulses produce a cumulative incision effect. To optimize highly localized photodisruption in transparent structures using laser induced breakdown we have also developed a new laser technology based on flashlamp pumped titanium-sapphire. Flashlamp pumped solid state lasers feature higher pulse energy and lower cost compared to cw laser pumped lasers. As described in another section of this report, the present laser produces picosecond pulses with energies in the microjoule range. Picosecond pulses in the millijoule range should become available as a result of our ongoing research.

In addition to direct modelocking we are also developing a variable pulse duration, regeneratively amplified titanium-sapphire laser. A 20 fs, 80 MHz, nJ energy Kerr-lens-modelocked oscillator is being incorporated as a seed laser for chirped-pulse regenerative amplification in the flashlamp-pumped titanium-sapphire rod. The completed laser will have variable pulse duration (20 fs - 300 ps), wavelength tunable (700-1000 nm) output. Using this source, studies will be performed to correlate tissue incision and collateral injury effects with laser parameters (pulse duration, wavelength, and repetition rate). Studies will include time resolved measurements of the fundamental physical processes in optical breakdown, as well as tissue effects. Special emphasis will be placed on investigating pulse durations shorter than 100 picoseconds in order to reduce the pulse energy needed to achieve


optical breakdown below the 70 microjoule level used in our previous study. This system will serve as a valuable and unique tool for optimizing the localizability and incision rate for ultrashort pulse intraocular laser surgery.

1.15 EUV Laser Studies

Sponsor
MIT Lincoln Laboratory
Contract BX-5098

Project Staff
James G. Goodberlet, Martin H. Muendel, Timothy A. Savas, Marc Fleury, Sumanth Kaushik, Dr. Santanu Basu, Professor Peter L. Hagelstein

During the past several years, our group has developed an experimental facility to test new low-power "table-top" EUV and soft x-ray lasers. In last year's RLE Progress Report, we described the first observations of significant gain (αL ~ 3) in Ni-like Nb at 204.2 Å. Efforts this year to improve on this result were unsuccessful, prompting us to re-evaluate our strategy for collisional lasers. The basic problem is that we require more intensity x length to achieve more gains-lengths using normal-incidence pumping; in response to this, we decided to implement a pulse-compression scheme to boost our available power by more than an order of magnitude. This conversion should be completed in the coming year. Experiments on a number of recombination lasers were performed, and these resulted in a very nice success: we observed significant gain (αL ~ 3 - 4) for the first time in H-like B at 262 Å.

1.16 Ni-like Nb Studies

Sponsor
MIT Lincoln Laboratory
Contract BX-5098

Project Staff
James G. Goodberlet, Martin H. Muendel, Sumanth Kaushik, Dr. Santanu Basu, Professor Peter L. Hagelstein

Our first successful experiments showing gain in Ni-like Nb at 204.2 Å (see figure 1) constituted a major step forward in collisional excitation lasers, bringing down by more than two orders of magnitude the pump energy requirements per gain length for this type of laser.77 Our initial attempts to improve upon these results led to damage to our preamplifier. After repairs, further attempts at extending our results were unsuccessful; nevertheless, these studies have suggested that we pursue a new strategy.

![Figure 1. Intensity versus length at 204 Å from Basu et al.](image)

After experiments at the highest power resulted in optical damage to the preamplifier, we studied our system to understand the reasons for the damage and also ways to avoid future damage. The initial results of our studies pointed to fluctuations in pulsewidth as being the prime culprit. The pulsewidth was stabilized, and new experiments were done searching for gain. The new experiments at conditions thought to be similar to our initial conditions failed to show high gain. Further experiments were done to understand what was different.

Experiments were done optimizing the line focus, varying targets between solid Nb and Nb on Si targets, varying pulsewidth, varying Z, varying the slit arrangement, and varying intensity. The pulsewidth and energy were carefully measured and remeasured; the linewidth of the focus was measured to check for beam degradation.

A key observation was that there was a considerable difference in the appearance of the data between the 1992 experiments and those in 1993. The best 1992 shots gave 1 ns streak camera images that often saturated the film, with the 204.2 Å line lasting for as long as 1.3 ns. The 1993 experiments deemed to be equivalent gave 1 ns streak camera images that often needed two shots for moderate film exposure, and the 204.2 Å line never lasted longer than 0.75 ns.

It was conjectured that the energy calibration was in error in 1992 and that instead of pumping with 1 Joule per pulse, the actual energy was higher. Experiments were conducted at higher intensity and shorter lengths to attempt to verify this. At higher intensity, the experiments at shorter length became comparable in both brightness and duration of 204.2 Å emission. Consequently, we determined that the energy calibration of the 1992 experiments was low by at least a factor of 2; in retrospect, our best 1992 data was taken at more like 2 Joules/pulse.

Theory predicts that higher total gain-lengths requires more intensity \( \times \) length product. In late 1992, initial experiments suggested that higher gain could be achieved at higher energy. This conjecture was tested by running at equal power with both 60 ps and at 120 ps pulses, with roughly comparable results at the two pulse intensities.

Although we were not able to improve upon our 1992 results in 1993, the results that we did get have very interesting consequences. If we could increase the pump power on target, then we should be able to increase the number of gain-lengths considerably. Our group has long considered the possibility of implementing a pulse compression scheme to increase the available energy. It is now clear that this, coupled with pulse shortening, will have a dramatic impact on the Ni-like Nb experiments.

### 1.17 H-like Boron Recombination X-ray Laser

**Sponsor**

MIT Lincoln Laboratory  
Contract BX-5098

**Project Staff**

James G. Goodberlet, Timothy A. Savas, Sumanth Kaushik, Dr. Santanu Basu, Professor Peter L. Hagelstein

One area of research on the MIT table-top x-ray laser involved a study of amplification in a recombining plasma of H-like boron. The recombination lasing scheme has been studied by several research groups in the field, and gain has been demonstrated at relatively low pump energies, < 6 J for H-like carbon. The spectral emission from a low Z scheme is much simpler than for a mid-Z Ni-like scheme, and data analysis is therefore simpler. Also, the ionization potential for H-like boron is about 30 percent lower than for the Ni-like Nb scheme, which would favor higher gain at lower pumping energy.

We report here a measured gain of \( \alpha = 3.72 \text{ cm}^{-1} \) in a recombining plasma of H-like boron. The low pumping energy, 1.1 J, and rapid firing rate, two

---


shots per minute, of our table-top system permitted a study of the plasma gain profile. We also mounted steel blades near the target to improve gain and tested several blade designs. We found the highest gain to be 400 μm from the target surface when a blade was positioned 500 μm from the surface.

The experimental arrangement (figure 2a) has three major components. A pumping laser system, operating at 1.05 μm, produces the x-ray laser plasma which is contained in a large vacuum chamber. X-rays emitted from the plasma are detected with a streaked concave grating spectrometer (SCGS) which consists of a near normal incidence concave grating and an x-ray streak camera.

The pump laser system® consists of a Lumonics mode-locked Q-switched Nd:YLF oscillator which produces 20 μJ, 60 ps pulses separated by 7.5 ns. A single pulse is selected from the pulse train for amplification and beam shaping. A CCD array is used to profile the pump laser beam, and the transverse intensity profile is constant to within 20 percent across 66 percent of the 1/e² beam width. Fluctuations in pulse energy and pulse duration from the oscillator were determined to be within 10 percent by monitoring fundamental and second harmonic signals as well as measuring pulse duration with a noncollinear second harmonic autocorrelator. Output from the oscillator was amplified in two stages to the 1 J level, and energy stability was determined to be within 10 percent after final amplification.

The pump beam was brought to an 8 mm length line focus on a polished solid BN target inside the vacuum chamber. The width of the line focus was measured to be 30 μm. A 100 μm x 3 mm slit was positioned 5 mm from the end of the line focus as shown in figure 2b and used to view different regions of the plasma. For most experiments, a steel blade was positioned just above the pump beam line focus as shown in the figure. Both target and blade were movable as shown. The target was translated 100 μm between shots. Plasma emission, viewed through the slit, was collected with a streaked concave grating spectrometer (SCGS) and recorded on Polaroid 667 film.

Near the target surface, a reflective steel blade was oriented to return unused pump beam radiation from the plasma back onto the original line focus (figure 2b). In addition to increasing gain via plasma cooling, we postulated that the blade might also initially heat the plasma to a higher temperature by reflecting the unused pump radiation back onto the target.

A series of experiments were performed to determine the optimal blade location and design. With the pumping energy at 1 J, we observed no gain when the blade was < 250 μm from the target surface. The 12° blade and bare target, arranged

![Figure 2a. The table-top XRL system. The x-ray laser is driven with an amplified, mode-locked Q-switched Nd:YLF oscillator. The pump laser beam is focused to a 30 μm wide line on the x-ray laser target.](image)

---

Chapter 1. Optics and Quantum Electronics

$a$ is the small signal gain coefficient and $l$ is the plasma length.

Digitized raw data from an intensity versus plasma length measurement is represented in figure 3 for four plasma lengths. The grey levels of the photos were digitally inverted for clarity. Line-outs for the spectra are also shown. The prevalent line in each photo is the 3d-2p BV line at 26.2 nm, and the weak line, noticeable in the 8 mm photo, is a NV 3s-2p line at 26.6 nm. The temporal range of each photo is approximately 1 ns. Nonlinear growth of the BV line is apparent from the raw data.

![Figure 2b](image)

Figure 2b. The table-top XRL system. A reflective steel blade is located near the target inside the vacuum chamber. Emission from the resulting plasma is viewed through an 100 μm wide slit with a streaked concave grating spectrometer (SCGS).

as shown in figure 2b, most reliably produced gain in the plasma. With the blade properly oriented, gain values over 3 cm$^{-1}$ were repeatedly measured.

The existence of gain was verified with several measurements. In addition to intensity versus plasma length measurements (see figures 3 and 4), gain was determined at several locations above the solid target surface for a constant surface-to-blade separation. Also, a comparison of emission from the 8 mm plasma source to emission from a point source was made (see figure 5).

Gain for the BV 3d-2p line was first determined by recording the line intensity, for an 1 ns time interval, at several different plasma lengths. The resulting line intensities at various lengths were fit to the Linford formula\(^\text{83}\) to estimate gain.

\[
I(l) = \exp \frac{[(\alpha l - 1)^{3/2}]}{[\alpha l \exp(\alpha l)]^{1/2}}
\]

\(\alpha\) is the small signal gain coefficient and \(l\) is the plasma length.

For all data points, the fit gives a small signal gain of $4.64 \pm .5 \text{ cm}^{-1}$. When the 6 mm data points are discarded, the fit gives $\alpha = 3.72 \pm .20 \text{ cm}^{-1}$. A data set taken several days later showed the same behavior. We believe this behavior is due to the 20 percent intensity depression at beam center, and conservatively estimate the gain to be nearer to the latter figure of $3.72 \text{ cm}^{-1}$.

In another experiment, gain was measured at different locations above the solid target surface. At each location, an 100 \(\mu\text{m}\) region of the plasma was viewed through the slit, and intensity versus plasma length measurements were made. The data, plotted in figure 6, shows an optimum region of amplification at 400 \(\mu\text{m}\) from the target surface. The optimum represents a compromise between an overdense, hot plasma near the target surface and a cold, underdense plasma near the blade. Such an optimum has been observed in other recombination lasers.\(^{84}\)

thickness and angle, the curvature at gain center can be precisely cancelled.

The 150 μm thick etalon was first placed in the oscillator cavity and the output pulse duration was measured for various etalon angles. The shortest measured pulse duration was 49 ps, which compares to 68 ps without the etalon. The measurements were made with a background-free second harmonic autocorrelator and are shown in figure 7. At the angle of greatest pulse shortening, the oscillator threshold increased by ~ 40 percent.

**Figure 6.** Gain versus distance from target. The gain was measured at several different heights above the solid target surface. The highest value of gain was recorded about 400 μm from the target surface.

**1.18 Pump Laser Conversion**

**Sponsor**

MIT Lincoln Laboratory  
Contract BX-5098

**Project Staff**

James G. Goodberlet, Timothy A. Savas, Martin H. Muenkel, Professor Peter L. Hagelstein

In order to improve x-ray laser performance, we have investigated several improvements to our pump laser system. Particularly, we have examined methods of enhanced mode-locking and chirped pulse amplification (CPA). These methods would improve the operation of our pump laser system from (1 J, ~ 60 ps) to ( > 3 J, < 15 ps). Such an increase in pump energy and intensity would substantially increase the x-ray laser gain-length products. Progress in the area of pump laser conversion, to higher energy and shorter pulse, is described in the following sections.

**1.18.1 Oscillator Enhancement**

It has been shown that an etalon can be placed in a mode-locked cavity to shorten the output pulse. Since the shortening is a result of bandwidth broadening of the oscillator, if the etalon's transmission minimum coincides with the peak gain wavelength of the oscillator, the net result is a broadening of the oscillator bandwidth. By adjusting the etalon thickness and angle, the curvature at gain center can be precisely cancelled.

The 150 μm thick etalon was first placed in the oscillator cavity and the output pulse duration was measured for various etalon angles. The shortest measured pulse duration was 49 ps, which compares to 68 ps without the etalon. The measurements were made with a background-free second harmonic autocorrelator and are shown in figure 7. At the angle of greatest pulse shortening, the oscillator threshold increased by ~ 40 percent.

**Figure 7.** Pulse shortening with an intracavity etalon. An 150 μm-thick etalon was placed in the Nd:YLF oscillator to broaden the gain bandwidth and shorten the output pulse duration. The shortest pulse measured was 49 ps which represented a 28 percent pulse width reduction.

**1.18.2 Chirped Pulse Amplification**

A method to shorten the pump laser pulse duration and increase the pulse energy is chirped pulse amplification (CPA). This method is based upon self-phase modulation (SPM) in nonlinear material, which increases pulse bandwidth. After bandwidth broadening, the pulse is stretched before amplification and then recompressed to a shorter pulse after amplification. The pulse stretching allows higher amplification before approaching the amplifiers intensity limit. Employing this method, we anticipate achieving pump pulse energies of 3 J and duration of 15 ps.

Since self-phase modulation also gives rise to self-focusing in bulk material, the pulse bandwidth cannot be arbitrarily broadened by increasing the material length. The self-focusing distance can be increased by increasing the beam waist, but this increase will also reduce the spectral broadening.
Chapter 1. Optics and Quantum Electronics

factor. We have optimized the beam waist and material length for our input pulse duration and energy. For our system, the beam waist in the SF6 is 550 μm, and the material length is 2.3 cm.

The first phase of converting our pump laser has been to generate increased bandwidth via self-phase modulation (SPM) in bulk glass and to compress the frequency broadened pulse. Using 60 ps, 8 mJ pulses which are focused to an intensity of $10^{10}$ W/cm$^2$ in SF6 glass, we have observed spectral broadening of the pulse bandwidth by more than a factor of two and compression of the pulse duration by a factor of three. Due to unavoidable optical losses, the net increase in pump intensity presently achieved is a factor of 2.4. We anticipate an overall intensity enhancement factor of 10 after completing the conversion.

Measurements of the oscillator pulse and compressed pulse from a SHG autocorrelator are shown in figures 8a and 8b. In agreement with theory, the pulse narrows and develops side lobes as the gratings are separated. The measured compression ratio is a factor of three. The measured compression ratio suggests that the pulse bandwidth has been broadened by at least a factor of three.

1.19 Development of a Densitometer

Sponsor
MIT Lincoln Laboratory
Contract BX-5098

Project Staff
Marc Fleury, James G. Goodberlet, Martin H. Muenchel

Spectral and temporal features of our laser plasma and x-ray emissions are recorded on film in conjunction with an x-ray streak camera. Two types of film are used for our analysis: Polaroid 667 and Kodak TMZ P3200. The Polaroid film has a dynamic range of 100 and is analyzed using a conventional flatbed scanner from which intensity information is obtained. The Kodak TMZ P3200 is a transparent film that has a much larger dynamic range (≈ 1000). However, since it is transparent, it is not well suited for flatbed scanning. Consequently, we have designed and built a densitometer to analyze the intensity data recorded on this film.

The densitometer consists of a 5 mW He-Ne laser focused to a slit of 500 μm by 2 inches. This slit is scanned across the film using a translation stage and the transmitted light is focused and collected by a photodiode. The position of the translation stage and the intensity are read by a PC using data acquisition hardware and software. The densitometer is sketched in figure 9.

---


130 RLE Progress Report Number 136
1.20 Quantum-Well X-ray Detector

Project Staff
Sumanth Kaushik, Professor Peter L. Hagelstein

We have designed a soft x-ray detector using multiple quantum wells that is designed to operate in the 70-500 eV range. The novel feature of this proposed detector is that it is based on optical effects, which have the potential to lead to excellent sensitivity and good spatial, temporal, and energy resolutions. The detector characteristics are summarized in table 1. For energies greater than 500 eV, the detector's energy resolution deteriorates, but its other features remain essentially unaltered. This kind of detector is presently unavailable; the development of such a device would have important applications in the areas of x-ray microscopy and plasma diagnostics.

The detector (see figure 11) is based on the concept of optically imaging the x-ray induced free-carriers in semiconductor quantum wells. The general idea is to select a probe beam tuned near the onset of the heavy hole excitonic peak where the change in the optical susceptibility is the largest. The lengths of the various layers are chosen in such a manner that under quiescent conditions, the reflection is near minimum (typically 5 to 10 percent). The presence of x-ray induced carriers modifies the optical susceptibility, thereby modulating the optical impedance of the multiple-

---

Figure 9. Schematic design of the densitometer.

Figure 10. Intensity versus wavelength from four shots of a H-like Boron plasma.

In figure 10, we show data obtained from our densitometer. The data traces represent the intensity as a function of wavelength for emission from a H-like Boron plasma. Plotted curve is that obtained from integrating the outputs from four shots. From these curves, we have computed the dynamic range of the densitometer to be 1300 with a resolution of $\lambda/\Delta\lambda = 250$.

---

quantum-well (MQW) structure. The modulation in the impedance appears as reflected light and is imaged on a CCD (charged coupled device) camera. By exploiting the natural resonances in the transmission and reflection coefficients in a multi-layer dielectric stack, contrast ratios as high as 10 to 20 percent can be achieved.

<table>
<thead>
<tr>
<th>Design Parameters</th>
<th>Value</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_4$ : Al$<em>x$GaAs$</em>{1-x}$</td>
<td>50Å</td>
<td>± 2.4Å</td>
</tr>
<tr>
<td>$L_w$ : GaAs</td>
<td>170Å</td>
<td>± 2.4Å</td>
</tr>
<tr>
<td>$L_5$ : Al$<em>x$GaAs$</em>{1-x}$</td>
<td>48.6 nm$^*$</td>
<td>± .5 nm</td>
</tr>
<tr>
<td>$\lambda_p$</td>
<td>818.55 nm</td>
<td>± 1 nm</td>
</tr>
<tr>
<td>$\theta_{in}$</td>
<td>0°</td>
<td>± 5°</td>
</tr>
<tr>
<td>$T$</td>
<td>150 K</td>
<td>10 K</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Detector Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Resolution: 1μ</td>
</tr>
<tr>
<td>Temporal Resolution: 20 ps</td>
</tr>
<tr>
<td>Sensitivity: 25 photons/μm$^2$</td>
</tr>
<tr>
<td>Energy Resolution: 85 eV</td>
</tr>
<tr>
<td>Contrast Ratio: ~ 10%</td>
</tr>
</tbody>
</table>

Table 1. Summary of the chief parameters and performance of the detectors. The definitions of the variables are found in the text. Note: Actual $L_w$ values can also be half-integer multiples of the probe wavelength added to the minimum value specified above.

Figure 11. Schematic of the multilayer stack. With each interface and layer, a transmission matrix is associated. The overall transmission matrix is expressed as a product of individual matrices. The cap layers, usually GaAs, prevent the oxidation of the surface and are typically very thin (~ 100 Å) and do not affect the optical properties of the quantum wells significantly.

---


132 RLE Progress Report Number 136
1.21 Lineshape Theory

We have studied the origin of non-Lorentzian features in the absorption lineshapes of excitonic transitions in semiconductors. One very pronounced feature of the absorption lineshape is that for energies well away from the resonance, the lineshape is found to decrease exponentially as a function of energy. This result, known as the Urbach's rule, was first observed experimentally, and is attributed to interaction of excitons with LO-phonons.

The principal result of our investigation is a new lineshape function that accounts for the interactions with LO phonons. Specifically, we have found that the lineshape function can be written as

\[
\Delta(E) = \frac{2\Gamma_0 t_0}{\sqrt{\Gamma_0 + t_0^2}} e^{\Gamma_0 t_0 K_1(t_0(E^2 + \Gamma_0^2)^{1/2})} \]

where

\[
t_0 = \frac{\Gamma_0}{\sum_{\lambda \sigma} |M_{\lambda \sigma}(q)|^2}
\]

The function \(K_1\) is the Bessel function of the second kind. The parameter \(\Gamma_0\) is the Lorentzian linewidth as measured or as computed using standard linewidth formulas. The parameter \(M_{\lambda \sigma}(q)\) is the phonon matrix element between the 1's exciton ground state and the excited state.

This lineshape function has the important property that it is Lorentzian near the resonance energy and decreases exponentially in energy for energies well away from the resonance (see figure 12). That is for \(t_0(E^2 + \Gamma_0^2)^{1/2}\) large:

\[
\Delta(E) \sim e^{-t_0E}.
\]

(Urbach)

This exponential dependence appears to be in agreement with the well known Urbach's rule. Although the exponential tail has been shown analytically to be a consequence of exciton-phonon interaction, the earlier arguments were based on studying asymptotic properties of the Green's functions. A simple analytical lineshape function of the form given in equation (1) has been hitherto lacking.

Our result is derived by writing the Green's function in the form

\[
G(t) = -i\theta(t) \exp(F(t)) = -i\theta(t) \exp(\Gamma(t) - \Sigma(t))
\]

where the function \(F(t)\) is derived using cluster expansion method. By noting the following very important property of \(\Gamma(t)\) and \(\Sigma(t)\) under time inversion

\[
\Gamma_a(t) = \Gamma_a(-t)
\]

\[
\Sigma_a(t) = -\Sigma_a(-t)
\]

the functions \(\Gamma(t)\) and \(\Sigma(t)\) can be approximated rather accurately. The Green's function is Fourier transformed to yield the lineshape quoted in (1).

1.22 Lattice-induced Reactions

Sponsor

Electric Power Research Institute
Contract RP3170-25

---

We have pursued mechanisms through which large energy transfer between a lattice and nuclei can occur. Energy transfer through conventional recoil mechanisms is well-known and does not lead to any new anomalous effects.

We investigated anomalous energy transfer through phonon mode frequency shifts in the case of neutron transfer reactions, and found that associated with a mode frequency shift of $\delta \omega$, there was a corresponding energy transfer of

$$\Delta E = N \hbar \omega$$

where $N$ is the number of phonons in the modes that shift. In the case in which the phonon modes jump across a band gap, the energy exchange per phonon can be on the order of a few meV. In the case in which the phonon mode is a continuum mode, the number of phonons can in principle be very great, resulting in anomalously large energy transfer. For example, if somehow $10^8$ phonons were initially in a gap-jumping mode, the total energy transfer could be as large as fractions of an MeV.

Recently, we have generalized this basic mechanism to the case of gap-jumping between impurity phonon bands due to host vacancies in metal hydrides. For example, in the case of PdH, it is known that the basic octahedral distribution of hydrogen atoms around a Pd atom within a unit cell is preserved in the presence of a Pd vacancy. In this case, the hydrogen atoms see a considerably softer potential, and we speculate that this leads to impurity phonon modes within the band gap. If this is true in the case of a PdD lattice, as illustrated in figure 13, then some rather interesting physics could result.
In this case, any process that creates a host lattice vacancy, or else changes the number of deuterons around a host lattice vacancy, will cause a change in the number of phonon modes in each impurity band. Large excitation of these modes will result in anomalous energy transfer associated with vacancy production. We have formulated a theory for this process and computed rates for a variety of decay channels.

For example, if the lowest optical phonon modes were very highly excited, then the production of a Pd vacancy would down-shift the highly excited modes, leading to energy transfer from the lattice to whatever mechanism causes the vacancy. In this case, the lattice energy can drive reactions that cause a vacancy to be produced. A dominant decay channel at large energy transfer is fast electron production, mediated by the Coulomb interaction between the nuclei and K-shell electrons. The decay rate can be calculated directly from

\[ \Gamma = \sum \sum \frac{2\pi}{h} | \langle L_{\text{f}}, 1s | \frac{Ze^2}{r_{12}} | L_{\text{f}}, k \rangle |^2 \]

\[ \delta(E^{(L)} + \hbar^2 \frac{k^2}{2m_e + l_k} - E^{(L)}) \]

Performing the integration over electronic coordinates leads to

\[ \Gamma = \frac{2\Pi}{h} \sum \sum | V(k) |^2 | \langle \psi^{(L)}(q) | e^{-i\delta_{qk} R_i} | \psi^{(L)}(q) \rangle |^2 \]

\[ \delta(l_k + \hbar^2 | k^2/2m_e + E^{(L)} - E^{(L)}) \]

where \( V(k) \) describes the Coulomb interaction

\[ V(k) = \frac{Ze^2}{\sqrt{V}} \left[ \frac{N_{1s}}{\sqrt{V}} \right] \frac{1}{\hbar^2} \left[ \frac{Z^2}{\pi a_0^3} \right] \frac{4\pi}{k^2 + (Z/a_0)^2} \]

The conventional recoil is accounted for through the appearance of \( e^{ik \cdot R_i} \) and \( e^{-i\delta_{qk}} \) is a Duschinsky operator that includes the effects of changes in the structure of the phonon modes.

Since only a few modes jump a band gap, these modes can be used to evaluate the Duschinsky matrix element approximately in the absence of recoil; this leads approximately to

\[ \Gamma = \frac{2\Pi}{h} \sum \sum | V(k) |^2 | \langle \psi^{(L)}(q) | e^{ik \cdot R_i} | \psi^{(L)}(q) \rangle |^2 \]

\[ \delta(l_k + \hbar^2 | k^2/2m_e + NH \delta \omega) \]

in the limit that the recoil has been sufficiently strong to create a new vacancy. If the anomalous energy transfer is large, then the recoil energy can be neglected, and the net effect of the recoil term is to determine what fraction of the time a new vacancy is created. If the probability of vacancy creation is | \( T(k) \) |^2, we obtain the final result

\[ \Gamma = \frac{24N_{1s}}{\pi Z} \frac{1}{\hbar} \left[ \frac{\varepsilon}{l_H} \right] \frac{1}{2} \frac{1}{(1 + (\varepsilon/Z^2 l_H))^2} | T(k) |^2 \]

where \( \varepsilon = NH \delta \omega - l_L \). This is plotted below in figure 14 for recoils from Pd and in figure 15 in the case of recoils from D.

These rates are observed to be quite fast. They are sufficiently rapid that it is a priori difficult to understand how such large phonon mode populations could build up in the presence of such fast reaction rates, especially if other decay modes are present that turn on at lower energy. Our picture is that many phonon modes are excited; modes not immediately next to the band gap can build up without the same limits as the lowest modes. When the lowest modes jump, they reveal new lowest modes which in turn may jump; these new modes may be quite highly excited, resulting in the possibility of large energy transfer. The associated rates as calculated by Fermi's Golden Rule are too large to be physical, implying that the decays will occur in bursts as a high-order Raman-type process.

![Figure 14. Predicted fast electron emission rate from deuterons as a function of lattice energy transfer.](image-url)
Of immediate interest are the decay rates for Coulomb-induced deuteron-deuteron recoil (shown in Figure 16), which is a dominant decay mechanism up to a few KeV of lattice energy transfer. This process results in the production of fast deuterons, up to energy transfer where fast electron decay as described above turns on. This is interesting because these deuterons are sufficiently energetic to cause fusion reactions to occur. The fusion rate per deuteron pair is shown in Figure 17; these rates are observed to be sufficiently high to account for the observation of dd-fusion neutrons claimed to have been observed in cold fusion experiments during the last few years.

The reactions described above are based on the premise that a very large number of phonons can be developed in a single phonon mode. The situation is much like the difference between broad band light from a hot thermal source, and the radiation field inside a laser cavity. The presence of photon gain inside the laser cavity results in the development of very large phonon populations in a single mode. The laser amplifier will add photons preferentially to modes within the gain linewidth that already have the most photons.

One way to generate large phonon populations in phonon modes is to create a phonon amplifier that will drive a "phonon laser". Phonon lasers have been built, but they have not operated previously in optical phonon modes. We are considering a novel mechanism to drive a phonon amplifier. If exothermic chemical reactions were to occur at a solid surface (at an interface with a gas or solid), and if the reactions involved an inhibiting potential barrier that phonons could help to overcome, then the reactions would tend to couple energy preferentially into the phonon modes that helped to stimulate the reaction. If the end product of the reaction were not stable at the surface, then a population inversion would be developed between the reactants and their products; this is the condition for phonon gain.

One interesting example of this mechanism is hydrogen desorption from a metal hydride. Absorbed and adsorbed hydrogen in metal hydrides is often localized as single atoms in local wells that occur in pockets between host lattice atoms. Gas desorption involves the tunneling of the hydrogen atoms to the bound molecular state away from the metal surface, where it is free to leave. In PdD, deuterium desorption is exothermic at very high D to Pd loading, which is where the anomalies are thought to occur.

In previous RLE reports, we have described at length novel neutron transfer reaction mechanisms. We would do so here as well, except that our report is already much too long. During the last year we have found that our model is very closely related to the Anderson model in condensed matter physics, that it can be solved approximately using perturbation theory, and that the reaction rates for the specific E1 and M1 transfer mechanisms described last year are too small to be observable. The reason that the rates are too small is that the electromagnetic coupling is very weak. A modified version of the theory that uses strong interaction neutron matrix elements, which in fact are a much closer...
analog to the equivalent electron matrix elements normally used in Anderson model studies, results in much stronger coupling and very large reaction rate estimates. This model looks very promising; it predicts heat due to neutron transfers from $^{106}$Pd to $^6$Li at 156 KeV lattice energy transfer per reaction to be the dominant heat-producing reaction in Pons-Fleischmann experiments. This work is currently being documented.
Professor Jeffrey H. Shapiro, Associate Head, Department of Electrical Engineering and Computer Science
Chapter 2. Optical Propagation and Communication

2.1 Introduction

The central theme of our research programs has been to advance the understanding of optical and quasi-optical communication, radar, and sensing systems. Broadly speaking, this has entailed: (1) developing system-analytic models for important optical propagation, detection, and communication scenarios; (2) using these models to derive the fundamental limits on system performance; and (3) identifying and establishing through experimentation the feasibility of techniques and devices which can be used to approach these performance limits.

2.2 Squeezed States of Light

Sponsor

Maryland Procurement Office
Contract MDA 904-90-C-5070
Contract MDA 904-93-C-4169

Project Staff

Professor Jeffrey H. Shapiro, Dr. Ngai C. Wong, Dr. Lance G. Joneckis, Scott R. Shepard, Ke-Xun Sun

The central focus of our research has been gain-saturated operation of an optical parametric amplifier (OPA). Under gain saturation, the amplified output intensity of an injected coherent-state signal becomes amplitude squeezed and the large-signal signal-to-noise ratio improves, an operating regime known as squeezed amplification. We have performed a detailed study of the mean-field characteristics, gain saturation, and quantum-noise spectra of an injection-seeded OPA including arbitrary cavity detunings of the three interacting waves. Comparison with our mean-field experimental results using a type-I phase-matched LNO3 OPA has shown excellent qualitative agreement. We have also successfully analyzed a thermal-hysteresis effect encountered in our experiment that can severely impact the tuning capability of the OPA.

As part of our squeezed-amplifier theory, we have established the quantum form of the Gaussian moment-factoring theorem and used this theorem to deduce the direct-detection photocurrent noise spectrum from a squeezed amplifier in the Gaussian-state approximation. We have also used this Gaussian-state formalism to explicitly unify the analyses of the three principal nonclassical manifestations of signal-idler state entanglement, i.e., nonclassical signal/idler photon correlation, nonclassical quadrature-noise squeezing, and nonclassical...
Chapter 2. Optical Propagation and Communication

signal/idler fourth-order interference. This work clarifies previously ill-understood limits on the semiclassical theory of fourth-order interference. In particular, we show that a quantum signature is demonstrable, in the low-photon-flux regime, at fringe visibilities substantially below 50 percent.

In another theoretical study, we have continued our development of quantum propagation theory for single-mode fiber. Using a coarse-grained time approach based on a phenomenological Kerr-effect time constant, we have been working to evaluate the homodyne-measurement noise spectra for a variety of input and fiber conditions. We have obtained results for continuous-wave (cw) inputs applied to dispersionless fiber in the absence and presence of loss. The latter case was handled using the terminated-cumulant expansion (TCE); we are now using this same approach to find the noise spectrum for dispersive propagation with a cw input. We have also begun exploring the value of local-oscillator optimization in pulsed-input homodyne measurements.

2.2.1 Publications


2.3 Multiresolution Laser Radar Range Profiling

Sponsor
U.S. Air Force - Office of Scientific Research
Grant F49620-93-1-0604

Project Staff
Professor Jeffrey H. Shapiro, Irene Fung

This effort is part of a collaboration on automatic target detection and recognition, with Professors Alan S. Willsky (MIT Laboratory for Information and Decision Systems) and W. Eric L. Grimson (MIT Artificial Intelligence Laboratory) and their students. The unifying theme of the collaboration is the use of multiresolution (wavelet) methods at every stage—from sensor front-end processing, through feature extraction, to the object recognition module—in an overall system.

Our group's activity centers on the application of multiresolution techniques to laser radars. Previously, we have studied maximum-likelihood (ML) estimation for planar range-profiling with a laser radar.7 The importance of the ML approach lies in its ability to suppress the range anomalies caused by laser speckle. The practicality of the ML approach derives from the utility of the expectation-maximization (EM) algorithm for this problem. In initial work on the multiresolution range-profiling problem, we have shown that the same EM approach can be applied, at any desired wavelet

---


scale, to perform ML profiling. We are now beginning performance studies on this scheme using Haar wavelets to fit simulated laser radar range data.

2.4 Optical Frequency Division and Synthesis

Sponsors
Charles S. Draper Laboratories  
Contract DL-H-441698  
MIT Lincoln Laboratory  
Contract CX-16335  
National Institute of Standards and Technology  
Grant 60-NANGD-1052  
U.S. Army Research Office  
Grant DAAL03-90-G-0128  
Grant DAAH04-93-G-0399  
Grant DAAH04-93-G-0187

Project Staff
Dr. Ngai C. Wong, L. Reginald Brothers, Bobby Y. Lai, Dicky Lee

Frequency division and synthesis in the optical domain play an important role in modern optical precision measurements, optical frequency standards, and coherent optical communication. We have demonstrated tunable optical frequency division using an optical parametric oscillator (OPO) approach based on an efficient, one-step parametric downconversion process. An OPO converts an input pump into two intense, coherent subharmonic outputs whose frequencies are tunable and whose sum frequency equals the pump frequency. By phase locking the output frequency difference to a microwave source, the output frequencies are precisely determined, and the OPO functions as an optical frequency divider. OPO frequency dividers can be operated in series or in parallel to measure, compare, and synthesize frequencies from optical to microwave, with high precision and resolution.

2.4.1 Terahertz Optical Frequency Comb Generation

In order to facilitate difference frequency measurements in the THz range, we have developed an optical frequency comb generator based on an efficient electro-optic phase modulator design. By incorporating a microwave waveguide resonator structure in a LiNbO$_3$ electro-optic modulator, the phase velocities of the microwave and optical fields can be matched to maximize the electro-optic modulation at a user-specified microwave frequency. The modulation is further enhanced by placing the modulator inside an optical cavity that is resonant for the input optical beam and generated sidebands. For 1 W of microwave power at 17 GHz, we have obtained an optical frequency comb with a 3-THz span, as shown in figure 1. We have employed the terahertz comb to measure a difference frequency of 1.45 THz between a YAG laser and the signal beam of an OPO with a measurement uncertainty of 1 MHz, limited only by the stability of the two optical sources. In addition to optical frequency metrology, terahertz optical frequency comb generation is potentially useful for frequency identification in a wideband optical communication network.

2.4.2 Tunable Optical Parametric Oscillator

In our first demonstration of a tunable optical frequency divider, the signal-idler difference frequency of a type-II phase-matched KTP OPO was phase locked to a microwave source at 12.3 GHz with excellent signal-to-noise ratio. The tuning range was limited by the 25-GHz bandwidth of the high-speed photodetector. For many applications, a tuning range that is much larger than the detector bandwidth is desirable. We have made use of the terahertz optical frequency comb generator to phase lock the KTP OPO at a difference frequency of 665 GHz. The idler beam was used to drive the terahertz optical frequency comb generator; the resultant beat frequency between the signal and the 39th upper sideband of the comb was detected by a 1-GHz photodetector. We have measured a residual phase noise spectral density of the phase-

---


Chapter 2. Optical Propagation and Communication

Figure 1. Optical spectrum analyzer trace of the optical frequency comb, with a mode spacing of 17.05 GHz. The inset shows an amplified portion of the spectrum.

The high spectral purity obtained in the experiment shows that phase locking at an even higher difference frequency is possible.

We have also developed a cw dual-cavity doubly resonant OPO (DRO) that provides a wider tuning range than the usual single-cavity OPO. Because of the stringent requirement of simultaneous cavity resonance of the signal and idler waves in a single-cavity DRO, mode hops usually occur if the cavity length is tuned by more than 1-2 nm, or ~ 10 MHz in output frequencies in our DRO system. To overcome this lack of tuning capability, signal and idler waves are separated internally by a polarizing beam splitter and resonated in separate cavities. In this way, the threshold remains low while continuous frequency tuning is achieved by varying the two cavities separately. Our preliminary results show that the cavity length could be tuned over 60 nm, or 600 MHz, in the output frequencies. The threshold for the dual-cavity DRO was 55 mW compared with 47 mW for a single-cavity DRO. The tuning range of the dual-cavity DRO is expected to be limited by the angle phase matching bandwidth of the KTP crystal of ~ 50-100 GHz. This development may lead to a wider use of cw OPOs as convenient and tunable radiation sources for many applications.

2.4.3 Publications


Lee, D., N.C. Wong, and L.R. Brothers. "Frequency Tuning and Stabilization of a KTP Optical

---


2.5 Analog Processing of Optical Wavefronts Using Integrated Guided-Wave Optics

Sponsor
U.S. Air Force - Office of Scientific Research
Contract F49620-90-C-0036

Project Staff
Dr. Robert H. Rediker, Boris Golubovic

In wavefront sensing and correction, it is envisioned that $10^3 - 10^4$ basic modules would be used. In integrated optics, as in integrated circuits, it is important to relax the requirements for individual components while requiring that operation of the integrated optics (circuits) is independent of significant component variations. The wavefront is sensed by interferometers between the multiplicity of through waveguides with the arms of the interferometers evanescently coupled to the adjacent waveguides. The input powers to the interferometer arms will not be equal as a result of (1) the input power to the waveguide array being non-uniform and (2) unequal coupling by the evanescent couplers.

In RLE Progress Report No. 135, we stated that the experimental results on the phase measurement and phase correction were independent of the power difference between the interferometer arms up to a ratio of greater than 50:1. We also wrote in last year's progress report that it was decided in 1992 to embark on building a complete optical module. The on-chip detector at the output of the interferometer was to also be incorporated in this planned module. In the previous work, the detector has been off chip. In this past year, a semiconductor processing technique was developed and used so that fabrication of all components (e.g., waveguides, couplers and detectors) of the module was compatible. The complete optical modules were fabricated; their performance was measured, and compared favorably with theoretical predictions.

We can now report in more detail on the components of the complete module. The phase on the output of the through waveguides was varied using a reverse biased p-n-p-n structure incorporated into a portion of the waveguide. A small dither was also impressed on the phase. A Y-junction interferometer was used to obtain the relative phase difference between the two waveguides. Integrated MSM photodetectors, 17x69 mm in size with total internal reflection mirror coupling, were used for signal detection. Both the interferometer and detector were fabricated and operated between the two 30 μm-separated through waveguides. The phase dither detection system made it possible to adjust appropriately the relative phase between the outputs of the two through waveguides, largely independent of the power ratio between the individual guides.

2.5.1 Publications


Chapter 3. High-Frequency (> 100 GHz) and High-Speed (<1 ps) Electronic Devices

Academic and Research Staff
Professor Qing Hu

Visiting Scientists and Research Affiliates
Dr. Thomas Schäpers, Dr. Simon Verghese

Graduate Students
James Ernstmeyer, Rajesh K. Gupta, Brian R. Jacobson, Jurgen H. Smet, Rolf A. Wyss

Undergraduate Students
David K. Chow, Elliot E. Hui

Technical and Support Staff
Barbara A. King

3.1 Facility for Millimeter-wave and THz Frequencies
Professor Hu's laboratory is equipped with various millimeter-wave and infrared sources which can generate coherent and incoherent radiation up to 30 THz. These include: Gunn oscillators at W-band frequencies (75-110 GHz); a frequency doubler, tripler, and quadrupler using Schottky diodes at 200, 300, and 400 GHz; an optically pumped far-infrared laser which generates coherent radiation up to 8 THz; and an infrared Fourier transform spectrometer which is capable of performing linear spectroscopy from 45 GHz to 30 THz and beyond. This laboratory is also equipped with various cryogenic millimeter-wave and infrared detectors. These include: Si composite bolometers, InSb hot-electron bolometers, superconductor-insulator-superconductor (SIS) receivers, and high-Te Josephson detectors. There are many infrared cryostats which can cool the devices from 0.3 K to 77 K.

3.2 Far-Infrared Studies of Antenna-coupled Quantum-effect Devices

Sponsor
MIT Research Laboratory of Electronics
Postdoctoral Fellowship
National Science Foundation
Grant DMR 90-22933

Project Staff
Rolf A. Wyss, Dr. Thomas Schäpers, Professor Qing Hu, in collaboration with Professor Jesús A. del Alamo and Jülich GmbH Company

Quantum transport has been one of the most active fields in solid-state physics in recent years. Advances in material preparation have made quantum phenomena profound in electron transport for many semiconductor quantum devices such as quantum point contacts, quantum dots, quantum wires, quantum wells, superlattices, etc. In clean samples and at low temperatures, electrons can travel through the whole sample without suffering phase-destructive scattering. Extensive work has been done to study various features of such phase-coherent quantum transport. However, the experiments reported so far are limited to measurements by dc transport and far-infrared spectroscopy.

---

1 Jülich, Germany.
Chapter 3. High-Frequency and High-Speed Electronic Devices

It is well known in the field of superconducting tunneling that photons can assist the tunneling process, provided that the tunneling is elastic. In a broad sense, elastic tunneling is a phase-coherent quantum transport process in a classically forbidden region. Therefore, all the results of photon-assisted tunneling can be applied to the study of photon-assisted quantum transport in semiconductor devices. This will provide a new dimension to study the exciting quantum transport phenomena. Novel long-wavelength optoelectronic devices may also emerge from this research.

In this project, we study the interaction between far-infrared photons and ballistic electrons in quantum point contact devices, whose schematic is shown in figure 1a.

We have fabricated several antenna-coupled quantum point contact devices using a combination of optical and electron-beam lithography. Figure 2 shows the SEM photographs of one of the devices. The split-gate electrodes (the vertical leads) also serve as the terminals of a log-periodic antenna, whose function is to couple the far-infrared radiation of millimeter wavelength to the point contact of submicron dimensions.

Figure 1. (a) Schematic of a quantum point contact. (b) and (c) Illustration of photon-assisted over-the-barrier quantum transport.

Figure 2. (a) SEM (with a magnification of 400) of a quantum point contact with a log-periodic antenna. (b) Central region of the quantum point contact. The opening of the split-gate electrodes is 0.15 μm.

The dc transport measurement of the drain/source transport showed fifteen quantized conductance steps, indicating the high quality of the devices. Under coherent far-infrared radiation at 285 GHz, a
pronounced photon-induced drain/source current is produced throughout a gate voltage range in which the device exhibits the behavior of a one-dimensional electron system. This photon-induced current is attributed to a bolometric effect as described in the previous report.

In addition to the photoconductive effect, we have recently discovered a pronounced photovoltaic effect in several antenna-coupled quantum point contacts irradiated from 280 GHz to 2.5 THz. Figure 3a shows the drain/source current of an irradiated quantum point contact without external drain/source bias. This induced current tracks the subband structure in a very regular manner that is similar to the transconductance of the device measured without radiation. We have identified this photovoltaic effect as a result of thermopower due to an asymmetric heating of the source and drain. This analysis is supported by experimental evidence that, by shifting the focal spot of the radiation, we can reverse the polarity of the photon-induced current, as shown in figure 3b. This is a direct observation of thermopower in a one-dimensional electron system.

We are currently fitting the measured photon-induced current with quantum transport theories by taking the finite drain/source thermal conductance into account and by using a nonlinear gate voltage versus Fermi energy relation.

To maximize the photon-assisted quantum process, we have recently started investigating the possibility of photon-excited bound-to-extended-state transition in an antenna-coupled lateral quantum-well device (or PRESTFET), as shown in figure 4. In this device, the dipole-transition matrix element is on the order of the well width, which is much greater than that in a quantum-point-contact device. Furthermore, the drain/source resistance of a lateral QW device is on the order 100 $\Omega$, which is much smaller than that of a quantum point contact, making impedance matching with the planar antennas much easier. Vertical quantum-well versions of this device have been well developed for detection applications around 10-$\mu$m wavelength region, and video cameras with $128 \times 128$ pixels have been made with temperature sensitivities around 10 mK.

Figure 3. (a) Drain/source current of an antenna-coupled quantum point contact irradiated at different power levels at 280 GHz with zero external drain/source bias voltage. (b) Same as (a), except the far-infrared radiation spot is shifted.

Figure 4. (a) Schematic of an antenna-coupled lateral quantum-well device irradiated at far-infrared frequencies. (b) Illustration of photon-excited bound-to-extended-state transition in these devices. Clearly, the operating frequencies of these devices can be tuned electrically by changing the barrier heights.
3.3 Femtosecond Dynamical Studies of Quantum Devices for Ultrafast Electronic Applications

Sponsor
MIT Lincoln Laboratory
Advanced Concept Program

Project Staff
Dr. Simon Verghese, Professor Qing Hu, in collaboration with Dr. Elliot Brown

Solid-state physics and electrical engineering communities are witnessing a trend in which electronic devices are becoming smaller and their speeds are becoming faster. This trend is driven by modern society's insatiable appetite for increasing capacity for information processing and transmission. As a result, the size of the smallest features of state-of-the-art devices has decreased to just several hundred angstroms. At this length scale, intersubband spacing and Coulomb interaction energy are comparable to the Fermi energy, and consequently the devices behave as quantum mechanical systems (or artificial atoms). In a related project (see section 3.2), we have studied transport properties of antenna-coupled quantum point contacts under coherent far-infrared radiation whose frequencies are comparable to the intersubband spacing in the point contact. In this ACC project, we will carry out transport studies in time domain by using subpicosecond electrical pulses generated by Auston switches to study the dynamical processes of electrons in quantum-effect structures. This work will substantially enhance our understanding of the dynamics of quantum structures, and novel ultrafast devices may also emerge from this research.

Our past work in frequency domain has yielded new information about quantum point contact devices (such as radiation-induced current and thermopower), and we intend to continue that effort to improve the device parameters (mainly to use better antenna structures to minimize heating and maximize the quantum process). A better approach would be to study the response of the quantum devices in time-resolved fashion by using a pump-and-probe method with a pulsed laser. A 100-fs laser pulse contains frequency components up to 10 THz, which should enable us to perform spectroscopic studies on quantum devices over a broad frequency range that covers all the interesting energy levels, namely the intersubband transition and Coulomb interaction energies. Furthermore, unlike a conventional far-infrared Fourier transform spectrometer, the brightness of a pulsed laser source will enable us to perform nonlinear spectroscopy. In addition, the short duration of the laser pulses minimizes heating effect and enhances quantum mechanical processes. Finally (and perhaps their most attractive feature), time-resolved studies can reveal information such as the time scale of transport process in quantum devices. One of the basic issues in determining the potential applications of quantum devices, this method may eventually answer some fundamental questions such as the traversal time of tunnel, which have intrigued physicists since the initial discovery of quantum mechanics.

We will pursue two types of measurements in this project. The first one involves a far-infrared spectrometer pumped by a mode-locked Ti:Sapphire laser which we have constructed recently, as shown in figure 5. In this set up, two antenna-coupled Auston switches are pumped by a Ti:Al$_2$O$_3$ pulsed laser. The THz electrical pulses are launched into free space and then combined by a beam splitter. The combined beam is a superposition of two coherent subpicosecond electrical pulses whose relative time delay can be varied. These combined pulses will then be focused onto an antenna-coupled quantum device to generate a dc electrical current in the device through the photon-assisted transport/tunnelling process. The generated electrical current is proportional to the time autocorrelation function of the whole system, whose Fourier transform will give the frequency response of the system. This measurement is essentially the same as that made with a conventional far-infrared Fourier transform spectrometer (FTS), which uses a Hg-arc lamp as its source. There are some major advantages of the pulsed laser system, however. First, its source is much brighter than that of the FTS so that far-infrared nonlinear spectroscopy can be performed. Second, the short duration of the laser pulses minimizes the heating effect and thus enhances the relative strength of the photon-assisted quantum process.

---

2 MIT Lincoln Laboratory, Lexington, Massachusetts.
Chapter 3. High-Frequency and High-Speed Electronic Devices

The second type of experiment involves a pump-and-probe method, by integrating two Auston switches monolithically with the quantum-effect (or Josephson) devices, as shown in figure 6. One switch will be used to apply a short electric pulse on the input, and the second one will be used to probe the output electrical current in a time-resolved fashion. This scheme differs from the first one in that the pump and probe beams are focused on different spots. In this way, we can measure the time scale of the transport process from the input to the output, as well as the spectroscopic information of the system. To achieve a high speed from the Auston switches, lattice-matched low-temperature-grown (L TG) GaAs materials which have subpicosecond recombination time will be used.

3.4 High-Tc Superconducting Josephson Devices

Sponsor
Advanced Research Projects Agency
Contract MDA972-90-C-0021

Project Staff
Rajesh K. Gupta, Professor Qing Hu, in collaboration with Twente University

Many superconducting analog devices have been demonstrated to have higher sensitivities, higher speed and frequency limit, and lower power dissipation than competing semiconductor devices. Among them, the most successful ones are superconducting quantum interference devices (SQUIDs) and millimeter-wave detectors.

The discovery of superconductors with their superconducting transition temperature higher than liquid nitrogen temperature (high Tc superconductors) has opened up new and exciting possibilities in electronic device technology. High-temperature versions of the superconducting devices mentioned above will find a much wider-range application wherever refrigeration is a problem. The key element of all superconducting analog and digital devices is the Josephson junction. This junction is formed by two superconductors weakly coupled together electrically. The supercurrent flowing...
through the Josephson junction oscillates at the Josephson frequency $f = 2 \, \text{eV}/h \,(500 \, \text{GHz}/\text{mV})$, when a voltage is applied across the junction. Most of the useful applications of Josephson devices, such as high-frequency and high-speed signal processors and high-sensitivity SQUID magnetometers whose high precision results from averaging over many cycles of the Josephson oscillations, result from this high-frequency oscillation.

In collaboration with the Dutch group headed by Professor Horst Rogalla, which is the leading group in fabricating high-quality high-$T_c$ Josephson junctions, we have studied the response of a YBCO/PBCO/YBCO ramp-type junction to coherent radiation at 176 GHz and 270 GHz. The $I-V$ characteristic of the junction closely resembles the prediction of the RSJ model (resistively-shunted-junction). The critical current and the normal resistance product $I_R$ of the junction is 0.25 mV at 5 K. The millimeter-wave radiation is coupled to the junction via a quasi-optical structure that focuses the radiation onto the junction through a yttrium-stabilized ZrO$_2$ substrate. At 176 GHz, we have observed as many as six Shapiro steps at the maximum power level of our Gunn Oscillator-pumped frequency doubler. This implies a Josephson oscillation up to $1 \, \text{THz}$. Shapiro steps are still visible up to 65 K.

The amplitudes of the zeroth, first, and second Shapiro steps, as functions of the square root of the radiation power, agree remarkably well with a Bessel function fit. This Bessel-like behavior is an indication that the junction is voltage-biased at the radiation frequency. This is in great contrast with other experiments carried at lower frequencies, in which a current-biased behavior is usually observed. For a voltage-biased RSJ-like junction, the detector response can be predicted analytically. The current responsivity of our device, defined as the induced dc current per unit RF power, is calculated to be $2 \times 10^3 \, \text{A/W}$, which is comparable to the quantum efficiency $e/h = 1.4 \times 10^3 \, \text{A/W}$ at this frequency.

Taking advantage of the ideal RSJ-like behavior of our high-$T_c$ Josephson junctions, we have studied the noise effect on their millimeter-wave response by using Ambegaokar-Halperin-Lee (AHL) theory which is well established for low-$T_c$ junctions. Such a study is crucial in developing high-sensitivity devices such as millimeter-wave receivers and SQUID amplifiers. According to AHL theory, noise effects on an RSJ junction behavior are exhibited by the rounding of the sharp corner transition from the constant voltage states near the critical current and the Shapiro steps to the "running states" that exhibit finite differential resistance. This rounding results from a dephasing of the Josephson pair current due to thermal activation. The susceptibility to noise is greatest in these transition regions. AHL theory predicts that the noise-rounded $I-V$ curve near a Shapiro step is given by:

$$
\delta V = \frac{\Delta I_n R_n}{2y} \left( 1 - e^{-2ny\delta I_n/\Delta I_n} \right) \frac{1}{T_1} \left( 1 + \beta_c \frac{T_2}{T_1} \right), \quad (1)
$$

where,

$$
T_1 = \int_0^{\pi} \left[ Y(y) e^{-2y\delta I_n/\Delta I_n} \right] \, d\theta, \quad (2)
$$

and

$$
T_2 = \int_0^{\pi} \sin \theta Y \left( Y \sin \theta e^{-2y\delta I_n/\Delta I_n} \right) \, d\theta. \quad (3)
$$

Here, $\delta V$ is the actual voltage measured from the center of a given step number $n$, $\Delta I_n$ is the total amplitude of the noiseless current step, $\delta I_n$ is the current measured from the center of the $n$th step for $\delta I_n = \Delta I_n/2$, and $R_n$ is the voltage-independent normal state resistance as found in the RSJ. The Stewart-McCumber parameter $\beta_c = \frac{h}{\Delta I_c} \frac{C}{\hbar}$, and $y$ is equal to $\hbar \Delta I_c/2e k_B T_N$. Together they give a complete description of the noise in the system. Figures 7 and 8 show the comparison between the measured and calculated $I-V$ curves near the zeroth, first, and second Shapiro steps. The noise response of the high-$T_c$ Josephson junction can be analyzed extremely well with AHL theory.

In figure 9, we plot the noise temperature $T_N$, which is used as a parameter in fitting the AHL theoretical calculation to our measured $I-V$ curves, as a function of the physical bath temperature $T_p$ that is monitored independently by a Si thermometer. Over a broad temperature range from 10 K to 70 K, $T_N$ is close to the physical bath temperature to within 10 percent, indicating that all high-$T_c$ superconducting active devices based on Josephson effect could reach thermal-noise limited sensitivity.

---

3 Twente University, Enschede, The Netherlands.
Figure 7. Experimental I-V curves (dots) at selected physical temperatures $T_p$ are taken near the critical current region and fitted with I-V curves calculated from Eqs. (1)-(3) (solid lines). The simulated noise temperature $T_n$ is in italics and the physical temperature is in normal text. The current and voltage scales are normalized, and the scaling parameters used are found in the inset.
Figure 8. Similar to figure 7, the first and second Shapiro steps under 176 GHz coherent radiation are plotted against the theoretical noise-rounded I-V curves.

Figure 9. Estimated noise temperature at the zeroth, first, and second steps are plotted versus physical temperature. Error bars result from possible errors in the measurement apparatus and do not represent inconsistencies of fitting data. The dashed line of slope 1 gives $T_n = T_p$. 
Chapter 3. High-Frequency and High-Speed Electronic Devices

3.5 Millimeter-wave and Infrared Superconducting Focal-plane Receiver Arrays

Sponsors
MIT Lincoln Laboratory
National Aeronautics and Space Administration
Grant NAG2-693

Project Staff
Brian R. Jacobson, Professor Qing Hu, in collaboration with MIT Lincoln Laboratory

Millimeter-wave and far-infrared frequencies remain one of the most underdeveloped frequency ranges, even though the potential applications in remote sensing and communication are obviously great. This is because the millimeter wave and far-infrared frequency range falls between two other frequency ranges in which conventional semiconductor devices are usually operated. One is the microwave frequency range, and the other is the near-infrared and optical frequency range. Semiconductor devices which utilize the classical diffusive transport of electrons, such as diodes and transistors, have a high frequency limit. This limit is set by the time it takes electrons to travel a certain distance. Currently, electron mobility and the smallest feature size which can be fabricated by lithography limit the frequency range to below 100 GHz. This limit probably cannot be pushed much higher. Semiconductor devices based on quantum mechanical interband transitions, however, are limited to frequencies higher than those corresponding to the semiconductor energy gap, which is higher than 10 THz for most bulk semiconductors. Therefore, a huge gap exists from 100 GHz to 10 THz in which very few devices are available.

The gap energies of conventional superconductors, such as Nb, are in the range of 100 GHz to 2 THz. This coincidence makes superconducting devices natural candidates for millimeter and submillimeter wave applications. At millimeter wave frequencies, the superconducting video detectors have been demonstrated to have quantum efficiency $e/h\omega$; that is, a transport of one electron for one incoming photon; while the superconducting coherent receivers have been demonstrated to have their sensitivities limited only by the zero-point fluctuation of vacuum. Such receivers have been used widely in astrophysical studies. More applications are feasible in space-based communication and far-infrared spectroscopy which requires the ultimate sensitivity.

We are currently developing a novel scheme to couple the millimeter-wave and infrared signals to the superconducting devices by using a micromachined horn antenna and a planar antenna supported by a thin (~1 μm) membrane, as shown in figure 10. This scheme combines the advantages of easy fabrication of lithographic thin-film structures and the high antenna efficiencies of horn antennas. Because of the absence of substrate losses in this scheme, it is expected that a THz receiver can be constructed using all-Nb superconductor-insulator-superconductor (SIS) junctions. More important, because of the cumbersome mechanical structures used in conventional waveguide technology, microwave receivers have been single-element devices, as opposed to CCD imaging arrays at optical frequencies. Spatial scan has been achieved mechanically. Using the novel quasioptical scheme mentioned above, focal-plane detector arrays can be fabricated lithographically on a single Si wafer, as shown in figure 11, and far-infrared imaging becomes feasible.

Figure 10. Example of anisotropic etching in a <100> silicon wafer. The opening of the wafer yields a pyramidal horn with a flare angle of 35.5 degrees.

Figure 11. Perspective view of a two-dimensional horn imaging array.
Chapter 3. High-Frequency and High-Speed Electronic Devices

The first step of this project is to demonstrate the feasibility of fabricating high-quality SIS devices on free-standing SiN membranes, and to demonstrate that these devices can survive thermal cycling. Using the microfabrication facilities in Group 86 at MIT Lincoln Laboratory, we have fabricated several wafers of SIS junctions with a critical current density of 4-6 kA/cm². We then etched the underneath Si wafer anisotropically in a KOH solution. A picture of an SIS junction with a dipole antenna on a 1-μm thick SiN membrane is shown in figure 12. Extreme care was taken to protect the SIS junctions from the KOH etchant during the anisotropic etching. Figure 13 shows the I-V curves of the SIS junction measured before and after the Si wafer was etched away. There are no noticeable changes in the I-V characteristics of the device. Clearly, our work has established the feasibility of building millimeter-wave SIS receivers using micromachined horn antenna structures. The referee of our paper, which was recently published in *Applied Physics Letters*, commended this work in his remark, “Very important work—This is the next generation of millimeter and submillimeter astronomical receivers. Publish!”

3.6 Far-infrared (THz) Lasers Using Multiple Quantum Wells

Sponsor
U.S. Army Research Office
Grant DAAL03-92-G-0251

Project Staff
James Ernstmeyer, Jurgen H. Smet, Professor Qing Hu, in collaboration with Professor Clifton G. Fonstad, Jr.

Semiconductor quantum wells are man-made quantum mechanical systems in which the energy levels can be chosen by changing the sizes of the quantum wells. Typically, the frequency corresponding to the intersubband transitions is in the far-infrared or THz range. Naturally, long-wavelength photoelectric devices, such as far-

---

Figure 12. Photograph of an SIS device with a dipole antenna and a micromachined horn structure. View of the device seen through the membrane using back illumination of a microscope.

Figure 13. DC I-V curves for a device on SiN membrane (a) before etching, and (b) after etching the Si wafer in KOH.

infrared lasers, which utilize the intersubband transitions have been proposed and subsequently studied. Significant progress has been made recently toward this goal. Large oscillator strengths of intersubband transitions have been observed in far-infrared absorption spectroscopy experiments. An intersubband spontaneous emission with a power level of $\sim 10^{-7}$ W has been observed.

Although these preliminary results are encouraging, two major challenges to building a quantum-well far-infrared laser still remain. One is to achieve a high degree of population inversion between two subband levels, and the other is the confinement of photons within the active region. Optical pumping has low efficiency and low modulation speed; thus it would be difficult to achieve sufficient gain to maintain lasing oscillation and would be unsuitable for communication applications. As far as optical confinement is concerned, the usual dielectric waveguide confinement method, commonly used in optical and near-infrared laser systems, is not applicable at far-infrared frequencies because the confinement (on the order of a wavelength) is too large compared to the dimensions of the active region.

In this project, we will use a novel multiple quantum-well (MQW) device to circumvent these two problems. In this device, whose energy band diagram is shown in figure 14 with and without a dc bias voltage, electrons are injected selectively into an upper subband level of a wide lasing quantum well through a narrow filter quantum well. After relaxing into a lower subband in the lasing well, the electrons are then removed selectively to a collector through another narrow filter quantum well. This filter well prevents the electrons in the upper subband level from tunneling out to the collector. Using this method of selective injection and removal, a high degree of population inversion can be achieved, provided the tunneling rate of the electrons from the ground state to the collector is greater than the relaxation rate of the electrons from the upper subband to the ground state. Our calculation indicates that this should be achieved easily.

In the proposed MQW laser device, a parallel-plate waveguide that confines the photons can be formed between a heavily doped semiconductor injector and collector. At a doping level of $2 \times 10^{19}$/cm$^2$, we can raise the plasma frequency in the injector and collector to about 17 THz. It has been reported that the electron mobility, $\mu$, for such a high doping level is $5 \times 10^8$ cm$^2$/Vs. This plasma is mostly reflective with a penetration depth as short as 0.8 $\mu$m for a lasing frequency in the range of 1-8 THz. This confinement is sufficiently tight so that the lasing threshold current density is on the order of several hundred A/cm$^2$, which should be achievable.

Figure 14. Band profiles (not drawn in proportion) of the proposed far-infrared laser device under (a) zero-bias, and (b) lasing condition, in which $V_{\text{bias}} = (E_g - E_1)/e = 20$ mV. The device consists of one narrow quantum-well energy filter on both sides of the lasing quantum well. The width of the narrow filter wells (made of In$_{0.5}$Ga$_{0.5}$As) is approximately 230 Å, so the energy of its first subband lies in between the first two subbands in the wide lasing well whose width is approximately 360 Å, which will have a radiation frequency of 5 THz due to the intersubband transitions.

The first step of this project is to fabricate multiple quantum-well structures with different well widths and barrier thicknesses, and to perform tunneling spectroscopy on them to determine the relative positions of the subbands in different quantum wells. Because of the small energy differences which are associated with the relatively wide quantum wells, the tunneling spectroscopy must be performed at liquid helium temperatures. Our initial measurements on the tunneling I-V characteristics revealed many features which could be attributed to elastic tunneling and LO phonon-assisted inelastic tunneling.

To interpret the fine features of the I-V characteristics with greater certainty, we have employed a powerful method, magnetotunneling spectroscopy, to study the relative positions of the subbands under concern. In addition to spectroscopy, this study has a practical purpose for laser applications. The nonradiative intersubband relaxation rate due to phonon emissions is much greater than that of the radiative relaxation rate. This is mainly because the transverse motion of electrons in the lasing quantum well is unrestricted, so there is a
large volume in the phase space into which the electrons are scattered. Similar to the concept employed in a tokamak to confine plasmas, a strong magnetic field will restrict the electron motions in the transverse directions. This magnetic field-induced transverse confinement, along with the longitudinal confinement provided by the quantum well, will effectively reduce the dimensionality of the electron system to that of a zero-dimensional system, just like that of quantum dot systems. Therefore, the nonradiative relaxation rate can be substantially reduced. Because the magnetic field does not affect the longitudinal motion of electrons, electrical pumping is still possible which is difficult to implement for quantum dot systems which are formed geometrically.

We have performed magnetotunneling spectroscopy on a double quantum-well structure which consists of 115-Å and 200-Å In$_{0.52}$Ga$_{0.48}$As quantum wells, both are confined by 30 Å of In$_{0.52}$Al$_{0.48}$As respectively on the top and bottom and separated by a 60-Å In$_{0.52}$Al$_{0.48}$As middle barrier. When a longitudinal magnetic field is applied, the tunneling current is modulated by the field, as shown in Figure 15a. The magneto-oscillation is periodic with the inverse of the magnetic field 1/B. Thus, a Fourier transform of this magneto-oscillation in 1/B gives a peak which centers at a characteristic field $B_0$, as shown in figure 15b.

The physical origin of this magneto-oscillation is scattering-mediated inter-Landau level tunneling. Suppose the energy in the first well is given by $E_1 + (n_1 + 1/2)\hbar eB/m^*$, and the energy in the second well is $E_2 + (n_2 + 1/2)\hbar eB/m^*$, where $E_1$ and $E_2$ are the energies of the subbands in the first and second well under concern. In the absence of LO phonon emissions, energy conservation requires the two energies to be equal. Thus, $(n_2 - n_1)\hbar eB/m^* = E_1 - E_2$, or $1/B = (n_2 - n_1)\hbar e/m^*(E_1 - E_2)$. This implies that whenever $1/B$ is an integer times the inverse of a characteristic field $1/B_0 = \hbar e/m^*(E_1 - E_2)$, the energy conservation is satisfied and the tunneling current will be enhanced, which gives the oscillatory behavior of the tunneling current as a function of $1/B$, as shown in Figure 15a.

Since the energy difference between the two subbands is related to $B_0$ through $E_1 - E_2 = \hbar eB_0/m^*$, this magnetotunneling spectroscopy provides us with a very accurate method to measure the relative positions of the subbands under concern. Figures 16 and 17 show the subband energy levels as functions of bias voltage in both forward and reversed polarities. These measured energy levels are within 5 percent from the designed values, which is remarkable considering the relatively small energy scales (~10 meV) for far-infrared applications. This accurate determination of subband energy levels is crucial in successfully developing far-infrared sources with specific frequencies.
Figure 16. The I-V characteristic (a) and the subband energies (b) as a function of applied forward bias at 4.2 K. The quasi-Fermi level in the emitter well is used as the energy reference. Features A, B, C, and D are unambiguously identified as A: $E_{m,1} \rightarrow E_{m,2}$; B: onset of LO-phonon emission process $E_{m,1} \rightarrow E_{m,2} + E_{LO}$; C: all electrons in subband $E_{m,1}$ can participate in LO-phonon emission process $E_{m,1} \rightarrow E_{m,2} + E_{LO}$; and D: $E_{m,1} \rightarrow E_{m,2}$. The measured (designed) intersubband energies are $E_{m,2} - E_{m,1} = 47 (48)$ meV, and $E_{m,2} - E_{m,1} = 113 (109)$ meV.

Figure 17. The I-V characteristic (a) and the subband energies (b) as a function of the applied reverse bias at 4.2 K. The quasi-Fermi level in the emitter well is used as the energy reference. Features A, B, C, D, and E are identified as A: $E_{m,1} \rightarrow E_{m,1}$; B: onset of LO-phonon emission process $E_{m,1} \rightarrow E_{m,1} + E_{LO}$; C: all electrons in subband $E_{m,1}$ can participate in LO-phonon emission process $E_{m,1} \rightarrow E_{m,1} + E_{LO}$; D: $E_{m,2} \rightarrow E_{m,2}$; and E: $E_{m,1} \rightarrow E_{m,2}$. The measured (designed) intersubband energies are $E_{m,2} - E_{m,1} = 52 (48)$ meV, and $E_{m,3} - E_{m,2} = 76 (70)$ meV.
3.7 Publications


Smet, J.H., C.G. Fonstad, and Q. Hu. "Magnetotunneling Spectroscopy in Wide In0.53Ga0.47As/In0.53Al0.47As Double Quantum Wells." \textit{Appl. Phys. Lett.} 63: 2225 (1993).


3.7.1 Meeting Papers


Theses


Section 4  Surfaces and Interfaces

Chapter 1  Statistical Mechanics of Constrained Electronic Systems and Semiconductor Structures

Chapter 2  Synchrotron X-Ray Studies of Surface Disordering

Chapter 3  Chemical Reaction Dynamics at Surfaces

Chapter 4  Semiconductor Surface Studies

Chapter 5  Step Structures and Epitaxy on Semiconductor Surfaces
1.1 Introduction

Sponsor
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Our objectives are to produce microscopic theories and to effect the statistical mechanics for electronic systems that are constrained by finite sizes, quenched disorder, and/or surfaces and for semiconductor systems that exhibit a multiplicity of structures on surfaces or in bulk. In these systems, the microscopic degrees of freedom, electronic or structural, are strongly coupled and therefore their correlated behaviors determine the mesoscopic and macroscopic properties. Examples are the enhanced conductivity of an electronic system, the different step arrangements of a vicinal semiconductor surface, or the different surface or bulk structures of a semiconductor alloy. All of these are among the results obtained in our research. The correlated fluctuations of the degrees of freedom are dominated, at low temperatures by quantum effects, at higher temperatures by thermal effects, and in mesoscopic systems by quenched impurity and boundary effects. Furthermore, competing interactions between these degrees of freedom often cause frustration, namely an inability to satisfy all of the interactions, thereby injecting entropy into the system and causing a highly complex and degenerate free energy landscape.

Renormalization-group theory, introduced by K.G. Wilson for critical phenomena, is a microscopic theory that can address these effects. In our research we have developed this theory for electronic conduction and semiconductor systems. Renormalization-group theory is, when systematically pursued, an almost guaranteed procedure for solving a complex physical problem by repeatedly thinning out the degrees of freedom. Because the calculation remains a physical description at every length scale, it is ideally suited for the inclusion of the effects mentioned above. Another used method that is well suited for finite-size boundary constraints, quenched impurities, and frustration, is our recently introduced hard-spin mean-field theory.

1.2 Renormalization-Group Approach to Highly Correlated Electronic Systems

Project Staff
Alexis Falicov, Professor A. Nihat Berker

The metal-insulator transition, metallic magnetism, heavy-fermion behavior, and high-\(T_c\) superconductivity are all finite-temperature effects that result from the strong correlation of electrons in narrow energy bands. It is therefore of significant interest to study finite-temperature phenomena in strongly correlated electronic systems. Accordingly, we have performed the finite-temperature statistical mechanics of the TJ model of electronic conduction in \(d\) dimensions, obtaining the phase diagrams, electron densities, kinetic energies, and correlation functions using renormalization-group theory. While zero-temperature properties of the TJ model have been studied by a variety of methods, the finite-temperature behavior of the model had been essentially unexplored, especially in \(d=3\), where we now obtain a rich structure.

We find that the renormalization-group method is very well suited for the statistical mechanics of an electronic conduction system. The work described
Chapter 1. Statistical Mechanics of Electronic Systems

here constitutes an important step for the statistical mechanics of mesoscopic systems and quantum dots, as described in section 4.3 below.

Our renormalization-group calculation automatically yields the global finite-temperature phase diagram and statistical mechanics of a generalized $tUJ$ model, defined, on a lattice with one spherically symmetric orbital at each site $i$, by the following Hamiltonian:

$$-\beta H = J \sum_{\langle ij \rangle} \langle \sigma \rangle \left( c_{i\sigma}^+ c_{j\sigma} + c_{j\sigma}^+ c_{i\sigma} \right)$$

$$- J \sum_{\langle ij \rangle} \bar{S}_i \cdot \bar{S}_j + V \sum_{\langle ij \rangle} n_i n_j + \mu \sum_{\langle ij \rangle} n_i,$$

where, at each site $i$, $c_{i\sigma}^+$ and $c_{i\sigma}$ are the electron creation and annihilation operators with $z$-component of spin $\sigma = \{ \uparrow, \downarrow \}$, $n_{i\sigma} = c_{i\sigma}^+ c_{i\sigma}$, and $\bar{S}_i$ are the electron density and spin operators, and $n_i = n_\uparrow + n_\downarrow$. In this Hamiltonian, the consecutive terms respectively represent the electron kinetic energy, nearest-neighbor antiferromagnetic interaction (for $J > 0$), nearest-neighbor Coulomb interaction, and chemical potential. The traditional $tUJ$ Hamiltonian obtains for $V/J = 0.25$.

We find no finite-temperature phase transition in $d=1$ and a finite-temperature critical point terminating a first-order boundary in $d=2$ (figure 1a). The first-order boundary in $d=2$ represents phase separation between electron-rich and hole-rich phases. As the hopping strength $t$ is increased, this phase separation vanishes around $t/J = 0.25$, as seen in figure 1b. This result addresses a currently controversial issue in the theory of strongly coupled electronic systems.

In $d=3$, we find a remarkably complex multicritical phase diagram, with a novel phase and multiple reentrances at different temperature scales, as seen in figures 2 and 3. This phase diagram exhibits antiferromagnetic (a), electron-rich disordered (D) and hole-rich disordered (d) phases separated by first- and second-order phase boundary lines. These boundary lines are punctuated by critical points (C), critical endpoints (E), tricritical points (T), and special multicritical points (M).

The antiferromagnetic phase of the filled system is unstable to a small amount of doping by holes (by about 5 percent in figure 3). One novel aspect is the appearance, close to a narrow phase separation of hole-rich and electron-rich phases, of a new phase (which we have called "r"), occurring at low temperatures in a narrow interval of doping. This phase, to our knowledge never seen before in finite-temperature phase transition theories, is the only volume of the extended phase diagram in which, after repeated length rescalings of the renormalization-group, the electron hopping strength $t$ does not renormalize to zero. In fact, all the interaction constants $(t,J,VJ,\mu)$ renormalize to infinite strengths, while their ratios eventually remain constant, a typical behavior for the renormalization-group sink of a low-temperature phase. A

![Figure 1](image)

Figure 1. (a) Typical cross-section of the calculated global finite-temperature phase diagram for the two-dimensional $tUJ$ model, with $t/J=0.2$ and $V/J=0.25$. Between the electron-rich (D) and hole-rich (d) disordered phases, a phase-separation boundary terminates at a critical point (C). (b) Calculated critical temperatures $1/T_c$ as a function of relative hopping strength $t/J$, in two dimensions. It is thus seen that finite-temperature phase separation occurs only for low values of $t/J$. 

162 RLE Progress Report Number 136
distinctive feature is that at this sink, which as usual epitomizes the entire thermodynamic phase that it attracts, the electron density \( <n> \) has the non-unit, non-zero value of 2/3. This feature makes strong-coupling conduction possible by having the system non-full and non-empty of electrons, and has also not been seen previously. The occurrence and characteristics of our newly discovered phase are quite suggestive.

Figure 2. Typical cross-section of the calculated global finite-temperature phase diagram for the three-dimensional J model, in the temperature and electron chemical potential variables. First- and second-order boundaries are shown respectively with dashed and full curves. Figures 2b and 2c are respectively the blow-ups of the small rectangles in figures 2a and 2b. The corresponding temperature versus electron density phase diagram is in figure 3.

Figure 3. Typical cross-section of the calculated global finite-temperature phase diagram for the three-dimensional J model, in the temperature and electron density variables. The phase separation boundaries of the first-order phase transitions are drawn with dashed curves; the unmarked regions inside these boundaries are the coexistence regions of the phases marked at each side of the regions. The second-order phase boundaries are drawn with full curves. The temperature versus electron chemical potential phase diagram corresponding to this figure is in figure 2a.
Another feature is the appearance of several islands of the antiferromagnetic phase. The islands are bounded by first- and second-order phase transitions adorned by the various special points already mentioned above. Thus, a multiply reentrant phase diagram topology obtains. The antiferromagnetic phase also occurs as a narrow sliver within the disordered phase reaching zero temperature between the antiferromagnetic and "T" phases. The appearance of the antiferromagnetic islands at dopings in the neighborhood of the "T" phase indicates that, when the electron hopping strength t strengthens under rescaling, antiferromagnetically long-range correlated states acquire substantial off-diagonal elements, which lowers the free energy of the antiferromagnetic phase. On the less dense side of the boundary of the "T" phase, a lamellar sequence of antiferromagnetic slivers and disordered inlets occurs, at several temperature scales, as seen in the blow-ups of figures 2b and 2c.

We have calculated electron densities, kinetic energies

$$< T_{ij} > =$$

$$< c_i^+ c_j + c_j^+ c_i + c_i^+ c_j + c_j^+ c_i >$$

and nearest-neighbor density-density and spin-spin correlation functions by summing along renormalization-group trajectories and matching to the densities of phase sinks. The kinetic energies reflect the conductivity of the system as seen from the sum rule

$$\text{Tr} \int_0^\infty dw \text{Re}[\sigma_{\omega}(w)] = -\frac{\pi e^2 \delta d}{2 \hbar a} < T_{ij} >$$

where $\sigma_{\omega}(w)$ is the frequency-dependent conductivity, $e$ is the electronic charge, and $a$ is the lattice spacing. Thus, the conductivity, as expected, goes to zero in the dilute and dense limits (figure 4).

1.4 Collective Phenomena in Systems with Quenched Impurities and Frustration

Project Staff
Alexis Falicov, Daniel P. Aalberts, Professor A. Nihat Berker

Our studies of systems with quenched randomness and frustration have yielded (1) new phenomena that are very generally applicable and (2) highly developed theoretical tools to deal with such effects in electronic and semiconductor systems. These theoretical tools consist of the renormalization-group theory, which we have perfected for the study of systems with quenched randomness, and the hard-spin mean-field theory, which for the first time conserves frustration and which in fact was introduced in these studies. Frustration, which signifies the lack of a unique local minimum-energy
state due to competing interactions, appears to
underly phenomena in systems ranging from spin-
glasses (where, as we have discovered, a chaotic
rescaling behavior results) to liquid crystals and to
high T_c superconductivity.

One of our general results is that, under quenched
random interactions, all first-order phase transitions
that involve a symmetry breaking are converted into
second-order phase transitions. This amounts to
the generation of infinite correlation lengths by the
effect of impurities! This prediction, which, for
example, raises the possibility of replacing the
uncontrolled hysteresis loops of probing devices by
infinite-response functions, was originally indicated
by our general domain wall arguments. Most
recently, our detailed renormalization group
calculations have confirmed the prediction, for example
with the phase diagram of a spin-1 magnetic model
(the Blume-Emery-Griffiths model), used in the
study of electronic systems and of semiconductor
alloys. The pure (non-random) system has first-
order phase transitions at low temperatures and low
densities, and second-order phase transitions at
high temperatures and high densities. Our calcula-
tions show that, with the introduction of quenched
randomness into the interactions, first-order transi-
tions are converted to second-order transitions,
which then reach zero temperature and low densi-
ties.

1.5 Surface and Bulk Structures of
Semiconductor Alloys

Project Staff

William C. Hoston, Asad A. Naqvi, Alkan
Kabakçıoğlu, Professor A. Nihat Berker, in collabor-
ation with Professor John D. Joannopoulos

Ternary and quaternary semiconductor alloys can
exist in zincblende, chalcopyrite, or possibly
stannite structures, involving two interpenetrating
fcc lattices on which up to four atomic species exist.
One atomic species occupies one of the fcc lattices
while three other atomic species may (chalcopyrite,
stannite) or may not (zincblende) order on the other
fcc lattice. The structural phase diagram of these
alloys as a function of atomic interactions, composi-
tion, and temperature is of interest. In recent work,
K.E. Newman and collaborators have adapted the
spin-1 Ising (Blume-Emery-Griffiths) model to the
study of zincblende to chalcopyrite or stannite tran-
sitions. In this model, the three spin values are
associated with a different species of atom, A,
B, or C, which exist on one of the fcc lattices. The
other fcc lattice is occupied by atomic species D.
The systems under consideration have the com-
position \((AB)_1(C_2)_3D_2\). Precise values of the model
interaction strengths applicable to the semicon-
ductor alloy systems will be provided, in our work,
from the total energy calculations of Professor J.D.
Joannopoulos. The regimes of the model that are
applicable to the alloy systems include negative
bilinear and negative biquadratic interactions, and

Figure 4. Electron kinetic energies in the tJ model, calculated by renormalization-group theory, at temperature 1/J=2
across the phase diagram of figure 3. These kinetic energies are related to the conductivities by the sum rule given
above.

165
therefore frustration, both features bringing interesting complications. Newman et al. obtained unusual phase diagrams using a cluster-variational theory, which does not include correlated fluctuations and frustration at the level of renormalization-group theory and hard-spin mean-field theory.

In our previous works, we have studied the Blume-Emery-Griffiths model extensively. Most recently we have studied the negative biquadratic interaction regimes, finding six new phase diagrams, with a novel multicritical topology and two new ordered phases. Thus, we have determined that the phase diagram of this simple spin system includes nine distinct topologies and three ordered phases. These results were obtained by a global mean-field theory with four independent order parameters. Our renormalization-group theory confirmed the occurrence of critical points inside an ordered phase, a question that had been raised by measurements on metamagnetic alloys.

However, since the fcc lattice is frustrated, analysis beyond standard mean-field theory is required. Our new method, "hard-spin mean-field theory", incorporates the hard-spin condition of local degrees of freedom and thereby conserves frustration. We have tested this method on frustrated triangular and stacked triangular lattices with the spin-1/2 Ising model, obtaining excellent results. We intend to apply the method to the Blume-Emery-Griffiths model on the fcc lattice, which is more frustrated. We also intend to use renormalization-group theory, which a preliminary study we have done indicates is feasible on this complicated physical system.

Furthermore, these semiconductor alloys are expected to have a variety of surface reconstruction phases on different crystal facets and a variety of step structures on vicinal surfaces. Our calculations will be extended to these effects, also using the meandering step-Hamiltonian method introduced in our previous study of vicinal silicon (100) surfaces. These calculations for bulk and surfaces of semiconductors integrate the electronic energy determinations of Professor J.D. Joannopoulos with statistical mechanics, and thereby are uniquely first-principles calculations for physical systems.

1.6 Publications


1.6.1 Meeting Papers


Professor Robert J. Birgeneau, Dean of the School of Science. (Photo courtesy of the Photography Division of the Brookhaven National Laboratory)
Chapter 2. Synchrotron X-ray Studies of Surface Disordering

Academic and Research Staff
Professor Robert J. Birgeneau, Dr. Barry Wells
Graduate Students
Monte J. Ramstad, Michael J. Young
Technical and Support Staff
Debra L. Harring

2.1 Introduction

Sponsor
Joint Services Electronics Program
Contract DAAL03-92-C-0001

In this research program, we use modern x-ray scattering techniques to study structures and phase transitions in thin films and on surfaces. We have two principal experimental facilities, one at MIT and the other at the National Synchrotron Light Source at Brookhaven National Laboratory. At MIT, we have four high-resolution computer-controlled x-ray spectrometers. The angular resolution can be made as fine as 1.8 seconds of arc, which enables us to probe the development of order from distances of the order of the x-ray wavelength 1 Å, up to 30,000 Å. The sample temperature can be varied between 2K and 500K with a relative accuracy of 2 x 10⁻³ K. At the National Synchrotron Light Source, in collaboration with IBM, we have three fully instrumented beam lines. Two of these beam lines allow us to make studies with photons varying in energy between 3 and 12 keV; the third has a fixed energy of 17 keV. These facilities make possible high-resolution scattering experiments with a flux more than three orders of magnitude larger than that of a rotating anode x-ray generator. In collaboration with Professor Simon G.J. Mochrie, we have now designed, built, and commissioned a second-generation extremely versatile x-ray surface facility which allows a wide range of surface scattering experiments.

Our basic scientific objective is to understand the morphologies and microscopic structures of simple semiconductor and metal surfaces at high temperatures. Possible phase changes include surface roughening, surface reconstruction, melting, amorphization, and dilution.

Semiconductor surfaces exhibit rich structural and morphological behavior in both equilibrium and non-equilibrium steady-state conditions. Equilibrium structures and phase transitions of high-symmetry facets are often closely related to important models in two-dimensional statistical mechanics. The physics of stepped (vicinal) surfaces, even in equilibrium, is subtle and poorly understood. For example, stepped surfaces should always be rough because of the freedom of the steps to wander. Step structures and spacing should be profoundly affected by any nearby high-symmetry surface reconstruction. Their association bears on issues of faceting and equilibrium crystal shapes. The steps themselves may exhibit phase transitions that involve dislocations and/or bunching in the step lattice as well as step-height changes; changes in the morphology, especially those involving vacancy aggregates, should be strongly affected by nearby steps.

Semiconductor surface physics involves fundamental issues in statistical mechanics that are important to a broad spectrum of other problems. Furthermore, elucidating the factors that control the morphology and perfection of vicinal semiconductor surfaces can be expected to lead to improved electronic device processing and ultimately improve semiconductor device performance and lifetime.

2.2 Vicinal Semiconductor Surfaces

Previously, our research on vicinal semiconductor surfaces was focused on the faceting behavior of vicinal Si(111) around the 1 x 1 to 7 x 7 reconstruction transition at 1100K. Above 1100K, the steps are on average uniformly spaced, but the surface is rough due to fluctuations in the positions of the steps. In the course of these studies, we discovered some novel non-equilibrium faceting phenomena which occur at higher temperatures and which appear to be driven by applied electric fields. We have been studying these phenomena using both x-ray and optical diffraction techniques. Our most recent results are as follows.
We have observed an intricate faceting behavior of vicinal Si(111) which depends on the direction of the current flowing through the sample. The behavior above 1375K (and perhaps lower) is qualitatively similar for miscuts of 3 and 5 degrees toward $<11-2>$ and 3 degrees toward $<-1-12>$. The faceting can be longitudinal (step bunching) or transverse (step wandering), apparently due to instabilities in the step flow during sublimation. In support of this hypothesis, our experiments indicate that these transformations are sharp in temperature and DC bias. For positive current (current flowing down the steps), there is a transverse instability between 1175K and 1575K. For negative current, the steps are stabilized transversely. Longitudinal instabilities occur for alternating current directions as a function of temperature. The instability exists for positive current at 1175K, for negative current between 1275K and 1525K, and for positive current above 1525K.

Most of our studies have been of the time and history dependence of these transformations. As expected, the rates increase with temperature and magnitude of DC bias. At the highest rates, the surface morphology will continue to evolve for at least 12 hours. However, the rate of change decreases exponentially with time. Two lengths are associated with faceting: the facet spacing and the correlation length of the facets. For transverse facets, the facet spacing initially increases rapidly then becomes almost constant in time. It also appears to decrease slightly with temperature. The correlation of the facets falls off exponentially with the distance between facets. The associated correlation length decreases with time.

The behavior of the steps also depends on the history of the sample. Although it is possible to facet and unfacet a surface, it appears that a surface which has never been faceted behaves differently from one which has been faceted. For negative current above 1525K, a surface will unfacet both longitudinally and transversely forming an optically flat surface. On the other hand, a never-faceted surface will become optically rough. Also, a surface which has been heated to 1475K for an extended period using AC current develops disordered facets which are difficult or impossible to remove.

We are continuing our studies of these most unexpected phenomena.

### 2.3 Chiral Melting of the Si(113) (3x1) Reconstruction

According to the concept of universality, the critical behavior near a continuous phase transformation depends only on the most general aspects of the system in question, such as the symmetry of the phases involved and the dimensionality of space. Thus, the phase transformations of many disparate materials fall into universality classes, within each of which the critical exponents and certain universal constants are independent of microscopic details and delineate the universality class. It is generally believed that two-dimensional (2D) phase transformations may be understood in terms of a limited number of well-known universality classes. Therefore, the appearance of a new 2D universality class is of special interest, particularly when the critical behavior corresponds to a simple and generic statistical mechanical model, and also is unusual in its own right. The experimental characterization of such a universality class is the topic of this research.

In collaboration with Professor Simon G.J. Mochrie, we have carried out a detailed x-ray scattering study of the (3x1)-to-disordered phase transformation of the Si(113) surface. A continuous commensurate-solid to incommensurate-fluid transformation at $T_c = 950 \pm 40$K is observed. At the transformation, the reconstructed layer becomes uniaxially incommensurate along the cubic (110)-direction (x-direction). It remains commensurate along the (332)-direction (y-direction). Critical scattering shows power-law behavior over nearly two decades of reduced temperature ($t = (T - T_c)/T_c$) with exponents $\beta = 0.66 \pm 0.05$ for the incommensurability ($\epsilon$), $\nu_x = 0.65 \pm 0.07$ for the inverse correlation length in the incommensurate direction ($k_x$), $\nu_y = 1.06 \pm 0.07$ for the inverse correlation length in the commensurate direction ($k_y$) and $\gamma = 1.56 \pm 0.13$ for the susceptibility ($\chi$).

Below $T_c$, the variation of the square of the order parameter, proportional to the peak intensity at the commensurate position ($I_0$), varies with an exponent $2\beta = 0.22 \pm 0.04$. It is noteworthy that the correlation lengths in the disordered phase scale anisotropically, that is $\nu_x \neq \nu_y$, and that the collected exponents do not conform to those of any previously known universality class. In addition to the critical exponents of the transformation, two universal constants have been measured. The ratio of the incommensurability and the inverse correlation length along the incommensurate direction in the disordered phase is found to be independent of temperature, that is $\beta = \nu_x$, and to have the value $\nu_0 = 1.6 \pm 0.2$. This is consistent with predictions...
Chapter 2. Studies of Surface Disordering

for a new two-dimensional chiral melting universality class. Also, the combination \( R_s = \chi K_s k_\beta l_\omega V_r \), where \( V_r \) is the two-dimensional resolution volume, is independent of the reduced temperature, consistent with the derived hyperscaling relationship \( \nu_s + \nu_f = \gamma + 2\beta \). According to the hypothesis of two-scale-factor universality, \( R_s \) is a universal constant, which we find takes the value \( R_s = 0.07 \pm 0.03 \).

2.4 Two-Dimensional Melting

In spite of many years of study, the nature of melting in two-dimensions continues to be a controversial issue. Many theories suggest that melting should be strongly first order as in three dimensions. Alternatively, melting in two dimensions may be continuous; the specific mechanism is the unbinding of pairs of point dislocations in a manner analogous to the proposed unbinding of vortices in two-dimensional superfluids.

We have carried out a study of the melting transition of near-monolayer xenon adsorbed onto the basal planes of single crystal graphite at a temperature of 140K. Xenon on graphite is widely regarded as a model system for the investigation of two-dimensional melting in the presence of a weak orientational ordering field. It has been suggested that it may represent an example in which melting occurs continuously through the unbinding of topological defects in a manner analogous to the Kosterlitz-Thouless transition in superfluids. Recent thermodynamic studies by Chan and coworkers have reported that the melting of near-monolayer xenon on graphite is always first order in contrast to the observations of previous x-ray scattering studies as well as compressibility studies by Suter and coworkers.

Our investigation confirms that at a temperature of 140K there is indeed a continuous evolution of the length scale of the positional fluctuations corresponding to a transition from a finite-size limited 2D solid to a well correlated (\(~100\) Å) orientationally ordered 2D liquid. The relevant data are shown in figure 1. This figure shows longitudinal and transverse x-ray scans through the (10) Bragg peak of the monolayer xenon triangular lattice. In this series of scans the temperature is held fixed while the xenon pressure is varied continuously. The actual melting transition occurs at 17.5 \pm 0.2 Torr. It can be seen from the figure that with decreasing pressure the diffraction peaks broaden progressively and the intensity decreases continuously. This is the signature of a continuous rather than discontinuous (first order) melting process, thus confirming that melting at 140K is indeed continuous.

One signature of a continuous transition in a finite-size limited 2D XY model is that the order parameter should simulate a smeared power law with exponent \( \beta = 0.23 \). We show in figure 2 the measured solid intensity as a function of chemical potential near the melting transition. The solid line is rounded power law with \( \beta = 0.23 \). Clearly, this model works very well.

2.5 Publications


Thesis

Chapter 2. Studies of Surface Disordering

Figure 1. Longitudinal scans obtained at 140K close to the melting transition, as analyzed in terms of an aligned liquid model. The transition is at ~ 17.5 Torr. Note that the peak half-width at half-maximum evolves continuously as the finite-sized solids melts into the aligned liquid phase. The dashed lines shown in the panels at 16.58 Torr and 16.27 Torr correspond to the lineshape predicted from a finite-sized analysis in the case that the exponent $\eta_0$ is constrained to its fitted value deep in the solid phase, $\eta_0 = 0.45$. It is clear that the data at lower vapor pressures can no longer be described by the solid lineshape.
Figure 2. Intensity, $S_0$, obtained from our finite-sized solid model versus reduced chemical potential. The solid line represents the results of a fit to the power law form $|\mu/\mu_0 - 1|^{1/\beta}$, with $\beta = 0.23$ and $\mu_0$ smeared with a Gaussian distribution of reduced chemical potentials, $\mu/\mu_0$, with width $\sim 0.005$. 

$140K \beta = 0.23 \pm 0.04$
Chapter 3. Chemical Reaction Dynamics at Surfaces

Academic and Research Staff
Professor Sylvia T. Ceyer, Dr. David P. Pullman, Dr. Athanassios A. Tsekouras, Dr. Arthur L. Utz, Dr. Zhe Zhang

Graduate Students
Sean P. Daley, David B. Gosalvez, Theodore R. Trautman, Julius J. Yang

Technical and Support Staff
Catherine Cooper

3.1 A New Mechanism for Dissociative Chemisorption on Si: Atom Abstraction

Sponsor
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff
Professor Sylvia T. Ceyer, Dr. David P. Pullman, Dr. Athanassios A. Tsekouras, Dr. Zhe Zhang, David B. Gosalvez, Julius J. Yang

The classic view of dissociative chemisorption of a molecule incident on a surface involves cleavage of the bond of an incident molecule by formation of two bonds to the surface. The result is two fragments adsorbed on the surface. The dissociation proceeds in this manner because the formation of two bonds is energetically necessary for bond cleavage. This classic mechanism for dissociative chemisorption is operative in many molecule-surface systems. In contrast, one can envision bond cleavage by formation of a single surface-atom bond if the energy released by this bond formation is greater than the bond energy of the incident molecule. Generically, this mechanism, known as abstraction, is a well-documented reaction mechanism between two gas phase molecules.

With JSEP support, we have recently documented the dissociative chemisorption of a molecule on a surface by an abstraction mechanism. Specifically, we demonstrated that a Si(100)2x1 surface abstracts one F atom from an incident F₂ molecule by detecting the complementary F atom scattered back into the gas phase with a triply differentially pumped, rotatable, quadrupole mass spectrometer in a molecular beam-surface scattering UHV apparatus. In addition, we coupled these experiments with He atom diffraction from the resulting fluorinated surface to demonstrate that the Si surface dangling bonds are responsible for the abstraction and are the sites for F adsorption. Because of the experimental complexities in the detection of reactive radicals such as F atoms, this mechanism has not been found in numerous previous studies of the interaction of fluorine and fluorine containing molecules with Si.

Briefly, the triply differentially pumped F₂ beam is formed by expansion of a mixture of 1 percent F₂ in Kr. The velocities of the incident and scattered beam are determined by cross correlation time-of-flight (TOF) methods. The beam is directed at an n-type Si(100) crystal that can be heated resistively and cooled to 150 K. The crystal is etched prior to insertion into the vacuum chamber where it is repetitively sputtered and annealed (850°C for 30 minutes) until Auger spectroscopy reveals carbon as the only contaminant, at a 1 percent or less level. The atom diffraction measurements show the surface to exhibit the well-documented (2x1) reconstruction associated with the formation of Si surface dimers. In real space, the surface exhibits rows of Si-Si dimers, with one dangling bond per Si atom.

Identification of the abstraction mechanism by detection of the scattered F atom is complicated by the fragmentation of the unreactively scattered F₂ in the electron bombardment ionizer of the mass spectrometer to produce both F₂ (m/e=38) and F⁺ (m/e=19) ions. Therefore, a signal observed at m/e=19 will contain a contribution from scattered F₂ as well as from F atoms, if present. We can distinguish between F⁺ arising from F atoms and that produced by the cracking of F₂ on the basis of the different velocities with which they scatter from the surface. Two features are clearly observed in the time of flight spectrum, a narrow one at short times and a broad feature at longer times. The TOF distribution measured at m/e=38, scaled by the F⁺/F₂ cracking ratio, can be superimposed on the m/e=19 spectrum. The single feature in m/e=38 distribution is observed to occur at the same flight time as the broad one in the m/e=19 distribution and its intensity, scaled to represent the component of the
m/e=19 signal attributed to F₂, matches well the intensity of the broad feature in the m/e=19 spectrum. These two observations lead to the conclusion that the broad feature arises from the cracking of the unreactively scattered F₂ while the narrow, short time feature arises from scattered F atoms. A point-by-point subtraction of the two distributions yields the net scattered F atom TOF distribution. From the fits to the distributions, the flux-weighted average velocities of F and F₂ are determined to be 1140±40 m/s and 430±32 m/s, respectively, for a surface temperature of 250K. The incident F velocity is 390 m/s. Measurements of the TOF distributions of F and F₂ scattered from a surface at a temperature of 1000 K show that the velocity of the scattered F is independent of temperature (1181±64 m/s) while the velocity of the scattered F₂ is higher (600±42 m/s). No ions are observed to desorb.

Additional confirmation of the presence of scattered F atoms comes from the F₂ exposure dependence of the scattered signal at m/e=19 and 38. The two signals have a very different dependence on F₂ exposure. While both signals become constant at high exposure, the m/e=19 signal rises significantly more rapidly at very low exposure than the m/e=38 signal. It is clear that the F⁺ signal at low exposure arises from a source different from the cracking of F₂ in the ionizer. We assign it to F atoms scattered from the surface. The F⁺ atom signal is a maximum at a low but non-zero exposure and then decays to zero with increasing exposure.

To probe the site of F adsorption, we have measured He diffraction spectra from the clean and F-covered Si(100)2×1 surface. They are recorded by directing a supersonic He atom beam (E=11 meV, ΔE=4 meV), incident on the 250 K surface at a 20° angle from the surface normal, and then detecting the scattered He in 0.5° increments. The He beam is modulated at 150 Hz for the purpose of background subtraction. The key result is that the half order diffraction beam which is the signature for the presence of the Si-Si dimers, is clearly present in the spectrum from the fluorinated surface, indicating that the adsorption of F does not break the dimer bond. Moreover, the diffraction intensities from the fluorinated surface remain unchanged upon further F₂ exposure, indicating that the surface is saturated and that the F is adsorbed as an ordered overlayer with a (2x1) unit cell. In addition, we have determined, as discussed below, that the coverage of this (2x1) overlayer is 1 ML, which is equivalent to one F for every Si dangling bond. These observations convincingly demonstrate that the reaction of F₂ with Si(100) is occurring at the dangling bond sites of the surface dimers, and that no Si-Si lattice bonds are broken. This structural determination does not suffer from the ambiguity in LEED generated by efficient electron stimulated desorption of F and is consistent with other results that identify SiF as the dominant species.

These results signal the observation of a new mechanism for dissociative chemisorption and we designate it F atom abstraction. As F₂ is incident on this surface, a Si dangling bond abstracts one of the F atoms while the other F atom is scattered away. The cleavage of the F₂ bond by the formation of a single F-Si bond is thermodynamically feasible because the energy released upon adsorption on a single Si dangling bond, which does not require cleavage of a Si lattice bond, is 5-6 eV compared to 1.5 eV for the F₂ bond energy. Some of the exothermicity of this bond formation is converted to translational energy of the scattered F, as evident from the large velocity of the scattered F compared to the incident F₂ and the independence of the velocity on surface temperature. Abstraction has been observed in a simulation of this system.

While the F atom that is not abstracted can scatter back into the gas phase, it does not necessarily do so. It may be caught on its outgoing trajectory by an adjacent dangling bond and adsorb there. It is also possible for both atoms to be simultaneously abstracted by two adjacent dangling bonds if the F₂ molecular axis is favorably aligned upon its initial collision. We can demonstrate that adsorption of both atoms also occurs by measuring the difference between the total F₂ adsorption probability, P, and the probability for single atom adsorption, P₁. P₁ is determined from 1-P, where P₁ is the probability for unreacted F₂ to be scattered from the surface, is the angle integrated F₂ flux to the incident F₂ and P₁ is the ratio of the angle integrated F flux to the incident F₂ flux. The difference P₁-P₂ is the probability for adsorption of both F atoms, P₂, and is plotted, along with P₁ and P₂ versus coverage. The coverage is determined by integrating P₂×0.5P₁ over F₂ exposure. In the zero coverage limit, P₂ contributes 0.83 and P₂ contributes 0.21 to the total P₁ of 0.95, in reasonable agreement with a simulation and with a previous experimental result for P₁. The adsorption probability is observed to decay to zero when the dangling bonds are saturated at 1 ML fluorine coverage.

The maximum in P₁ at a non-zero coverage is also consistent with the presence of a reaction channel in which both F atoms are adsorbed. P₂ is small in the limit of zero coverage because the dangling bonds adjacent to the abstraction site of the first F atom are likely unoccupied and, therefore, have a high probability of trapping the second F. As the coverage increases, the probability of having unoccupied adjacent sites decreases, thereby leading to
a higher probability that the second F does not adsorb and scatters back to the gas phase. As the coverage increases beyond 0.6 ML, the effect of the diminishing number of sites for abstraction of the initial F dominates, leading to an overall decay of $P_1$. In short, the maximum in $P_1$ arises from the competition between the number of abstraction sites available to the incident $F_2$, which decreases with coverage, and the number of occupied sites (which cannot trap the second F) adjacent to the abstraction site, which increases with coverage. A simple analytical model which incorporates these physical features well and describes these data is presented in the next section. Note that $P_1$ is a lower limit to the abstraction probability because its value is based only on those complementary atoms which successfully scatter back to the gas phase.

In summary, we have identified and characterized a new mechanism for dissociative chemisorption, atom abstraction. This mechanism is a consequence of the high exothermicity of the reaction between $F_2$ and Si and should be observable in other highly exothermic molecule-surface interactions. This mechanism may also be an additional source of radicals that have been previously unaccounted for in models for plasma etching.

### 3.1.1 Publications

#### Journal Article


#### Chapter in Book


### Meeting Papers


### Theses


### 3.2 A Model for Atom Abstraction by Surfaces

#### Sponsor

Joint Services Electronics Program
Contract DAAL03-92-C-0001

#### Project Staff

Professor Sylvia T. Ceyer, Dr. David P. Pullman, Dr. Athanassios A. Tsekouras, Dr. Zhe Zhang, David B. Gosalvez, Julius J. Yang

The probability for atom abstraction, $P_1$, exhibits a very non-langmuir dependence on fluorine coverage. Specifically, our abstraction probability first increases, goes through a maximum and then decreases to zero at a coverage of one monolayer of fluorine while a langmuir dependence of an adsorption probability decreases strictly linearly with coverage. We have developed a simple analytical model based on lattice gas kinetics to understand this peculiar dependence on coverage. Consider a $F_2$ molecule approaching the surface with its molecular axis roughly parallel to the surface. The probability of the first fluorine atom being abstracted depends linearly on the coverage, a $(1 - \theta)$ dependence, where $\theta$ is the fractional coverage. Once the first F atom is abstracted, the second F atom is propelled away by the exothermicity of the reaction. However, because of the horizontal orientation of the incident $F_2$ molecule, the second F atom is propelled along the surface. The probability of it showing up in the gas phase where it contributes to our measurement of $P_1$ is proportional to the occupancy of the adjacent sites which has a $\theta$
dependence. That is, if the fluorine atom is propelled in the direction of a unfilled dangling bond, it will adsorb and will not scatter into the gas phase. But, if it is propelled in the direction of a filled dangling bond, it cannot adsorb and will scatter into the gas phase. Therefore, the \((1 - \theta)\) dependence must be multiplied by a \(\theta\) dependence resulting in a quantity proportional to \(P\) that is zero at zero coverage and that has a maximum at \(\theta = 0.5\). Therefore, it is clear that the maximum in the abstraction probability occurs as a result of the competition between the number of unfilled sites for abstraction of the first \(F\) atom which decreases with coverage and the number of filled sites to scatter the second \(F\) atom into the gas phase which increases with coverage.

However, there is an additional contribution to the abstraction mechanism which shifts the coverage at which the maximum occurs to lower values and which results in a non-zero coverage at zero fluorine coverage. This additional contribution arises from those \(F_2\) molecules that are incident on the silicon surface with their molecular axes perpendicular to the surface. In this case, the probability for abstraction of the first \(F\) atom again has a \((1 - \theta)\) dependence, but because of the vertical orientation of the incident \(F_2\), the second \(F\) atom is propelled away from the surface. The probability of its appearance in the gas phase is independent of the occupancy of the adjacent sites. Therefore, this \((1 - \theta)\) dependence for the vertically oriented molecules must be added to the \(\theta(1 - \theta)\) dependence for the horizontally oriented molecules. The resulting prediction for the abstraction probability \(P\) versus coverage agrees reasonably well with the data. The good agreement with this lattice gas model indicates that the mechanism for dissociative chemisorption indeed proceeds as a two step process involving first an abstraction step followed by the adsorption of the second \(F\) atom if it is propelled across the surface in the direction of an empty dangling bond. If it is propelled in the direction of a filled site, it scatters into the gas phase away from the surface. Whether the second \(F\) atom is propelled across the surface or away from the surface depends on its incident orientation.

We have also carried out a Monte Carlo simulation of the abstraction probability versus coverage. The results of this simulation show that steric hindrance between the incident \(F_2\) and the filled dangling bond sites at intermediate coverages plays a large role in the rapid drop of the adsorption probability of \(F_2\) with coverage. That is, even though the first abstraction step requires only a single empty dangling bond site, it is difficult for an incident \(F_2\) molecule to access this site if it is surrounded by sites on which \(F\) atoms are already adsorbed.

3.2.1 Publication


3.3 Atom Abstraction and its Relationship to Thin Film Growth

Sponsor

Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff

Professor Sylvia T. Ceyer, Dr. David P. Pullman

We have recently demonstrated by Monte Carlo simulation, that our newly discovered mechanism, atom abstraction, provides a suitable physical explanation for non-statistical island growth. Specifically, Brune et al. employed scanning tunnelling microscopy to show that the spatial distribution of \(O\) atoms on \(Al(111)\) after adsorption of \(O_2\) at low exposures was essentially random. This is a surprising result because within a conventional view of dissociative chemisorption, the adsorbing atoms are expected to reside on neighboring sites. The resulting spatial distribution is therefore non-random and is manifested in a pair correlation function that peaks at the first nearest-neighbor shell. They explained these observations by a hot adatom model in which the reaction exothermicity is channelled into adatom translational motion parallel to the surface plane. If the transfer of energy into translation is \(l a r e\), the mean flight path of adatoms on the surface will likewise be large and the pair correlation function will be structureless.

We have shown however, that the structureless pair correlation function at low coverage and the non-statistical island growth at high coverage are also consistent with abstraction of one of the \(O\) atoms from the incident \(O_2\) molecule while the unabtracted atom is either propelled into the gas phase or propelled along the surface. The trajectory of this latter atom may be halted if it collides with an island next to which there is an empty site but if not, it will scatter into the gas phase. As we have pointed out, the abstraction of an \(O\) atom from \(Al\) is thermodynamically feasible as it is in the interaction of \(F_2\) with \(Si(100)\). Our analysis demonstrates the generality of the atom abstraction mechanism and the importance of this mechanism to thin film growth.
3.3.1 Publication


3.4 Etching of Si(100) by Energetic Fluorine

Sponsor

Joint Services Electronics Program
Contract DAAL03-92-C-0001

Project Staff

Professor Sylvia T. Ceyer, Dr. David P. Pullman, Dr. Athanassios A. Tsekouras, Dr. Zhe Zhang, David B. Gosalvez, Julius J. Yang

The reaction of \( \text{F}_2 \) with Si(100) ceases after the dangling bonds are saturated at one monolayer of coverage. This lack of additional reactivity precludes the build up of a sufficient layer of fluorine to produce the volatile etch product \( \text{SiF}_4 \), and therefore the use of \( \text{F}_2 \) as an etchant of Si. However, we have shown that if the kinetic energy of the incident \( \text{F}_2 \) molecule is increased above a threshold value of 4-5 kcal mol\(^{-1}\) (0.2 eV), the dissociation probability of \( \text{F}_2 \) on a fluorinated surface increases linearly with the normal component of the kinetic energy. This additional reactivity at higher translational energies, which results in fluorine coverages greater than one monolayer, is the result of sur-mounting barrier to dissociation on the fluorinated surface. During this funding period we coupled our measurements of the reactivity of \( \text{F}_2 \) as a function of energy with a structural technique, He atom diffraction. The goal was to identify the sites for the reaction of the translationally activated \( \text{F}_2 \) with the fluorinated surface. Our measurements of the He atom diffraction spectra as a function of the incident F kinetic energy indicate that the barrier is associated with dissociative chemisorption on the second layer of Si atoms as well as on Si dimer atoms. These conclusions are based on the identical rate of linear decay for all diffraction features with increasing \( \text{F}_2 \) normal energy above the threshold. There is no preferential cleavage of the Si dimer atoms. These observations may represent the first experimental correlation of the barrier for dissociation with the sites for dissociation.

These results also provide conclusive evidence that Si can be etched with low energies, two orders of magnitude below the energies used in plasmas, using molecular beam techniques, and without the use of plasmas. The low energies afforded by molecular beam techniques preclude the introduction of radiation damage or defects into the Si lattice, which is a typical consequence of plasma etching.

3.4.1 Publication

Chapter 4. Semiconductor Surface Studies

4.1 Introduction

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001

Understanding the properties of the surfaces of solids and the interactions of atoms and molecules with surfaces has been of extreme importance both from a technological and an academic point of view. The advent of ultrahigh vacuum technology has made microscopic studies of well-characterized surface systems possible. The way atoms move to reduce the energy of the surface, the number of layers of atoms involved in this reduction, the electronic and vibrational states that result from this movement, and the final symmetry of the surface layer are all of utmost importance in arriving at a fundamental and microscopic understanding of the nature of clean surfaces, chemisorption processes, and the initial stages of interface formation.

The theoretical problems associated with these systems are quite complex. However, we are currently at the forefront of solving the properties of real surface systems. In particular, we are continuing our efforts to develop new techniques for calculating the total ground-state energy of a surface system from "first principles," so that we can provide accurate theoretical predictions of surface geometries and behavior. Our efforts in this program have concentrated in the areas of surface growth, surface reconstruction geometries, structural phase transitions, and chemisorption.

4.2 Defects on Surfaces

Real surfaces typically contain a variety of imperfections that can be important in determining their electronic and chemical behavior. A particularly interesting system is the vicinal Si(100) surface. The Si(100) surface has been an object of intensive study for the last three decades in part because it is technologically important in semiconductor industry, and in part because its reconstructions are simple enough to be tractable to theoretical investigations.

Although much is understood about the electronic and structural properties of Si(100), there has been relatively little work done to study the nature of intrinsic defects on this surface. Interestingly, scanning-tunneling-microscopy (STM) images of this surface consistently show the presence of a large number of what appear to be missing dimers or dimer vacancies (DV). In fact, a typical STM image of this surface exhibits three striking features. The first is the very large number of missing dimers present on the surface. The second is that the missing dimer defects tend to congregate into small clusters, and the third is that the clusters appear to form distinctive complexes that dominate over others. These results are summarized for a typical image (which had about 22,000 dimer sites in total) in figure 1. The data shown here are from a surface annealed at 700°C and "flashed" up to 1100°C for several minutes. The STM image, however, is taken at room temperature after the sample has cooled down. The STM image therefore represents a quasi-equilibrium "snap-shot" at some freezing-in temperature below 700°C.

In order to unravel this striking behavior, we combined ab-initio total energy calculations of the energetics of over 20 defect complexes with statistical mechanical and kinetic arguments to construct a meaningful and realistic theory of this system. Examples of some of the defects and defect complexes studied are shown in figures 2-4. In figure 5
Chapter 4. Semiconductor Surface Studies

<table>
<thead>
<tr>
<th>Pattern</th>
<th>number of defects</th>
<th>fraction: $n (10^{-3})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-DV</td>
<td>148</td>
<td>6.7</td>
</tr>
<tr>
<td>2-DV</td>
<td>78</td>
<td>3.5</td>
</tr>
<tr>
<td>1+1-DV</td>
<td>244</td>
<td>1.1</td>
</tr>
<tr>
<td>1+3-DV</td>
<td>6</td>
<td>0.3</td>
</tr>
<tr>
<td>2-DV</td>
<td>23</td>
<td>1.0</td>
</tr>
<tr>
<td>1+3-DV</td>
<td>54</td>
<td>2.5</td>
</tr>
<tr>
<td>1+3+1-DV</td>
<td>100</td>
<td>4.5</td>
</tr>
<tr>
<td>4-DV</td>
<td>0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**Figure 1.** The total number and fraction of various DV, DV clusters, and DV cluster complexes on the surface.

**Figure 2.** Total energy and restoring force as a surface dimer penetrates the bulk to attempt to form an interstitial dimer. (a) Schematic diagram illustrating three stages of the penetration process. Open circles represent the surface atoms and shaded circles, bulk atoms. (b) The force and formation energy of the dimer as a function of penetration depth. Note the absence of a metastable interstitial-dimer configuration.

we summarize the results of our total energy calculations for some of the most important defects and defect complexes.

Let us now address the problem of predicting the distribution of DV clusters and complexes on the Si surface. The first phase in the analysis of the defect distribution is to determine whether the kinetics of the anneal at 800-1100°C are sufficient to establish a thermal distribution among the defects or whether some trace remains of the initial distribution from cleavage. The kinetic pathways which contribute to the distribution after the initial anneal must have barriers sufficiently low that they may be expected to be traversed during the time scale of the anneal. Taking the waiting time for a transition event to be on the order of 1 minute, the attempt rate to be near the Si optic phonon frequency (the final answer is not sensitive to these values as we eventually take a logarithm), and the probability of a successful traversal to be the Boltzmann factor, we conclude the maximum barrier height for allowed paths in this experiment is 4.0 eV at the peak anneal temperature.

In figure 6, we show the pathways whose energy barriers we find to be near or less than the 4.0-eV cutoff. Note the defects we have considered in this work form a connected Markov chain which is tied to the sink and the source of DVs by diffusion of single DVs to the step edges. Therefore, the annealing process will be sufficient to remove all traces of the DV distribution of DV clusters and complexes.

After the initial anneal is completed, the sample is then cooled to room temperature before being scrutinized by the STM. Understanding the behavior of the distribution during this process requires a more detailed look at the transitions between the DV
defects. Three broad groups make up these transitions. The first group, indicated by the solid lines in figure 6, have relatively low energy barriers. These transitions all involve processes whereby an isolated dimer vacancy can migrate to an adjacent site. Using our estimates of the three basic types of second-layer rebonds, we find that all the transitions in this class traverse barriers of about 2.5 eV.

The second class of transition, indicated by the dashed lines in figure 6, have significantly higher barriers because they can occur only by disturbing the strong second-layer rebonding adjacent to the moving dimer. The barrier heights for these transitions are more difficult to estimate quantitatively, but our bond counting puts them in the range of 3.5-4.0 eV. The broken solid line in figure 6 represents an example of a third class of transition which we believe to be rare despite its energy barrier being in the first class. This is the transition by which the DV on the far right of the (1+2+1)-DV complex stretched in figure 5 migrates to the left, leaving behind a (1+2)-DV complex. This motion alone results in a high-energy metastable (1+2)-DV complex which can decay into the stable (1+2)-DV complex only by the independent migration of both second-layer rebonds across the 2-DV cluster. This particular transition therefore requires three independent motions to occur in concert (the initial dimer migration plus the migration of the rebonds), and so we suggest in our model that the attempt frequency for this transition may be so much reduced from the optic phonon frequency that this pathway is not relevant in the system.

We are now ready to understand the system's behavior as it cools from the anneal. As the system cools, the higher-energy transitions rapidly become inaccessible to the system, severing the Markov chain into two disconnected parts by removing the dashed pathways shown in figure 6. As a consequence, the system of complexes comprising the upper part of the chain in the figure loses thermal contact with the rest of the system as the temperature drops below 900-1100°C (corresponding to 3.5-4.0 eV). Although low barrier transitions among these complexes continue to occur down to about 570°C (the freeze-in temperature for the lower barrier transitions by the familiar kinetic argument), they cannot change the total number of these complexes, which will remain fixed as the system cools. These transitions, however, will affect the relative distribution among those complexes, which will track the surface temperature down to the lower freeze-in temperature.

The lower part of the chain, however, remains in contact with the DV source and the sink at the step edges through lower-energy single DV migrations and thus remains in overall thermal equilibrium with the surface until the lower freeze-in temperature.

In figure 7 we confirm the predictions of our theory with a plot of our calculated defect energies, $E_i$, against $-\ln(n_i/g_i)$, where $n_i$ is the experimentally observed defect density. We choose this scaling so that Boltzmann distributed data appear as straight.
Chapter 4. Semiconductor Surface Studies

Figure 6. Illustration of the Markov chain formed by the principal DV defects and the primary pathways among them. Dashed lines represent transitions with relatively large energy barriers (~3.5-4.0 eV) which cease to be relevant around 900-1100°C, while solid lines represent transitions of lower energy near 2.5 eV, which essentially shut off around 570°C. The transition with the broken solid line also has a barrier near 2.5 eV, but it is expected to be forbidden by a low attempt frequency.

Figure 7. Plot of the \( \text{ab initio} \) formation energy of each defect vs. the logarithm of the experimental defect density, \( n_i \), normalized by appropriate degeneracy factors, \( g_i \). Data with absolute Boltzmann distribution at 570, 900, and 1100°C fall along the straight lines indicated in the figure which pass through the origin. A relative Boltzmann distribution (see text for definition and discussion) at 570°C is sketched through the data clustered at the top of the figure. Note that the two groupings of data correspond exactly to the severed parts of the Markov chain in Figure 6.

4.3 Cross-sectional Scanning Tunneling Microscopy

Cross-sectional Scanning Tunneling Microscopy (X-STM) is a potentially powerful experimental tool for investigating interface structure in heteroepitaxial growth. An unfortunate problem with STM, however, has been the unambiguous identification of chemical species when several types of atoms are present in a variety of bonding configurations. For example, there are two serious questions which arise when considering an experiment using X-STM to investigate the initial stages of growth in an actual heteroepitaxial system such as GaAs on Si. First, it is not obvious that the isolated Ga and As atoms on the surface of a Si substrate could be uniquely identified, either by the contrast they produced in an X-STM image (induced by charge-density fluctuations) or by changes in the I-V signature of the tunneling current in the vicinity of such atoms. Second, the interactions between the charge-density fluctuations from the cleaved surface of the bulk and the growth Si-doped surface and their impact on the resultant X-STM image are not well-understood. For these two reasons, we have chosen to take a simplified approach as a first step in our investigation of the initial stages of heteroepitaxy of GaAs on Si. We have chosen first to examine, by theoretical methods, the chemical interactions taking place in the vicinity of a plane of impurity atoms embedded in a host crystal, and to predict whether or not unambiguous chemical identification can be performed on an isolated impurity...
atom in such a system. By choosing this prototypical "embedded-surface" approach, we can eliminate the additional complications provided by the growth surface. Because we are interested in the atomic interactions between Ga, As and Si, for this work we have chosen to examine the use of Si atoms to dope a GaAs crystal. Doping of semiconducting materials, in which dopant atoms are intentionally confined to a single atomic plane in the host crystal, has been widely applied to the exploration of quantum effects in two-dimensional structures (where the doping profiles can be narrower than the de Broglie wavelength of electrons in the material), and has also led to important technological developments such as high mobility transistors, non-alloyed contacts, Schottky-gate transistors, and doping superlattice devices such as spatial light modulators and lasers.

The results of our calculations of the electronic structure of the clean and doped GaAs(110) cleaved surface, and associated theoretical X-STM images at varying bias voltages, have led to several interesting predictions, including: (1) Si does not behave as a conventional donor on the surface; (2) the electronic state associated with Si lies deep in the gap and is very localized on the surface; (3) at negative bias voltages, the signature of the Si atom has enhanced intensity at 3rd-nearest neighbor sites; and (4) for positive bias, the signature of Si would be that of a missing atom on the surface. The latter result is shown in figure 8. These predictions have recently been experimentally confirmed by Professor E. Weber's group at the University of California at Berkeley.

4.4 Publications


Chapter 5. Step Structures and Epitaxy on Semiconductor Surfaces

5.1 Project Description

In this program, our goal is to obtain a detailed, experimental description of the temperature-dependent structure, phase-behavior, and morphology of semiconductor, especially silicon, surfaces. The results obtained in this way will be essential to theoretical efforts aimed at understanding silicon surface energetics. To this end, we employ high-resolution x-ray scattering, using an ultra-high-vacuum (UHV) apparatus for surface x-ray diffraction,1 which was funded in part by JSEP. The apparatus is located at the National Synchrotron Light Source (NSLS) and may be employed either at the MIT-IBM bending-magnet beamline, X20, or at the NSLS X25 wiggler beamline, which provides a five-fold intensity improvement compared to X20. The MIT UHV apparatus, in combination with either X20 or X25, constitutes the world's premier facility for surface x-ray scattering. In 1993, we continued to exploit its unique capabilities.

Specifically we have completed a comprehensive x-ray scattering study of the structure of the Si(113) surface. Current crystallographic techniques make it possible to determine the full three-dimensional atomic structure of a surface, with accuracies approaching 0.01 Å. Knowledge of the atomic structure is critical for evaluating calculations of surface energetics. It is surprising that the Si(113) surface has received much less attention up to now than the Si(001) and (111) surfaces. Nevertheless, the Si(113) surface is particularly interesting for a number of reasons which involve both possible technological applications and the basic physics of surfaces.

First, although Si(113) is a relatively high-index surface, it has been found to be remarkably stable with a surface energy comparable to that of the (111) or (001) orientation. Specifically, recent measurements of the shape of voids, created by He-ion implantation inside a Si crystal, have established that the surface energy of Si(113) is less than 1 percent higher than that of the Si(001) surface.2 Understanding the geometric and electronic structure of such a low energy silicon surface is an important challenge in surface physics. Reports of superior epitaxial growth on Si(113)3 make the structure of this surface and its temperature dependence especially worthy of study.

The rationale that Si(113) may be particularly suitable as substrate for epitaxial growth of III-V semiconductor films follows: An unreconstructed (113) surface may be envisaged as a sequence (001) terraces separated by (111) steps. For Si(113), the terraces are a single atomic row wide. The corresponding surface unit cell contains two different types of surface atoms: (a) two-fold coordinated atoms with two dangling bonds, deriving from the (001) terrace, and (b) three-fold coordinated atoms with a single dangling bond, from the (111) step (see figure 1a). Thus, there are two different adsorption sites. It seems likely that one of the sites will be preferred by one of the epitaxial

---

atomic positions at the Si(113) surface are described in section 2. Our efforts in this regard are coordinated with those of Professor J.D. Joannopoulos, who is presently carrying out a total energy calculation of the Si(113) (3x1) structure.

In addition, we have initiated a study of the orientational phase diagram of stepped Si(113) surfaces between 300 and 1200 K. Stepped (or vicinal) Si(113) surfaces afford the opportunity to deepen our general understanding of the morphology of crystal surfaces. As described in detail in section 3, we have found that steered Si(113) surfaces undergo a so-called facetting transformation. Facetting is accomplished through the rearrangement of surface steps, so that there are regions of the surface with a high step density, together with step-free Si(113) facets of macroscopic extent. We believe that the theoretical studies of the possible facetting of stepped Si(001) surfaces carried out by Professors A.N. Berker and J.D. Joannopoulos may lead to a detailed understanding of Si(113) facetting. Evidently, surface morphology is key to a number of technologically important processes, including epitaxial growth, etching, electromigration, and corrosion. An essential prerequisite for controlling these processes is an understanding of the equilibrium morphology of the surface. For example, our results demonstrate that large, step-free Si(113) surfaces may be prepared by high-temperature annealing. One may hope that the use of such surfaces as substrates for epitaxial growth might lead to interfaces of unprecedented perfection, which in turn may possess particularly favorable electronic transport properties.

5.2 The Structure of the (3x1) Reconstruction of the Si(113) Surface

X-ray diffraction measurements to determine the full three-dimensional atomic structure of the Si(113) surface at 800 K were performed at X20 using our UHV apparatus. Our initial measurements were carried out in a glancing incidence geometry (in-plane measurements), so that the information gained concerned the atomic positions projected onto the surface plane. Relative structure factors at 130 surface Bragg reflections were collected, from which was derived the pair correlation function for atoms within the surface unit cell. The pair corre-

5.3 Morphology of Stepped Si(113) Surfaces

Equilibrium facetting is a remarkable phenomenon in which a crystal surface increases its area to decrease its free energy. Facetting is accomplished through the rearrangement of surface steps, so that their distribution across the surface is no longer uniform. Instead, there are macroscopic regions of high and low step density. Our present understanding of equilibrium facetting is based upon a description of the surface free energy in which the surface orientation is a key thermodynamic variable in addition to the temperature. In this context, facetting is an example of thermodynamic coexistence among different phases with different orientations. Moreover, such a description motivates the construction of orientational phase diagrams, within which are recorded the phase boundaries between different surface phases as the temperature and surface orientation are varied.

Experiments carried out at X20A using our UHV apparatus reveal a complex phase diagram for silicon surfaces oriented between (113) and orientations tilted roughly 6 degrees away from (113) towards (001), as illustrated in figure 3.7 In figure 3, solid lines indicate phase boundaries and dashed lines indicate triple points. (The solid circles correspond to surface orientations found experimentally for the sample used in this study.) Within the range of temperature and tilt angle investigated, we have found three distinct orientational phases. Below $T_3 = 1130$ K, phase separation into the (113) orientation and the (114) orientation, which lies 5.7 degrees away from (113) towards (001) occurs for surfaces with mean tilt angle lying between (113) and (114). The scattering lineshapes of the (113) and (114) surfaces indicate that both are atomically smooth facets in this temperature range. However, at $T_3$ there occurs a phase transformation at which the (114) facet disappears and are replaced by a surface with a tilt angle of 6.5 degrees. Our measurements indicate that for this phase the scattering lineshape becomes increasingly broad with increasing scattering angle, which is the behavior expected for a rough surface. Therefore, we believe that the 6.5 degree-tilted surface corresponds to a rough phase. All three phases, namely the (113) facet, the (114) facet, and the rough phase, may coexist at $T_3$, which thus is a triple point temperature. Above $T_3$, there is coexistence between the (113) facet and the rough phase. For temperatures increasing above $T_3$, the tilt angle of the rough surface decreases with increasing temperature while the intensity of the scattering from the (113) facet decreases. Eventually, at 1200 K, the tilt angle of the rough surface reaches the value determined by the macroscopic cut of the sample, which for the data shown in figure 3 was 3.8.

---

Chapter 5. Semiconductor Surfaces

Figure 3. Orientational phase diagram of the Si surface near the (113) direction for tilt angles towards (001). Single-phase regions are hatched, and two-phase regions are plain (unhatched). Solid lines are boundaries between two phases, and dashed lines represent triple points. Solid circles represent tilt angles measured for a sample with a mean tilt angle (miscut) of 3.8 degrees.

degrees, and the (113) facet essentially disappears. On further increase in the temperature, the tilt angle remains fixed because the surface lies in a one-phase region of the orientational phase diagram shown hatched in figure 3. We have extrapolated the boundary of the (113) facet-rough phase two-phase coexistence region to zero tilt angle at $T_0 = 1245$ K. The character of the transformation from (113) facet to rough phase at this point is presently undetermined. Future experiments will study the interesting possibility that it is a critical point.

5.4 Publications


Thesis

Panel I: Application

Section 1: Atomic Structure and Nuclear Physics

Section 2: Plasma Models

Section 3: Electromagnetics

Section 4: Photo Astronomy
Section 1  Atomic, Molecular and Optical Physics

Chapter 1  Quantum Optics and Photonics

Chapter 2  Basic Atomic Physics
Chapter 1. Quantum Optics and Photonics

1.1 Data Storage Using Raman Induced Optical Spectral Holeburning

Sponsor
U.S. Air Force - Electronic Systems Division
Contract F19628-92-K-0013

Optical spectral holeburning in cryogenic materials has generated much recent interest because of its potential for high density optical storage, both in the time and frequency domain and in high speed optical processing. In holeburning materials, storage capacity is determined by the ratio of optical inhomogeneous to homogeneous linewidths (typically $10^6$) at a single position. Optical data rates can be as large as the optical inhomogeneous linewidth (usually in excess of several GHz). In addition, spatial resolution is determined by the optical wavelength, so that the total storage density can be very large, and entire holographic images can be processed at high speed. Although this is impressive, there is still room for substantial improvement.

We have developed a novel technique for enhancing the storage capacity of optical spectral holeburning materials, using Raman coherent population trapping. This technique is based on two properties of the Raman transparent state, namely, the equivalence of the Raman transparent state to a ground-state (spin) coherence and the sensitivity of this transparent state to the difference phase of the optical fields. The first property permits the optical excitation and detection of ground-state (spin) echoes, and the second property allows time-separated optical fields to interfere even after the optical coherence has decayed. Together, these two properties can be used to increase the storage capacity of optical spectral holeburning beyond the current fundamental limit given by the ratio of optical inhomogeneous to homogeneous linewidths.

Raman population trapping, illustrated in figure 1, allows both constructive and destructive interference to occur between time-separated optical fields, even though the optical coherence has long decayed. In the presence of ground-state inhomogeneous broadening, this permits time-domain optical storage. We have demonstrated this effect experimentally using a sodium atomic beam.

Figure 2 shows storage and recall of data using Raman induced spin coherence for a variety of conditions.

---

1. Rome Laboratory, Hanscom, Massachusetts.
2. Harvard University, Cambridge, Massachusetts.
3. Tufts University, Medford, Massachusetts.
Figure 1. (a) Schematic of the resonance Raman interaction. (b) The Raman interaction in the superposition state basis. The $|\rightarrow\rangle$ state is transparent to the resonant optical fields.

Figure 2. Observation of data storage and retrieval using Raman induced spin echo: (a) unequal bits, (b) equal bits, (c) out of phase bits, and (d) $\pi/2$ phase shifted bits.

To show the potential for increased storage density, note that the Raman-excited rf echo storage of figure 2 requires only low-intensity Raman-resonant optical pulses. Thus, for an inhomogeneously broadened optical transition, only atoms corresponding to a single optical spectral hole need participate. This means that many bits of optical data can be stored in a single optical spectral hole. The maximum number of bits per optical spectral hole is ultimately limited by the ratio of the inhomogeneous to homogeneous linewidths of the rf transition. Since the storage capacity of optical spectral holeburning is calculated based on only one data bit per homogeneous width spectral hole, it is clear that the ratio of optical inhomogeneous to homogeneous linewidths no longer defines the storage limit.

Raman population trapping can also be used to increase the storage capacity of time-domain optical holeburning. To see this, recall from figure 1 that the $|\rightarrow\rangle \leftrightarrow |e\rangle$ transition is optically excited and can therefore be used to perform time-domain spectral holeburning, just as any inhomogeneously broadened optical transition. In particular, consider the case of long-term data storage, employing Raman-resonant optical write and read pulses. In this case, long-term information storage is accomplished with a modulated population difference between the $|\rightarrow\rangle$ and $|\leftrightarrow\rangle$ states. But this population difference is also an rf coherence, which dephases in the presence of ground-state inhomogeneous broadening. As a result, the Raman resonant read pulse cannot produce optical echoes unless the ground-state coherence is first rephased, for example by an rf $\pi$-pulse. This suggests that it may be possible to selectively store and recall more than one data set by using time-domain ground-state spectral holeburning to store and retrieve various $|\rightarrow\rangle$ and $|\leftrightarrow\rangle$ state population modulations. These population modulations in turn correspond to the Fourier transforms of individual sets of optical data pulses.

Verification of multiple data set storage and recall, using both Raman and optical time-domain spectral holeburning is accomplished with the pulse sequence of figure 3. As shown, the input consists of two time-separated data/write pulse sets, where each set consists of a pair of Raman resonant optical data pulses followed by a Raman resonant write pulse. Rephasing is accomplished with an rf $\pi$-pulse, and optical echoes are produced by two Raman resonant read pulses applied at appropriate times after the $\pi$-pulse, as illustrated. Note that the rf field excites all the atoms in the material. If spatially localized re-phasing is needed, the rf field can be replaced by intense off-resonance Raman excitation.
been shown. For comparison, figure 3c shows calculated echoes for the case of conventional stimulated optical echo storage, modeled by a three-level system, with only one ground-state coupled to the excited-state, as illustrated. In this case, the ground-state populations cannot dephase so that the population modulations created by the two optical data sets are simply summed. Recall of this data at any time simultaneously recalls the data contained in both optical data sets, leading to the crosstalk evident in the echo signals. In the general case of unequal Rabi frequencies, the excitation pulse sequence of figure 3a will produce echo signals which look like a mixture of the two limiting cases of figure 3b and figure 3c (i.e., crosstalk will appear, roughly in proportion to the Rabi frequency difference).

In summary, we have shown how Raman coherent population trapping can be used to increase the optical storage capacity of optical spectral holeburning materials beyond the limit imposed by the ratio of optical inhomogeneous to homogeneous linewidths. This increased storage is based on the ability of Raman coherent population trapping to allow interference even after decay of the optical coherence.

1.2 Observation of Cooling Assisted Velocity Selective Coherent Population Trapping

Sponsor
U.S. Air Force - Electronic Systems Division
Contract F19628-92-K-0013

In recent years, there has been a great deal of interest in velocity selective coherent population trapping (VSCPT). However, since VSCPT occurs via a random walk in momentum space, in order to achieve significant VSCPT in three dimensions, it is necessary to precool and confine atoms close to the recoil limit. We recently predicted that this type of cooling and velocity confinement assisted VSCPT can be achieved in one, two or three dimensions when a $\Lambda$ system atom is excited by a

---


---

197
pair of blue detuned Raman resonant standing waves, with a phase difference of $\theta = \pi/4$.

We have recently observed this effect in one dimension using a beam of metastable He atoms.\(^9\)

Figure 4 illustrates schematically the experimental setup, where a diode laser array pumped LNA laser is used to excite the $2^3S_1 - 2^3P_1$ transition at 1.083 micron.

![Figure 4. Schematic diagram of experimental setup.](image)

In figure 5, the curves labeled A show the beam profile when the laser is off. With laser on resonance and $\theta = \pi/4$, we studied first the efficiency of VSCPT as a function of the saturation parameter, given by $s = r/(1 + 4\delta^2)$, where $r$ is the laser intensity in units of the saturation intensity, and $\delta$ is the detuning in units of natural linewidth. For $s = 2.0$, the result is shown by curve B, where VSCPT is

---


198 RLE Progress Report Number 136
manifested by the appearance of the two peaks at ± 2k. When s is reduced, the degree of VSCPT is decreased, as expected, since the pumping rate into the dark state is proportional to the mean excited state population, which in turn is proportional to s. Curve C shows the result for s = 0.4, where the VSCPT peaks are barely visible. We then blue detune the laser by one linewidth (Δ = 1) and adjust the intensity to get the same degree of saturation (s = 0.4). The resulting distribution is shown in curve D. As can be seen, there is strong precooling, with a width of ± 2k, in agreement with our previous prediction. In addition, the precooling enhances the degree of VSCPT compared to that in curve C, even though the saturation is the same.

Due to the constraints of the present setup, our interaction time is limited to only about 6 recoil times. While this is long enough to observe the precooling, it is too short to see the full extent of the enhancement of VSCPT due to the precooling. Nonetheless, our observation validates the mechanism we had predicted previously, rendering credence to its claim that this scheme would enhance the rate of VSCPT in three dimensions by a factor of 200. This mechanism would allow us to reach a temperature a factor of ten below the recoil limit in about 90 msec.

1.3 Suppression of Absorption of Resonance Fluorescence in a Folded Three-Level Atom

Sponsor

U.S. Air Force - Electronic Systems Division
Contract F19628-92-K-0013

In recent years, there has been a lot of interest in using the Λ system atom to reach a very cold temperature via velocity selective coherent population trapping (VSCPT). However, since VSCPT takes a long time, the system is bright initially. To determine what kind of density can be achieved in such a system, it is therefore necessary to look at the density limiting effects during the initial stage when significant VSCPT has not yet taken place. The most important effect is due to the heating caused by absorption of the emission of neighboring atoms.

We have recently found that this absorption is strongly suppressed in the Λ system under a very general set of conditions. Tabosa et al. considered various schemes that also experience such suppressions. However, the system we are considering is of great practical interest since a trap employing Λ system atoms may be cooled to a temperature substantially below the recoil limit. In addition, this result bodes well for the density achievable in a rectified force trap of Λ atoms.

Figure 6a shows how a true Λ system can be realized, and figure 6b illustrates the parameters of a Λ system. We consider, as an example, the case where the common detuning Δ is blue, as necessary for the scheme of cooling assisted VSCPT. The difference detuning, Δ is non-zero, as would be the case for the atoms that have not yet undergone VSCPT, and are moving relatively fast. Without affecting the basic feature of the result, we have, for clarity, also assumed that the ground state splitting matches Δ. Figure 7a shows the resonance fluorescence spectrum (RFS) of the Λ atom, including the delta function peak at the driving frequency. Figure 7b shows the corresponding gain spectrum. As can be seen, in contrast with a two-level atom, the strongest absorption peaks are far removed from the significant part of RFS. Figure 8 shows the gain spectrum in greater detail. Note that (figure 8c) there is no absorption at the central peak of RFS, and one of the big sidebands of RFS sees loss while the other sees gain. The net result is that the heating due to rescattering force in such a system would be far smaller than that in a two-level system.

---

Chapter 1. Quantum Optics and Photonics

Figure 6. (a) The ideal $\Lambda$ system. (b) Parameters of the $\Lambda$ system.

Figure 7. (a) Resonance fluorescence spectrum for $g_1 = g_2 = 5.0, \delta = 10.0, \Delta = \pm 4.0$, all in units of $\Gamma$. (b) Gain spectrum corresponding to Figure 7a, in arbitrary units.

Figure 8. Expanded view of the gain spectrum: (a) the two absorption peaks, (b) the three central peaks corresponding to the three biggest peaks in the fluorescence spectrum, (c) two very small gain peaks. Note the different vertical scales, normalized to the height of the large absorption peaks in (a).
We find that these results are independent of the sign of $\Delta$, and the relative values of the various parameters. Thus, this suppression of heating applies to the non-zero (positive and negative) velocity atoms in the initial stage of cooling assisted VSCPT. Note also that for the atoms that have undergone VSCPT, the gain spectrum would be similar to figure 7b, while the RFS spectrum vanishes. Thus, these zero velocity atoms see very small heating from the scattering of the hotter atoms. Finally, this scheme may be easily generalized to three dimensions using a $J = 1 \leftrightarrow J' = 1$ transition.

1.4 Inversionless Raman Lasing

Sponsor
U.S. Air Force - Electronic Systems Division
Contract F19628-92-K-0013

There has been much recent interest in the concept of inversionless lasing, primarily for application to high frequency lasers where it is difficult to achieve population inversion. Conventional Raman lasers, while inversionless in the sense that no excited state population is required to produce gain, are not very useful for up-conversion because they still require a population inversion between initial and final states.

Recently, we have studied an inversionless raman laser (IRL) which can not only be used for up-conversion, but has promising potential application to optical phase conjugation and laser frequency translation. The simplest form of IRL is illustrated in figure 9. Here, the Raman laser is pumped by the field at $\omega_a$ and gain is experienced for properly chosen probes at $\omega_a'$ and $\omega_a''$, where the low-frequency pumps at $\omega_a$ and $\omega_b$ empty the final state via coherent population trapping (CPT). This is analogous to gain without inversion (or gain with "hidden" inversion) in the double $\Lambda$ system, and in fact the calculated "optimum" gain coefficients are of the same order of magnitude. The difference is that it is possible to achieve the optimum gain coefficient over a much wider range of atomic systems. For example, there are no constraints on the relative populations of states $|e>$ and $|f>$ in order to produce gain.

Aside from up-conversion, IRL can be used for precise laser frequency translation. This is illustrated in figure 10a. Here, Raman lasing occurs at $\omega_a'$, where $\omega_a'$ acts as the pump and $\omega_a$ is a single, low-frequency optical pumping laser which creates the necessary initial to final state population inversion. Unlike the conventional Raman laser, the IRL scheme of figure 10a allows cw lasing to occur, even if states $|a>$ and $|b>$ are long-lived. For application to optical phase conjugation, the more general four-level IRL scheme of figure 10b is required. This is equivalent to the scheme in figure 4a except that there are two low-frequency pumps so that the necessary inversion is created by CPT, rather than simple optical pumping. This apparently minor distinction leads to significantly larger probe gain coefficients, especially for weak probes. To see why this is the case, recall that in a Raman laser, the gain can be explained in terms of the pump laser scattering off of the ground-state coherence. However, this ground-state coherence must first be generated by the pump and probe beams via CPT, and hence in the weak probe limit,

---


Chapter 1. Quantum Optics and Photonics

Magnetic field and two standing wave laser fields can provide such a beam splitter. Unfortunately, the required magnetic fields are undesirable in many interferometer applications, and few atomic systems have a suitable transition.

We show that a similar beamsplitter for a three-level atomic system can be obtained in the absence of a magnetic field as long as the excitation fields interacting with the two transitions are standing waves with a differential detuning. This effect can be explained in terms of atomic scattering from dressed state potentials which form a blazed grating, or as the result of a synchronized cycles between Rabi flopping and the precession between between excited states induced by the field detuning. This technique could be used on transitions other than J = 1 ↔ J' = 1.

Consider a three-state system in the basis |a⟩, |b⟩, |e⟩, where the fields interact with the transitions |a⟩ ↔ |e⟩ and |b⟩ ↔ |e⟩. For equal Rabi frequencies and a phase difference of π/2, the eigenenergies of the Hamiltonian as functions of the atomic position z are shown in figure 11. For states which approach the ground states when the Rabi frequencies tend to vanish, the potentials are given by ± (Δ/2 + |sin(kz)|), where Δ is the difference detuning, and k is the wave number. In contrast, the state which approaches |e⟩ has a nearly triangular potential.

Finally, the ability of IRL to perform optical phase conjugation is a consequence of the differential optical phase sensitivity of Raman gain in the double Λ system. For example, by choosing Ωα and Ω′α to be counterpropagating plane waves, and Ωβ to be the input signal beam, IRL generates gain at Ω′β only for a beam propagating in the conjugate direction.

1.5 Creating a Large Angle Coherent Atomic Beamsplitter without a Magnetic Field

Sponsor

U.S. Air Force - Electronic Systems Division
Contract F19628-92-K-0013

Many atom interferometer applications would benefit from large angle atomic beam splitters with narrow distributions. The interaction between a

Figure 10. Schematic of the four-level double Λ system exhibiting inversionless Raman lasing. (a) Limiting case of single low-frequency pump laser. (b) More general case of two low-frequency pump lasers.
Figure 11. Eigenenergies as a function of position in the standing waves.

Figure 12. Theoretical momentum distribution for a monoenergetic atomic beam. The interaction time is $7.5(2\pi/\Delta)$ and each Rabi frequency is $\sqrt{2} \Delta$. 
Chapter 1. Quantum Optics and Photonics

1.6 Stimulated Brillouin Scattering Ring Laser Gyroscope

Sponsor
Charles S. Draper Laboratory
Contract DL-H-441690

Research is in progress on a new fiber optic ring laser gyroscope based on two counter-propagating stimulated Brillouin scattering (SBS) lasers which are generated in the same fiber optic ring resonator. The use of SBS as the laser gain medium is crucial to the operation of this gyroscope because a conventional solid-state gain medium cannot support simultaneous bidirectional lasing due to gain competition between the counter-propagating lasers. The directionality of the SBS gain medium, however, prevents gain competition and thus allows stable, simultaneous bidirectional lasing.

In the presence of an inertial rotation normal to the plane of the ring laser's resonator, a difference frequency is automatically generated between the counterpropagating SBS lasers which is proportional to the applied rotation rate, as predicted by the Sagnac effect. The operation of this solid-state gyroscope is very similar to that of the bulk-optic ring laser gyroscope (RLG) based on the He-Ne gain medium. It should be noted that gain competition in a gaseous medium is avoided by using two partially overlapping, Doppler broadened gain media.

In contrast with either the interferometer or passive resonator gyroscopes, both the SBS RLG and the He-Ne RLG do not require any external means to measure the nonreciprocal phase shift that is induced by rotation.

Figure 14 shows a simplified schematic diagram of an SBS ring laser gyroscope. Light from a 1 mW He-Ne laser at 1.15 μm is split into two pump beams, labeled P1 and P2, frequency shifted by acousto-optic modulators, and coupled into counter-propagating directions of the same ring resonator. For maximum effective pump power inside the

---


resonator, the pump lasers are held at the centers of their respective cavity resonances using servos, not shown in the figure, and are matched to an eigen-polarization of the cavity resonance.

However, it is important to note that for a range of low rotation rates in figure 15b the difference frequency between the SBS lasers is zero. This "lock-in" zone is caused by the frequency locking of the counter-propagating SBS lasers through backscattering within the cavity and has been extensively studied in the bulk-optic RLG where it is removed using mechanical dither.

Mechanical rotation dither minimizes the amount of time that the gyroscope spends in the lock-in zone, thus reducing the errors due to lock-in. Because of its much smaller mass, mechanically dithering the fiber RLG would be much simpler than dithering the bulk-optic RLG.

Another method for removing lock-in is by generating the SBS lasers with a large frequency separation between them which would prevent the lasers from frequency locking. The tunability and directionality of the SBS gain medium allows the generation of the counter-propagating SBS lasers in different longitudinal modes of the resonator by either frequency shifting one of the pump lasers, or by using two separate pump lasers.

Figure 16a shows the difference frequency between two SBS lasers generated in different longitudinal modes of the same ring resonator, as a function of the applied rotation angle (figure 16b). In this case, the two lasers are separated by 9 longitudinal modes and, as seen in the figure, no lock-in is observed in the beat frequency.

Chapter 1. Quantum Optics and Photonics

While this mode of operation eliminates errors due to lock-in, there are a number of additional sources of error unique to it. One important additional error source is due to dispersion pulling of the SBS laser frequencies resulting from intensity backscattering of the pump lasers.

Consider a single pump laser, P1, in the counter clockwise (CW) direction and the resulting SBS gain curve and SBS laser, B1, in the clockwise (CW) direction, as shown in figure 17. If there is any pump backscattering, a small amount of pump will also be present in the CW direction of the cavity, labeled P1'. Though this pump will be below the SBS lasing threshold, it will still generate a small SBS gain curve, labeled B1'. In a single direction SBS laser, this extra gain curve has no effect since it is in the opposite direction of the SBS laser.

![Figure 17. Generation of an extra SBS gain curve due to pump backscattering.](image)

However, as shown in figure 18, if two counterpropagating SBS lasers, B1 and B2, are generated in the same cavity, this extra SBS gain curve, B1', due to its dispersion, will pull the frequency of the B2 SBS laser. Thus, as the intensity of P1 changes, or as the backscattering at the pump frequency changes, the size of the dispersion pull will also change and cause a bias drift in the gyroscope. Further, since the intensity backscattering in the cavity is highly variable, the size of this dispersion pull is also highly variable.

![Figure 18. Extra SBS gain curve due to pump backscattering.](image)

The effect of this additional SBS gain curve can be most easily demonstrated and measured by introducing a third pump laser into the cavity. In this way, the size and frequency separation of this third gain curve can be easily controlled. Figure 19 shows the effect of varying the intensity of the third pump laser on the beat frequency between the two SBS lasers. As seen in the figure, the frequency pull increases linearly with increasing pump power, as expected, and the maximum pulling was measured to be 680 Hz μW of input power. This is in good agreement with the predicted value of 700 Hz μW for this cavity.
To reduce the errors due to dispersion pulls, SBS laser frequency spacings that are comparable to the half-width of the SBS gain curve should be avoided. Thus, either very large frequency spacings, or very small frequency spacings should be used.

Backscattering at the SBS laser frequencies also can lead to frequency pulls on the SBS lasers themselves. In this case, the frequency pull is the result of backscatter-induced changes in cavity loss at the SBS frequency which cause a change in the self dispersion pull due to the SBS laser's gain curve.

Normally, the internal losses of the cavity at the SBS frequency are considered to be constant. However, variations in backscattering can lead to significant variations in the cavity loss as a function of the cavity environment. This loss variation will cause a corresponding change in the SBS gain, since the gain always equals the loss while the SBS laser is lasing. Thus, if the cavity mode is not exactly at the center of the SBS gain curve, this variation in the SBS gain will cause a variation in the dispersion pull of the SBS gain curve. For example, for a cavity mode 70 MHz from the center of the SBS gain curve with a 100 µW threshold, a 1 percent variation in the cavity loss will result in a 4.680 Hz offset change.

This source of drift can be eliminated by thermally tuning the SBS gain curve so that the cavity resonance is at the center of the gain curve, since the dispersion is zero there.

Presently, though thermally and disperively driven drifts appear to limit the performance of the separated frequency SBS fiber RLG, there are a number of other error sources which may become important as the performance improves. These errors include eigenpolarization mismatch between the SBS lasers which would reduce reciprocity between the SBS lasers, eigenpolarization mismatch between the pump and the SBS laser which could lead to polarization pulling, frequency pulls due to frequency dependent backscattering, and excess phase noise due to finite common mode suppression of cavity jitter.

In order to perform an independent test of the null stability, we are going to mechanically dither the SBS gyro. Mechanical dither, as in the case of the bulk-optic ring laser gyro, may very well prove to be the best alternative for eliminating lock-in.
Professor Daniel Kleppner, Associate Director of the Research Laboratory of Electronics
Chapter 2. Basic Atomic Physics

Academic and Research Staff
Professor Daniel Kleppner, Professor David E. Pritchard, Professor Wolfgang Ketterle

Visiting Scientists and Research Affiliates
Dr. Theodore W. Ducas, Dr. Fred L. Palmer, Dr. H. Joerg Schmiedmayer, Stefan Wehinger

Graduate Students

Undergraduate Students
Ilya Entin, Philip M. Hinz, Wan Morshidi, Amrit R. Pant, J. David Pelly, Abraham D. Stroock, Bridget E. Tannian, Stanley H. Thompson, John J. Wuu, Peter S. Yesley

Technical and Support Staff
Carol A. Costa

2.1 Testing Quantum Chaos with Rydberg Atoms in Strong Fields

Sponsors
National Science Foundation
Grant PHY 89-19381
Grant PHY 92-21489
U.S. Navy - Office of Naval Research
Grant N00014-90-J-1322

Project Staff
Michael W. Courtney, Hong Jiao, Neal W. Spellmeyer, Professor Daniel Kleppner

Interest in the behavior of Rydberg atoms in strong fields has resulted in the generation of high resolution experimental spectra and accurate calculations for atoms in electric and magnetic fields and also experimental data in the low field regime for electric and magnetic fields. The principal motivation for this work has been to understand atomic systems under external perturbations similar in magnitude to the unperturbed energy. Rydberg atoms are central to this research, because field strengths available in the laboratory are comparable with atomic fields of highly-excited states. We have chosen to work with hydrogen because it is the simplest system to study. Alkali metal atoms are also of interest, because they break the zero field degeneracy of hydrogen and are much more tractable experimentally.

The quantum mechanics of alkali Rydberg atoms in strong electric fields is now believed to be well understood. Much has been learned about diamagnetic hydrogen, but some important regimes remain to be explored, and our knowledge of hydrogen needs to be extended to other alkalis. For example, we know little about the alkali Rydberg atoms in parallel electric and magnetic fields and even less about them when the fields have arbitrary or perpendicular orientation.

With the exception of hydrogen in an electric field, the classical analogues of all these systems undergo a transition from order to chaos. In seeking to understand the connection between quantum mechanics and classical chaos, the problem of hydrogen in a magnetic field has been extensively studied. However, the more general problem of Rydberg atoms in strong fields provides a richer testing ground for theories describing "quantum chaos."

We have developed techniques for carrying out high resolution laser spectroscopy on the lithium

---

1 Physics Department, Wellesley College, Wellesley, Massachusetts.
atom in strong static fields. In addition, we have developed techniques for computing the spectra of alkali atoms in both strong electric and magnetic fields. The differences between lithium and hydrogen are minor for many cases of interest. In other cases, the core potential itself is the cause of classical chaos and becomes essential to consider.

In the experiment, a lithium atomic beam, excited by a laser, transfers atoms from the 2S state to the 3S state by a two-photon transition, and a second laser excites the atoms to Rydberg states.

The Rydberg atoms are detected by electric field ionization. Two methods for relating quantum spectra to the classical dynamics system are (1) periodic orbit theory and (2) use of statistical correlations. In the past year, we have made progress in testing both these methods which characterize quantum chaos.

Classically chaotic systems possess a proliferation of periodic orbits. Periodic orbit spectroscopy is a technique used to relate the quantum mechanical spectrum to the periodic orbits of the corresponding classical system.\(^2\)

The hamiltonian of hydrogen in a magnetic field is given by

\[
\mathcal{H} = \frac{p^2}{2} - \frac{1}{r} + \frac{1}{2} L^2 B + \frac{1}{8} B^2 \rho^2 \tag{1}
\]

This hamiltonian can be rescaled using the relations: \(r = B^{-3/2}r, \rho = B^{1/2}\rho, \text{ and } L = B^{-1/2}L\). This gives the scaled hamiltonian,

\[
\tilde{\mathcal{H}} = \frac{\tilde{\rho}^2}{2} - \frac{1}{\tilde{r}} + \frac{1}{2} \tilde{L}^2 + \frac{1}{8} \tilde{B}^2 \tag{2}
\]

where \(\tilde{\mathcal{H}} = B^{-3/2}\mathcal{H}\). As a result, the classical dynamics depends only on the parameter \(\tilde{\varepsilon}_B = B^{-3/2}\varepsilon_0\) and not on \(\varepsilon_0\) and \(B\) separately.

The semi-classical Bohr-Sommerfeld quantization condition for the two non-separable scaled coordinates \(\tilde{\rho}, \tilde{z}\) is

\[
\tilde{S} = (2\pi)^{-1} \oint (\tilde{p}_\rho \, d\tilde{\rho} + \tilde{p}_z \, d\tilde{z}) = nB^{-1/3}, \tag{3}
\]

where \(\tilde{S}\) is the scaled action of a closed classical orbit. As a result, \(\tilde{S} = nB^{-1/3}\) gives rise to equidistant lines on a scale \(B^{-1/3}\) in a spectrum taken at constant scaled energy \(\varepsilon_B\). In other words, each peak in the Fourier transform of a constant scaled energy spectrum corresponds to the scaled action of a periodic orbit.

This has been shown experimentally.\(^3\) In addition, Delos and his colleagues have developed the techniques of periodic orbit theory which use classical dynamics to compute the actions of the periodic orbits and semi-classical techniques to compute the strength of the periodicities. The strength of the periodicities is called the recurrence strength, because an outgoing spherical wave will recur at times corresponding to the periods of the periodic orbits of the classical system. Figure 1 shows the Fourier transform of a computed constant scaled energy spectrum compared with the results of periodic orbit theory (courtesy of Dr. J.A. Shaw of the College of William and Mary). The recurrences from periodic orbit theory are in good agreement, except for peaks near scaled actions of 4, 8, 12, 16, and 20, which are systematically much stronger in the periodic orbit theory results. These peaks are due to the 5th, 10th, 15th, 20th, and 25th return of the orbit in the z=0 plane. These orbits have a bifurcation very close to the scaled energy shown here which makes their recurrence strength difficult to accurately compute.

We are planning periodic orbit spectroscopy both experimentally and computationally. Periodic orbit spectroscopy requires varying the energy and the magnetic field simultaneously so that scaled energy \(\varepsilon_B\) remains constant. Since our computational techniques have been verified by our experiments, and since it is easier to accurately vary \(E\) and \(B\) while maintaining a constant \(\varepsilon_B\) numerically than experimentally, we have proceeded computationally in the regions where the computations have been verified. In addition, our experiment is limited to exciting only odd-parity states of lithium. Computationally, we have been able to explore both odd and even parity hydrogen and lithium. In this way, we

---


have made several discoveries with implications for periodic orbit theory.

It was previously believed that if an orbit lies on the node of the outgoing wavefunction, the recurrence strength is zero. Since odd parity wavefunctions have a node in the z=0 plane, it was believed that the recurrence strength of the orbits in the z=0 plane are zero. Our quantum computations show that the recurrence strength is small but non-zero. J.A. Shaw and J.B. Delos have derived a new semiclassical formula for the recurrence strength of a periodic orbit at a node of the quantum wavefunction. The results are compared with the quantum computation in figure 2.

The classical dynamics of both lithium and hydrogen in a magnetic field show a transition from order to chaos as the scaled energy is raised. However, lithium's core causes it to become chaotic at a much lower $\varepsilon_0$ than hydrogen. As a result, lithium is completely chaotic at $\varepsilon_0 = -0.6$, a value at which hydrogen is completely regular. This presents a paradox for periodic orbit theory, because it is known that the spectra of odd-parity diamagnetic lithium is nearly identical to that of hydrogen. It has also been shown that the energy level statistics for odd-parity lithium corresponds to a regular classical system. This is expected, since the odd parity wavefunctions have a node in the region of the chaos-causing core.

In contrast, the wavefunctions of even-parity lithium do not have a node at the core. As a result, the spectra of even-parity lithium have chaotic energy-level statistics and are very different from the spectra of even-parity hydrogen, which has regular energy level statistics. We arrive at a paradox,

---


5 Private communication from Dr. J.A. Shaw, College of William and Mary, Williamsburg, Virginia.


because classical dynamics does not "know" about parity. As a result, periodic orbit spectroscopy should yield the same periodic orbits for both odd and even parity hydrogen. Similarly, periodic orbit spectroscopy in lithium should give the same periodic orbits, regardless of parity. In both hydrogen and lithium, the recurrence strengths may vary between the odd and even parity cases, because these are computed by including parity in the semi-classical methods. However, since odd parity hydrogen and lithium have nearly identical spectra, we expect the periodic orbits of lithium and hydrogen to be nearly identical. This seems to be the case from looking at the recurrences in figure 3.

However, we do not expect a completely chaotic system to have the same periodic orbits as a regular system. In addition, from a quantum mechanical point of view, we have no reason to expect that the spectrum of even parity lithium will exhibit the same periodicities as even parity hydrogen. However, figure 4 shows that many of the recurrences are the same. This is a victory for periodic orbit theory. The case of even parity lithium also has many extra recurrences, indicating extra periodic orbits. This is a challenge to periodic orbit theory, because to resolve this paradox, periodic orbit theory must find these extra periodic orbits and compute their recurrence strengths in even-parity lithium. Furthermore, it also needs to be shown that all these extra periodic orbits have a very small recurrence strength in the odd parity case.

In the case of Rydberg atoms in a strong electric field, parity is not a good quantum number. Hydrogen in an electric field is regular and chaos is caused completely by the core. The Hamiltonian of hydrogen in an electric field is given by

$$\mathcal{H} = \frac{p^2}{2} - \frac{1}{r} + Fz,$$

(4)

and this Hamiltonian can be rescaled by substituting $r = F^{-1/2}r$, $p = F^{1/2}p$. As a result, the classical dynamics of hydrogen in an electric field depends only on the parameter $\xi_F = F^{-1/2}E$ and not on $E$ and $F$ separately. In analogy with the magnetic field case, each peak in the Fourier transform of a spectrum taken at constant $\xi_F$ corresponds to the scaled action of a periodic orbit. These classical scaling rules are exact for hydrogen and good approximations for lithium.

In the case of lithium in an electric field, we have done experiments for both the cases $m = 0$ and $m = 1$. The $m = 1$ case is regular and nearly identical to hydrogen, because the angular momentum barrier causes the wavefunction to have a node at the core.
The $m=0$ case is chaotic for all regions of interest. The experimental result for $\varepsilon_f = -0.3$ is compared with quantum computations for lithium and hydrogen in figures 5 and 6. Notice that lithium and hydrogen have many recurrences in common, but just as in the case of even-parity diamagnetic lithium, there are extra recurrences which need to be computed by periodic orbit theory. Several of these extra peaks occur near a scaled action of 14.

Spectral correlations have also been useful in characterizing spectra in regions of classical chaos. Computing the spectra in regions of total chaos has enabled the testing of a new class of universal correlations in spectra of chaotic systems as a parameter is changed. Our computations enabled MIT Professor Boris Altshuler's group to verify their new class of universal correlations on a real physical system. This new class of correlations has been confirmed in the chaotic regions of both diamagnetic hydrogen and lithium.

Figure 3. Recurrence strength for $\varepsilon_s = -0.6$, $m = 0$, odd parity. Notice that it shows the same periodic orbits for hydrogen and lithium.

---

Figure 4. Recurrence strength for $\varepsilon_0 = -0.6, m = 0$, even parity. Notice that it shows that lithium has many periodic orbits in common with hydrogen, but that there are extra periodic orbits present in lithium.

Figure 5. Recurrence strength of lithium for $\varepsilon_f = -3.0, m = 0$. The experiment is compared with quantum computations for lithium.
Chapter 2. Basic Atomic Physics

2.2 Millimeter-Wave Frequency Measurement of the Rydberg Constant

Sponsors
Joint Services Electronics Program
Contract DAAL03-92-C-0001
National Science Foundation
Grant PHY 89-19381
Grant PHY 92-21489

Project Staff
Joel C. DeVries, Dr. Theodore W. Ducas, Jeffrey R. Holley, Robert I. Lutwak, Professor Daniel Kleppner

The Rydberg constant $R_n$ relates the wavelengths of the spectrum of atomic hydrogen to practical laboratory units. It is a so-called fundamental constant, for it relates numerous quantum properties to laboratory units. Recent advances in optical wavelength metrology have enabled experiments using laser spectroscopy to measure $R_n$ with accuracy approaching two parts in $10^{11}$.

Although $R_n$ is the most accurately measured fundamental constant, forthcoming high-precision experiments which will depend on $R_n$ as an auxiliary constant, demand even more accurate measurement. Because the practical limit of measuring wavelengths, about 1 part in $10^{10}$, has been reached, progress in the measurement of $R_n$ requires measuring the frequency of spectral lines. Because the speed of light is now defined, by measuring $cR_n$, the "Rydberg frequency," a new standard is created.

Our approach involves preparing highly excited "Rydberg" states of atomic hydrogen, $n = 29$, and measuring millimeter-wave transitions to nearby states. Because the millimeter wave signal is generated coherently from a frequency standard based on an atomic clock, our measurement can make use of the high precision of frequency metrology.

The goals of our experiment are three-fold: First is the reevaluation of $R_n$ itself, providing an independent check, in a different regime, of other evaluations based on optical wavelength metrology. Second is the measurement of the ground state

---

Lamb shift. Because our measurements involve high angular momentum states for which the Lamb shift is extremely small, our results may be compared with optical measurements of transitions between low-lying states to yield an improved measurement of the Lamb shift. Third is to provide a frequency calibration of the spectrum of hydrogen, enabling the creation of a comprehensive frequency standard extending from the radio-frequency regime to the ultra-violet.

Our experiment employs an atomic beam configuration to reduce Doppler and collisional perturbations. Atomic hydrogen is excited to the low angular momentum \( n=29, m=0 \) state by two-photon stepwise absorption. The excited atoms are then transferred to the longer lived \( n=29, m=28 \) "circular" state by absorption of circularly polarized radio-frequency radiation.\(^{10}\) The atoms enter a region of uniform electric field in which the frequency of the resonant transition \( n=29, m=28 \rightarrow n=30, m=29 \) is measured by the method of separated oscillatory fields. The final state distribution is measured by a state-sensitive electric field ionization detector. The resonance is manifested by a transfer of atoms from the \( n=29 \) state to the \( n=30 \) state as the millimeter-wave frequency is tuned across the transition.

Figure 7 illustrates the apparatus. Atomic hydrogen or deuterium is produced by dissociating \( \text{H}_2 \) or \( \text{D}_2 \) in a radio-frequency discharge. The beam is cooled by collisions in a cryogenic thermalizing channel in order to slow the beam and thereby increase the interaction time. The atoms enter the circular state production region, where they are excited from the 1s ground state, through the 2p state, to the \( n=29, m=0 \) state by two-photon stepwise excitation. This is performed in an electric field to provide selective population of the particular \( n=29, m=0 \) level required for subsequent microwave excitation of the circular states. The electric field is then rapidly reduced to an intermediate value as the atoms pass through the center of a square configuration of four electrodes. The electrodes are excited by a 2 GHz RF source with a 90 degree phase delay between the adjacent pairs (figure 8). The circularly polarized field drives the atoms into the \( m=28 \) circular state through the stepwise absorption of 28 photons. A detector in the circular state production region monitors the efficiency of the optical excitation and angular momentum transfer processes.

After the atoms are prepared in the \( n=29 \) circular state, the beam enters the interaction region. Because Rydberg atoms interact strongly with external fields, accurate measurement of the energy level structure requires careful control of the interaction environment. Thermal radiation is reduced by cooling the interaction region to \(-10\)K by a liquid helium flow system. The ambient magnetic field is shielded out by a double-wall high-permeability shield. A small electric field, which defines the quantization axis of the atoms, is applied with high uniformity by field plates above the interaction region.

---

and below the atomic beam. The millimeter-waves intersect the atomic beam at two locations separated by 50 cm. The millimeter-wave optical system was described in a previous Progress Report.

The state distribution of the atoms emerging from the interaction region is analyzed by a state-selective electric field ionization detector. Within the detector, the atoms enter a region of increasing electric field produced by a pair of symmetric ramped plates held at constant potential. Atoms in different states are selectively ionized at different fields and the charged nuclei are detected at different positions.

We have previously presented initial data on a Ramsey interference curve. The past year was devoted to making improvements in the apparatus necessary for an accurate measurement at the level of 1 part in 10^11. To extract the information necessary to understand systematic sources of error, it is necessary to increase the signal-to-noise ratio of the experiment by a factor of 1000. We are accomplishing this by increasing the excitation efficiency of the atoms and decreasing the signal loss due to interaction with blackbody radiation. We have acquired an ultraviolet excimer pump laser to replace our aging YAG laser and have begun development of an ultraviolet dye laser system.

This system will operate at 120 Hz instead of 10 Hz, increasing the duty cycle of the experiment from about 5 percent to 60 percent. In addition, we expect the new laser system to provide a greater spectral density and, hence, improved efficiency of atomic excitation.

On the second front, we are redesigning the atomic beam environment to reduce the temperature of the radiation fields within the interaction region. In taking preliminary data, we used the method of crossed electric and magnetic fields to transfer the atoms from the low angular momentum state to the circular state.11 This method requires a magnetic field. Because the field could perturb the measurements, the production region had to be isolated from the interaction region. We believe that this configuration was the source of leaks of blackbody radiation into the interaction region. We have now successfully produced circular states without a magnetic field using the method of adiabatic rapid passage in a radio-frequency field (figure 9). The new atomic beam apparatus will incorporate the production, interaction, and detection of Rydberg states into one uniformly shielded region. We hope to complete the new interaction region and ultraviolet laser system in the near future, and begin performing high precision millimeter-wave spectroscopy.


Figure 8. Schematic diagram of the RF circular state production method.
2.3 Precision Mass Spectroscopy of Ions

Sponsors
Joint Services Electronics Program
   Contract DAAL03-92-C-0001
National Science Foundation
   Grant PHY 89-21769

Project Staff
Michael P. Bradley, Frank DiFilippo, Vasant Natarajan, Fred L. Palmer, Abraham D. Stroock, Professor David E. Pritchard

In 1993, we developed and applied a new technique for precisely comparing the masses of two ions of widely differing atomic weight. Together with previously developed techniques, we completed a program of mass comparisons designed to determine ten atomic masses important for determination of fundamental constants or metrology. This program resulted in a new high accuracy atomic mass table with typical accuracy about 10^{-10} (an improvement of one to three orders of magnitude over previously accepted values). It also provided numerous quantitative tests of our internal consistency that virtually eliminate the possibility of unknown systematic errors. This capability has allowed us to contribute to several important experiments in both fundamental and applied physics, including:

- Contributing to a recalibration of the current x-ray wavelength standard by weighing the energy differences that go into γ-rays with wavelengths that will be measured at the National Institute of Standards and Technology;
- Determination of the molar Planck constant $N_A \hbar$, by weighing γ-rays; this will also provide an alternate determination of the fine structure constant;
- Determination of the atomic weight of $^{28}$Si, which is part of a program to replace the "artifact" kilogram mass standard by a crystal of isotopically pure $^{28}$Si, effectively creating an atomic standard of mass.

In addition, we are laying the groundwork for two future advances: (1) increasing the versatility of our apparatus by loading ions externally, and (2) improving our level of precision another order of magnitude. The latter goal will be achieved by comparing two simultaneously trapped ions (to eliminate the problems due to field drift) and using squeezing techniques to reduce the influence of thermal noise on the measurements. These advances will allow us to:

- Measure the $^3$He$^+$ - $^4$He$^+$ mass difference, important in ongoing experiments to determine the electron neutrino rest mass;
• Improve some traditional applications of mass spectrometry due to our orders of magnitude improvement in both accuracy and sensitivity;

• Determine the molar Planck constant \( N_A \), by measuring the atomic mass and recoil velocity of an atom that has absorbed photons of known wavelength; and

• Determine excitation and binding energies of atomic and molecular ions by weighing the small decrease in mass, \( \Delta m = E_{\text{bound}} c^2 \) (we must reach our ultimate goal of a precision of a few parts in \( 10^{-12} \) to make this a generally useful technique).

In our experimental approach, we measure ion cyclotron resonance on a single molecular or atomic ion in a Penning trap, a highly uniform magnetic field with confinement along the magnetic field lines provided by much weaker electric fields. We monitor the axial oscillation of the ion by detecting the currents induced in the trap electrodes. Working with only a single ion is essential because space charge from other ions leads to undesired frequency shifts. This work in trapping and precision resonance draws on techniques developed by Hans Dehmelt at the University of Washington and Norman Ramsey at Harvard, for which they shared the 1989 Nobel Prize.

We have developed techniques for driving, cooling, and measuring the frequencies of all three normal modes of Penning trap motion. Thus, we can manipulate the ion position reproducibly to within 30 microns of the center of the trap, correcting for electrostatic shifts in the cyclotron frequency to achieve greater accuracy. We use a \( \pi \)-pulse method to coherently swap the phase and action of the cyclotron and axial modes.\(^\text{12}\) Therefore, although we detect only the axial motion directly, we can determine cyclotron frequency by measuring the phase accumulated in the cyclotron motion in a known time interval.

In the past two years, we have rebuilt our apparatus with a new Penning trap and quieter rf SQUID detector. We have implemented a new signal processing algorithm to improve our phase estimation by a factor of two to three. We can measure the phase of the cyclotron motion to within 10 degrees, leading to a precision of \( 1 \times 10^{-10} \) for a one minute measurement. Our entire ion-making process has been automated, and the computer can cycle from an empty trap to having a cooled single ion in about three minutes under optimal conditions. With the spatial imperfections in the electric and magnetic fields made as small as possible our systematic errors are well below \( 5 \times 10^{-11} \). However, the typical statistical fluctuation in our magnetic field is \( 2.4 \times 10^{-10} \) between measurements (over the smooth drift in the field). Thus, with the ability to achieve ~20 alternate loadings of two different ion species in one night, our overall uncertainty can be as small as \( 8 \times 10^{-11} \) (see figure 10). We have found that the distribution of field variation is not Gaussian, but instead has additional outlying points. This has led us to use robust statistical analysis of field fluctuations; in particular, a generalization of least squares fitting called "M-estimates" in which outlying points are deweighted in a manner determined by the observed excess number of outliers. Using this analysis has eliminated arbitrary decisions about dropping "bad points" from our data sets and has increased the stability of the fit.

![Figure 10. Cyclotron frequency as a function of time for alternate \( \text{N}_2 \) and \( \text{CO}^+ \) ions in our Penning trap. The frequencies are obtained after a 50s integration of cyclotron phase (see text). The solid line is a polynomial fit to the drift in the field common to both ions.](image)

Chapter 2. Basic Atomic Physics

masses by direct comparison to carbon. Indeed, we have compared CD$_4$ and CD$_3$ to C$^+$ to obtain two independent determinations of the atomic weight of deuterium. In general, mass ratios of doublets cannot be inverted to obtain absolute masses without comparing some species to carbon, which by definition has an atomic mass of exactly 12. This is traditionally done using the ratios of very heavy ions such as C$_{13}$H$_{14}$/C$_{12}$H$_{25}^+$ to determine H, or C$_2$D$_2$/C$_3$D$_4^+$ to determine D. We have circumvented this difficulty by developing a scheme involving only doublet comparisons that uses the known Ar*/Ar$^{**}$ ratio to obtain the absolute masses of several light atoms—H, D, N, O, Ne, Ar—to about an order of magnitude better than the current standard mass table (see table 1). The value of the deuterium mass obtained using this method is consistent with the value from the direct non-doublet comparisons, giving us high confidence in our results.

An extensive series of quantitative systematic checks of our results involving repeated checks of identical ratios (some interspaced by a significant reduction of our field inhomogeneities); checks of circular ratios (e.g., A/B, B/C, and C/A); checks of related ratios (e.g., CO/N$_2$ and CO$_2$/N$_2$O); and known ratios (e.g., Ar*/Ar$^{**}$) has shown us that there are no unknown systematic errors in our results. H and D have been measured by several independent routes, because the mass of many interesting ions can be determined by comparison with a suitable organic compound of the same nominal mass.

To extend the applicability of our current mass measurement capabilities, we have built an external ion source and associated optics that can produce and select ions with a resolution of 0.1 in mass-to-charge ratio. When this comes on line, we should be able to greatly reduce the percentage of "bad" ions (wrong nominal atomic weight) and also alleviate the problem of residual neutral gas when using volatile species such as helium.

In order to proceed to much higher levels of mass measurement precision, we must eliminate the random fluctuations of the magnetic field in our measurements. We have explored two ways to achieve this by putting two ions in the trap at the same time and have decided to pursue a method

<table>
<thead>
<tr>
<th>Atom</th>
<th>Mass (amu)</th>
<th>Accepted values (amu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1$H</td>
<td>1.007 825 031 6 (5)</td>
<td>1.007 825 035 0 (120)</td>
</tr>
<tr>
<td>n</td>
<td>1.008 664 923 5 (23)</td>
<td>1.008 664 919 0 (140)</td>
</tr>
<tr>
<td>$^2$H</td>
<td>2.014 101 777 9 (5)</td>
<td>2.014 101 779 0 (240)</td>
</tr>
<tr>
<td>$^{13}$C</td>
<td>13.003 354 838 1 (10)</td>
<td>13.003 354 826 0 (170)</td>
</tr>
<tr>
<td>$^{14}$N</td>
<td>14.003 074 004 0 (12)</td>
<td>14.003 074 002 0 (260)</td>
</tr>
<tr>
<td>$^{15}$N</td>
<td>15.000 108 897 7 (11)</td>
<td>15.000 108 970 0 (400)</td>
</tr>
<tr>
<td>$^{16}$O</td>
<td>15.994 914 619 5 (21)</td>
<td>15.994 914 630 0 (500)</td>
</tr>
<tr>
<td>$^{20}$Ne</td>
<td>19.992 440 175 4 (23)</td>
<td>19.992 435 600 0 (22000)</td>
</tr>
<tr>
<td>$^{28}$Si</td>
<td>27.976 926 532 4 (20)</td>
<td>27.976 927 100 0 (7000)</td>
</tr>
<tr>
<td>$^{40}$Ar</td>
<td>39.962 383 122 0 (33)</td>
<td>39.962 383 700 0 (14000)</td>
</tr>
</tbody>
</table>

Table 1. Atomic mass table. The center column lists the atomic masses determined from our experiment, and the right column lists the accepted atomic masses from the 1983 evaluation. [A.H. Wapstra and G. Audi, "The 1983 Atomic Mass Evaluation (I). Atomic Mass Table," Nucl. Phys. A 432:1 (1985).] For the purpose of comparison, zeros have been added to the numbers in the right column so that the number of digits are equal.
we described last year.\textsuperscript{13} If this works well, the primary source of measurement noise will be the special relativistic mass shift due to thermal fluctuations in cyclotron amplitude.

We have proposed a scheme of classical squeezing with parametric drives to reduce amplitude fluctuations.\textsuperscript{14} In 1993, we were successful in demonstrating a simplified version of the quadrature squeezing of the thermal motion of the ion. Figure 11 shows how the rms fluctuations of the cyclotron amplitude vary when we apply a cyclotron pulse of adjustable phase to a previously squeezed distribution. Note that at $\theta = 90$ degrees, the fluctuations are reduced to approximately one half of their unsqueezed value.\textsuperscript{15}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{squeezing_graph.png}
\caption{Demonstration of classical amplitude squeezing. The observed amplitude noise is plotted as a function of the relative phase. At a drive phase of 90 degrees, the noise is minimized. A best fit to the data shows that a factor of two reduction has been achieved.}
\end{figure}

\textbf{Thesis}

Natarajan, V. \textit{Penning Trap Mass Spectroscopy at 0.1 ppb.} Ph.D. diss., Dept. of Physics, MIT, 1993.

\section{2.4 Atom Interferometry}

\textbf{Sponsors}

Joint Services Electronics Program
Contract DAAL03-92-C-0001
U.S. Army - Office of Scientific Research
Grant DAAL03-92-G-0229
U.S. Navy - Office of Naval Research
Grant N00014-89-J-1207

\textbf{Project Staff}

Michael S. Chapman, Christopher R. Ekstrom, Troy D. Hammond, Amrit R. Pant, Richard A. Rubenstein, Dr. H. Joerg Schmiedmayer, Bridget E. Tannian, Stefan Wehinger, Professor David E. Pritchard

During 1993, we refined our atom interferometer\textsuperscript{14} and started to perform experiments with spatially separated beams. The interferometer is now operating with smaller period gratings,\textsuperscript{17} providing greater beam separation. The experiments were performed with the aid of an interaction region that inserts a thin metal foil between the beams (figure 12). This allowed us to manipulate the atomic wave function in one arm of the interferometer without affecting the other portion.


Chapter 2. Basic Atomic Physics

Figure 12. A schematic of our atom interferometer showing the active vibration isolation system and the interaction region (not to scale). The 10 $\mu$m copper foil is shown between the two arms of the interferometer (thick lines are atom beams). The optical interferometer (thin lines are He-Ne beams) is used for active vibration isolation. The 200 nm period atom gratings are indicated by a vertical dashed line, and the 3.3 $\mu$m period optical gratings by a vertical solid line.

Approximately six atom interferometers have been demonstrated since our pioneering demonstration of an interferometer\(^\text{16}\) made with three nanofabricated atom gratings.\(^\text{17}\) These instruments will make possible new types of experiments involving inertial effects, novel studies of atomic and molecular properties, and tests of fundamental physics. They could also make possible the development of ultrasmall structures using atom lithography. Specifically:

- The relatively large mass and low velocity of atoms makes atom interferometers especially sensitive to inertial effects such as rotation, acceleration, and gravity. Sagnac rotation has been observed in accord with theoretical predictions\(^\text{18}\) and sensitivity to gravitational acceleration at the $3 \times 10^{-8}$ level has been demonstrated.\(^\text{19}\) Atom interferometers may become the best absolute accelerometers and gravimeters within the next few years.

- Atom interferometers can be applied to a number of experiments in fundamental physics: tests of quantum mechanics such as the Aharonov-Casher effect;\(^\text{20}\) geometric phases and the measurement process; measurement of the equality of proton and electron charges; and a precise measurement of the momentum of a photon. This latter measurement should produce a new high precision value for the molar Planck constant $N_A\hbar$.

- Interferometers for atoms and molecules will provide more accurate ways to measure intrinsic properties of these particles, like their polarizability. They will also open up new areas of study, such as measurements of the "index


of refraction" of a gas for a propagating particle beam.

The key component of our interferometer is the set of three matched transmission diffraction gratings with 2000 Å period which we constructed at the National Nanofabrication Facility (NNF) at Cornell University. The process that we developed allows fabrication of precisely positioned openings in thin silicon nitride membranes mounted in silicon frames. Grating patterns in the membranes are generated by electron beam lithography, making the process quite versatile. This process was used to create a variety of diffraction gratings used in the interferometer. In addition, several zone plates (atom lenses) were also built and demonstrated.

During 1993, we made several modifications to our atom interferometer, including the addition of a new computer interface that allows a Macintosh Quadra with Labview software to control the experiment and analyze the data in real time. These have resulted in larger signals and greater reliability for this instrument, which is the only atom interferometer in which interfering beams are physically separated. We have exploited this capability to perform several unique experiments.

We have used this interferometer to determine the polarizability of the ground state of sodium to 0.3 percent by applying a uniform electric field to one component of the atom wave and measuring the resulting phase shift of the interference pattern. This type of measurement determines the ground state polarizabilities rather than the difference of ground and excited state polarizability which can be measured with spectroscopic or interferometric techniques in which the interfering paths are not spatially isolated. The precision of our measurement was about five times better than previous determinations.

In another experiment, performed with a new interaction region that can confine a gas sample so that only one of the atom beams in the interferometer passes through it, we observed both the attenuation and the phase shift of an atomic sodium matter wave when it was transmitted through a number of atomic (He, Ne, Ar, Kr, Xe) and molecular gases (N₂, CO₂, NH₃, and H₂O). From the perspective of wave optics, the passage of a wave through a medium is described in terms of a complex index of refraction where \( N \) is the density of the medium and \( n \) is the scattering amplitude. The phase shift (attenuation) is proportional to the real (imaginary) part of the forward scattering amplitude, allowing us to determine both of these quantities.

This is the first time that the real part of the forward scattering amplitude has been measured for atom-atom scattering. An indication of the relative sensitivity of the real part to the interatomic potential is that this quantity varies over more than an order of magnitude in the systems studied. In contrast, the imaginary part of the forward scattering amplitude, which is proportional to the total scattering cross section, varies by about a factor of less than two. We therefore considered several simple models that predict the ratio of real to imaginary parts of the forward scattering amplitude in order to elucidate the basic physics of this quantity. All of these models are based on the partial wave expression for scattering by a central potential, a treatment appropriate for our situation where the wavelength of the scattered wave is much smaller than the size of this potential.

If the collision is modeled as a hard sphere, the partial wave scattering phase shifts are nearly

---


random, and the ratio of the real to imaginary parts of the forward scattering amplitude is predicted to be about 0.01 for He, not too far from the measured value (0.1). A van der Waals potential should produce a ratio of 0.79, in rough accord with the rare gas atoms with the largest polarizability, Xe(0.7) and Kr(0.6). Understanding the results for Ne(1.0) and Ar(0.5) will require fuller treatment, based on more realistic interatomic potentials.

Recently, we have changed the velocity of the sodium in our interferometer by changing the mass of the carrier gas in our seeded supersonic oven. This has enabled us to measure the dispersion of the index of refraction (i.e., its dependence on deBroglie wavelength) over a factor of two in wavelength.

We have also applied our atom optics techniques to molecules. Molecules were produced by optimizing the operating parameters of our seeded supersonic source so that sodium dimers constituted 30 percent of the detected beam intensity. The atoms were then removed from the beam by deflecting them with resonant laser light which was frequency stabilized to our atomic beam. Figure 14 shows diffraction of molecules and atoms and diffraction of only molecules from a 100 nm period diffraction grating.

We have used this molecular beam in our separated beams interferometer, measuring the ratio of real and imaginary parts of the forward scattering amplitude for collisions of these dimers with He and Ne. The results are the same within experimental error as for atomic sodium.

Lastly, we have demonstrated the Talbot effect, the self-imaging of a periodic structure, with atom waves. We have measured the successive recurrence of these self-images as a function of the distance from the imaged grating. This is a near field interference effect that has several possible applications in the rapidly growing field of atom lithography.

2.4.1 Publications


Chapter 2. Basic Atomic Physics


Thesis


2.5 Cooling and Trapping Neutral Atoms

Sponsors

Joint Services Electronics Program
Contract DAAL03-92-C-0001
U.S. Navy - Office of Naval Research
Grant N00014-90-J-1642

Project Staff

Michael R. Andrews, Kendall B. Davis, Ilya Entin, Philip M. Hinz, Michael A. Joffe, Marc O. Mewes, Wan Morshidi, J. David Pelly, Stanley H. Thompson, John J. Wu, Peter S. Yesley, Professor Wolfgang Ketterle, Professor David E. Pritchard

Our current objective is to obtain samples of atoms at very high density and ultra-low temperatures. This goal is pursued by using a high flux slower for atoms, a light trap to stop and compress the atoms, and a magnetic trap for the final confinement and cooling.

Experiments with dense samples of cold neutral atoms promise exciting new discoveries in basic and applied physics. Due to the considerably reduced thermal motion of atoms, they are ideal for high resolution spectroscopy and for more accurate atomic frequency standards. In addition,

- Collisions of ultra-cold atoms in such samples are characterized by a long deBroglie wavelength and are dominated by weak long-range interactions. Since the collision duration for slow atoms greatly exceeds the radiative decay time, stimulated and spontaneous radiative transitions can take place during the collision. Slow collisions are therefore fundamentally different from fast collisions studied so far and are becoming an exciting new field of atomic physics.24
- High density samples of atoms open possibilities for observing quantum collective effects such as Bose-Einstein condensation and collectively enhanced or suppressed radiative decay.

In 1993, we completed the work on a dark light trap which enabled us to confine more than $10^{16}$ atoms at densities one to two orders of magnitude higher than achieved before. The key feature of the dark trap is that the atoms spend most of their time in a dark hyperfine state which does not scatter the trapping light.25

The main focus of our work has been to obtain high densities and long trapping times in a magnetic trap which was loaded from the dark light trap. The product of density and trapping time is the crucial parameter for achieving evaporative cooling.

2.5.1 Magnetic Trapping of Sodium Atoms

We have captured up to $10^{11}$ atoms in a dark light trap and transferred them into a magnetic trap. By counting the number of atoms before and after the transfer, we determined the transfer efficiency to be 25 percent—in coarse agreement with the 33 percent expected for trapping one out of three equally populated $F=1$ hyperfine states. The transfer was accomplished in a "mode-matched"

---


way—both confinement volume and temperature were preserved by carefully adjusting the strength of the magnetic trapping field. Atomic densities in the magnetic trap were $10^{11}$ cm$^{-3}$.

Long trapping times were achieved by considerably improving the vacuum. Optimized differential pumping between the oven and trapping chambers allows pressure ratios of 1000. Ultimate pressures are below $10^{-10}$ Torr, even with the intense atomic beam on, resulting in trapping times approaching one minute. Trapping times in the light trap are longer by roughly a factor of five because of larger trap depth.

2.5.2 Dark Molasses

The temperature of atoms captured in a spontaneous light force trap can be further reduced by applying a cooling scheme called polarization-gradient molasses after switching off the magnetic field. Until now, one has encountered increased ultimate temperatures for the highest densities of atoms.26

We are currently investigating the idea that these high density effects can be avoided by using a dark version of molasses where most of the atoms are optically pumped into a dark hyperfine state, with only a small fraction of them in the optically cooled hyperfine state. This process ensures transparency of the atom cloud for the cooling light despite very high column densities. So far, we have successfully cooled samples to 60 $\mu$K. This temperature is below the Doppler limit of 240 $\mu$K, but higher than temperatures observed for Na at very low atomic densities (20 $\mu$K) using polarization gradient molasses.27 We are currently studying the limits in the ultimate temperature of dark molasses. Our atom clouds have optical densities of 200 for resonant repumping light and would be opaque for the cooling light without optical pumping into a dark state.

2.5.3 Rf Induced Evaporative Cooling of Atoms

We hope to reach the nK regime using evaporative cooling instead of optical cooling. Evaporative cooling of atomic hydrogen below the recoil limit has been achieved by Professors Kleppner and Greytak,28 but their method is highly specialized and limited exclusively to hydrogen. We believe that it is important to pursue similar goals with other atoms, not only because different atomic species such as the alkalis will show different physical effects, but also because of their ease of manipulation and observation. The cryogenic methods that are employed for hydrogen are cumbersome in contrast to the laser cooling methods that are used in the early stages of our cooling technique. Furthermore, there is evidence that the collisional cross sections for evaporative cooling are more favorable for alkali atoms than for hydrogen.29

Evaporative cooling is accomplished by repetitively removing the high energy "tail" of the thermal distribution of atoms in the trap. The remaining atoms then cool collisionally as the high energy tail is repopulated. The essential condition for evaporative cooling is that the collision rate be sufficiently high for many collisions to occur within the lifetime of the atoms in the trap. Until now, this has not been possible using laser-cooled atoms because the atomic densities have been too low. We are confident that our dark cooling and trapping techniques provide sufficiently high densities to observe evaporative cooling.

Beside high initial density and long trapping times, evaporative cooling requires a method for selectively removing hot atoms from the trap. Our approach is to use resonant rf radiation which selectively spinflips hot atoms and thus removes


them from the trap.30 The advantage of this strategy over simply lowering the magnetic trap depth is that confinement is not weakened. So far, we have tested the rf setup and shown that we can change the energy distribution function of trapped atoms with rf radiation.

2.5.4 Work In Progress

- Spin Flip Zeeman Slower
  We currently load our trap using an inverted "Zeeman slower" that has produced the largest flux of slow atoms achieved to date.31 We have a new design that may improve the effective flux by one or two orders of magnitude by overcoming effects due to transverse heating. The key new idea is to pump the atoms from the weak field seeking hyperfine level into the strong field seeking one during the slowing process.

- Magnetic Trap with High Field Gradients
  We have designed a new magnetic trap which will produce magnetic field gradients up to 1000 G/cm, a factor of ten improvement over our current setup. Such high field gradients will be used for adiabatic compression of the trapped atoms. This will result in an increased speed of evaporative cooling.

2.5.5 Publications


Chapter 2. Basic Atomic Physics

\textbf{Theses}


\textbf{2.5.6 Meeting Paper}

Section 2  Plasma Physics

Chapter 1  Plasma Dynamics
Chapter 1. Plasma Dynamics

Academic and Research Staff
Professor George Bekafi, Professor Abraham Bers, Professor Bruno Coppi, Professor Jonathan S. Wurtele, Dr. Stefano Miglioulo, Dr. Abhay K. Ram, Dr. Barrett Rogers, Dr. Linda E. Sugiyama, Ivan Mastovsky

Visiting Scientists and Research Affiliates
Dr. Carson C. Chow, Dr. Vladimir Fuchs, Dr. Paolo Detragiache, Dr. Marco Nassi, Dr. Jiří Ullschmied, Franco Carpignano, George M. Svolos

Graduate Students

Undergraduate Students
Beth L. Chen, Jonathan C. Doan, David H. Hijirida, Hana Ohkawa, Kenneth N. Ricci

Technical and Support Staff
Felicia G. Brady, Laura B. Doughty, Catherine Lorusso, Laura M. von Bosau, Miriam Weiner

1.1 Relativistic Electron Beams

Sponsors
National Science Foundation
Grant ECS 89-02990
U.S. Air Force - Office of Scientific Research
Grant F49620-93-1-0108
U.S. Army - Harry Diamond Laboratories
Contract DAAL02-92-K-0037
U.S. Department of Energy
Grant DE-FG02-91ER-40648
U.S. Navy - Office of Naval Research
Grant N00014-90-J-4130

Project Staff

1.1.1 Observation of Frequency Chirping and Phase of a Free Electron Laser Amplifier

Temporal variations of frequency and phase are important parameters of pulsed free electron lasers (FEL) in many proposed applications such as, for example, in radar and communications and as drivers in novel high gradient accelerators. Such temporal variations (chirping) have been observed in wave propagation in passive media and can occur as a result of a time-dependent dielectric coefficient caused by changes in plasma density, or as a result of a rapidly moving plasma-vacuum

1 University of Colorado, Boulder, Colorado.
2 Centre Canadien de Fusion Magnétique (CCFM), Quebec, Canada.
3 E.N.E.A., IEN "Galileo Ferraris", Turin, Italy.
4 Institute of Plasma Physics, Czech Academy of Sciences, P.O. Box 17, 182 00 Prague 8, Czech Republic.
5 A.S.P. Associazione per lo Sviluppo Scientifico e Tecnologico del Piemonte, Turin, Italy.
boundary. In this paper, we report frequency and phase shift measurements in an active, lasing medium in which the wave amplitude exhibits exponential growth. Here, the time-varying dielectric medium represents the interaction between the pulsed relativistic electron beam of the free electron laser (FEL), the magnetic wiggler, and the copropagating electromagnetic wave launched into the system. The theoretical and computational modeling is described elsewhere.

A schematic of the FEL amplifier is shown in figure la. A mildly relativistic electron beam (750 keV) is generated by a Marx capacitor bank (Physics International Pulserad 110A). The electrons are emitted from a hemispherical graphite cathode by an explosive field-emission process. The graphite anode acts as an emittance selector, allowing only a small fraction of the essentially monoenergetic electron beam to propagate through its 2.54-mm radius and 62-mm-long aperture.

The 50-period bifilar helical wiggler produced by current-carrying helical wires has a period of 3.18 cm and provides a magnetic field of uniform amplitude whose magnitude on axis is adjustable up to 1.8 kG. The wiggler field intensity is slowly increased over the initial six periods, and provides an adiabatic input for the electron beam. The system, including the gun, is immersed in a uniform axial guide magnetic field generated by a solenoid. The intensity of this field can be varied up to a maximum of 11.8 kG.

The 2-m-long stainless-steel drift tube has an internal radius of 0.51 cm and acts as a cylindrical waveguide whose fundamental TE_{11} mode has a cutoff frequency of 17.2 GHz. The system is designed to operate in this lowest waveguide mode.

A high-power magnetron operating at 33.39 GHz is the input power source for the FEL amplifier. The wave launcher consists of a short section of circular waveguide of radius 0.31 cm into which ~ 12kW of power is coupled from a standard Ka-band rectangular waveguide. This section of circular waveguide supports only the fundamental TE_{11} mode for the operating frequency. Its radius is then adiabatically increased to the radius of the drift tube. A linearly polarized wave is thereby injected into the interaction region.

---


Chapter 1. Plasma Dynamics

"anechoic chamber." A small fraction of the radiation is then collected by a receiving horn, passes through precision calibrated attenuators, and a 1.7-GHz-wide band-pass filter. The power levels are finally determined from the response of calibrated crystal detectors.

The radiation intensity, frequency and phase are measured as a function of interaction length within the wiggler. This length is varied by changing the distance that the electron beam is allowed to propagate in the drift tube. Application of a strong transverse magnetic field locally is sufficient to deflect the electrons into the waveguide wall and thereby terminate the interaction at that point.

All studies reported in this report were carried out in an FEL subjected to a combined axial magnetic field ($B_z$) and a helical wiggler field ($B_{w}$). The solid dots of figure 2 illustrate the optimal operational parameters in the three well known groups indicated in the figure. Group I refers to the situation where the cyclotron wavelength $l_c = 2\pi v_{c}m/eB_z$ in the guide field exceeds the wiggler wavelength $l_w$ such that $l_c > l_w$. Group II corresponds to the case where the cyclotron wavelength is less than $l_w$, namely $l_c < l_w$; and the Reversed Field Regime is one in which the electron orbital motions in the guide and wiggler fields oppose one another ($-l_c < l_w$). The three operating regimes are separated by two singular, unstable regions. The one, referred to as gyroresonance, occurs where $l_c \approx l_w$; the other so called antiresonance occurs at $-l_c \approx l_w$.

Early studies of the frequency spectrum and chirping have been determined by a heterodyne technique illustrated in figure 1b. A crystal rectifier was used as a mixer for the 33 GHz FEL radiation and for radiation from a variable frequency 23 GHz local oscillator. The resulting beat wave was amplified and sent to an array of band-pass waveguide filters. Thus, the calibrated response of the various channels provides, on a single shot basis, a direct histogram of output power in terms of radiation frequency.

The frequency spectra in the saturation region, corresponding to a measurement with a full length wiggler of $\sim 180$cm are shown in figure 3. We observe that in Group I the spectrum is shifted from the input magnetron frequency by $\sim 100$MHz. The 80 MHz width of each band-pass filter does not allow better resolution. Figure 3 also illustrates that, within the sensitivity of this technique, no fre-

Figure 1. (a) Experimental arrangement of the FEL amplifier, (b) spectrum analyzer, and (c) interferometer.

The output power from the FEL is transmitted by means of a conical horn into a reflection free

---

Figure 2. Orbit calculations of the axial electron velocity as a function of axial magnetic field. The solid points show the operating regimes at maximum measured RF output.

Frequency chirping is observed in either Group II or in the Reversed Field Regime.

In order to improve sensitivity and study frequency chirping (if any) in the two other FEL regimes, we resorted to the method shown in Figure 1c. In this interferometric technique, the total phase difference between the reference and FEL input signals into the hybrid tee as a function of time $t$ is found to be

$$\cos \phi(t) = \frac{P_1(t) - P_2(t)}{2 \sqrt{P_{\text{FEL}}(t)P_{\text{REF}}(t)}}$$  \hspace{1cm} (1)$$

Here $P_{\text{FEL}}(t)$ and $P_{\text{REF}}(t)$ are the instantaneous power inputs into the hybrid tee (see Figure 1c) during the FEL voltage pulse; and $P_1(t)$ and $P_2(t)$ are the corresponding outputs from the tee. For every phase shifter setting $\phi_0$, three shots are taken, and $P_{\text{FEL}}$, $P_{\text{REF}}$, $P_1$ and $P_2$ recorded. The phase shifter is then advanced by 60 degrees, and the procedure is repeated for a total of six successive settings of the phase shifter. A least squares algorithm then fits a sinusoid to the data recorded on the right hand side of Eq. 1 for all six phase shifter settings $\phi_0$ (Figure 4).

The procedure is repeated with the electron beam turned off, and a second sinusoid is generated. The phase slippage between the two sinusoids thereby derived then yields the desired phase change $\Delta \phi(t)$ caused by the FEL interaction.

Figure 5 illustrates the temporal behavior of the phase change $\Delta \phi(t)$ for a discrete set of time values chosen during the FEL pulse. The instantaneous frequency shift can then be obtained from the slope of these curves, namely

$$\Delta f = \frac{1}{2 \pi} \frac{d \Delta \phi(t)}{dt}.$$  \hspace{1cm} (2)$$

The large chirping observed in Group I and the relatively weak chirping in Group II and the Reversed Field Regime confirms the much cruder measurements depicted in Figure 3. We note that Group I is characterized by strong frequency upshifts as well as downshifts, unlike Groups II and Reversed Field regimes where mostly upshifting is seen.

---

Figure 3. Frequency spectra for five successive shots.

The results shown in figure 5 were made for the full 160 cm interaction length $z$. Figure 6 illustrates how the magnitude of chirping varies with interaction length $z$. It is seen to be roughly linear with distance and remains so irrespective of whether the measurements are made in the small signal regime ($z < 120$ cm) or in the nonlinear saturated regime ($z > 120$ cm).

Since the phase $\Delta \phi(t,z)$ has been measured at different times, $t$ during the pulse and at different interaction lengths $z$, we can readily derive the phase shifts as function of $z$ averaged over the duration of the pulse. The results are illustrated in figure 7 for the Reversed Field Regime, together with the corresponding power measurements. The curves represent simulations from the time independent WTDI code.\(^{12}\)

Figure 4. Typical sinusoidal curve fit to determine phase shift at a given time during the RF pulse.

In summary, we have carried out what we believe are the first spatio-temporal measurements of frequency $\Delta f(t,z)$ and phase $\Delta \phi(t,z)$ in a pulsed free electron laser amplifier. Such measurements can provide very sensitive tests of FEL theory. Moreover, they may also prove to be useful in future applications of free electron lasers in situations where very tight control of frequency and phase are demanded.

Erratic behavior of $\Delta f$ may also be used as a signature of unstable FEL operation. To test this notion we have studied chirping at the "antiresonance" (see figure 2). Figure 8 (top) illustrates the lack of reproducibility of $\Delta f$ on a series of successive shots. This is to be compared with the good reproducibility shown in figure 8 (bottom) taken in a FEL parameter range slightly removed from the precise position of the antiresonance. We note that such measurements are not possible at gyroresonance because here the instability is so strong that it drives the electrons into the waveguide walls.

---

Figure 5. FEL phase shift at discrete times during the RF pulse (z = 180 cm).

Figure 6. Frequency chirping as a function of interaction length. In Group I the maximum observed upshift, \( \frac{df}{dz} = 50 \text{ MHz/m} \) and downshift \( \frac{df}{dz} = -32 \text{ MHz/m} \). In Group II the upshift, time averaged over the pulse, \( \frac{df}{dz} = 10 \text{ MHz/m} \). In the Reversed Field Regime the time average upshift \( \frac{df}{dz} = 8 \text{ MHz/m} \).
1.2 Plasma Wave Interactions—RF Heating and Current Generation

1.2.1 Introduction

Sponsors
National Aeronautics and Space Administration
Grant NAGW-2048
National Science Foundation
Grant ECS 88-22475
U.S. Department of Energy
Grant DE-FG02-91-ER-54109

Project Staff
Professor Abraham Bers, Dr. Abhay K. Ram, Dr. Carson C. Chow, Dr. Vladimir Fuchs, Kai P. Chan, Steven D. Schultz, Luigi Vacca

The research of this group is concerned with studies on the electrodynamics of plasmas. Attention is directed toward understanding the nonlinear dynamics of plasmas driven by high-frequency electromagnetic fields (as in RF heating and current drive of magnetically confined plasmas, or in laser-plasma interactions), and the generation and propagation of unstable radiations from laser-plasma interactions and anisotropic electron distributions in space and astrophysical plasmas.

In the following, we report on our continuing study of the synergistic enhancement of current drive efficiency with the use of multiple types of waves. We also report on further progress in understanding web-chaos induced by waves on ion-orbit dynamics that can lead to direct ion heating. The third report concerns further studies of an anomalous electron transport phenomenon we discovered a few years ago: stochastic streamers in electron guiding center dynamics interacting with electrostatic waves. The last report gives our initial results on arriving at a description of transport in intense ion-cyclotron heating of a high-temperature tokamak plasma.
**Chapter 1. Plasma Dynamics**

### 1.2.2 Lower Hybrid Current Drive In the Presence of Waves In the Ion Cyclotron Range of Frequencies

**Sponsor**

U.S. Department of Energy

Grant DE-FG02-91-ER-54109

In the previous RLE Progress Report, we discussed some recent experiments on the Joint European Tokamak (JET) where it was observed that the lower hybrid (LH) current drive (CD) efficiency was significantly enhanced when the lower hybrid current drive (LHCD) waves were launched in the presence of waves in the ion cyclotron range of frequencies (ICRF).

The initial ICRF power is carried from the antenna into the center of the plasma by fast Alfvén waves (FAW). Our previous calculations have shown that the FAW diffusion coefficient is very small compared to the LH diffusion coefficient so that the FAW could not significantly modify the electron distribution function and enhance the CD efficiency. However, the conditions for which the CD efficiency was observed to be enhanced allowed some of the input ICRF power to be mode converted to ion-Bernstein waves (IBW). The mode conversion occurs at ion resonance layers located inside the plasma and the IBWs then propagate away from these resonance layers towards the high magnetic field. The mode conversion is most efficient for small $k_i$'s ($k_i$ is the component of the total wave vector along the toroidal magnetic field). The power mode converted to IBWs depends on the concentration of the minority ion species in a plasma. Mode conversion calculations show that, for JET-type parameters with a hydrogen minority in a deuterium plasma, less than 25 percent of the incident FAW power, depending on the minority concentration $\eta = n_d/n_e$, is mode converted to IBWs:

$$C(\%) = \begin{bmatrix} 0.1 & 2.5 & 12 & 25 \\ \end{bmatrix}$$

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>0.1</th>
<th>0.05</th>
<th>0.03</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(%)</td>
<td>0.1</td>
<td>2.5</td>
<td>12</td>
<td>25</td>
</tr>
</tbody>
</table>

The mode conversion coefficient $C$ decreases with increasing minority concentration $\eta$, with increasing $k_i$, and with increasing minority temperature:

$$C(\%) = \begin{bmatrix} 13.4 & 10.8 & 8.8 & 4.5 \\ 3.6 & 1.7 & 1.0 & 0.3 \\ \end{bmatrix}$$

<table>
<thead>
<tr>
<th>$T_i$ (keV)</th>
<th>4</th>
<th>7</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C(k_i = 1$ m$^{-1}$)</td>
<td>13.4</td>
<td>10.8</td>
<td>8.8</td>
<td>4.5</td>
</tr>
<tr>
<td>$C(k_i = 4$ m$^{-1}$)</td>
<td>3.6</td>
<td>1.7</td>
<td>1.0</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Since low $k_i$'s undergo mode conversion more effectively, a significant portion of the IBW power spectrum is at values of $\nu_i$ that lie above the LH spectrum. The number of electrons resonant with these IBWs is negligibly small at such high parallel (to the total magnetic field) phase velocities of the waves. Thus, the IBWs cannot, initially, interact with the electron distribution function. The IBW spectrum also cannot interact with the ion distribution function as the mode conversion point is sufficiently far away from the ion cyclotron resonance. However, as the IBWs propagate away from the mode conversion region, toroidal effects upshift the $|k_i|$'s so that eventually the IBW spectrum can interact with the electron distribution function and, in particular, with the energetic electron tail pulled out by the LHCDs.

As shown in our previous RLE report, the CD efficiency is determined by solving the Fokker-Planck equation for the flux-surface averaged electron distribution with the wave-electron interaction being described by the appropriate diffusion coefficient. For typical JET-type parameters, we find that:

---


\[
\frac{D_{\text{FAW}}}{D_{\text{LHW}}} = \frac{1}{4} (k_{\text{FAW}}^2 \rho_e^2) \left( \frac{u_\perp}{u_\parallel} \right)^2 \frac{\omega_{\text{LHW}}}{\omega_{\text{ICRF}}} \ll 1 \quad (1)
\]

where \(\rho_e\) is the electron Larmor radius. Equation (1) shows that the FAW diffusion coefficient is very small compared to the LHW diffusion coefficient. As pointed out before, the FAW, due to its small diffusion coefficient, cannot by itself account for the enhancement in the CD efficiency.

The IBW diffusion coefficient does not lend itself to a simple analytical form. We can express the IBW diffusion coefficient\(^{18}\) as \(D_{\text{IBW}} = D_{\text{IBW}}^0 \Delta f\) where \(D_{\text{IBW}}^0\) is referred to as the "form factor." In figure 9, we show \(cD_{\text{IBW}}\) (where \(c\) is the speed of light) as a function of the magnetic field. We find that \(D_{\text{IBW}}^0\) increases with increasing magnetic field and, for a fixed magnetic field, \(D_{\text{IBW}}^0\) decreases with the electron temperature. Thus, IBWs have to propagate away from the mode conversion region before the corresponding diffusion coefficient becomes large. Also, as the electron temperature increases, the IBWs have to propagate even farther in order to enhance the diffusion coefficient. Figure 10 shows the electron distribution function in the presence of LHWs alone and in combination with IBWs. For illustrative purposes, we have assumed that \(D_{\text{IBW}}^0/\Delta f = 20\) and \(D_{\text{LHW}}^0/\Delta f = 10\) (\(u_e\) is the electron thermal velocity, \(D_e = m_e u_e^2\), \(m_e\) is the electron mass, and \(\nu\) is the bulk electron-electron collision frequency). Figure 10 clearly shows that IBWs push electrons to higher momenta than for LHWs alone. In the case of LHWs alone, the calculated current drive efficiency is 0.39 A/W-m\(^2\), while in the presence of IBWs, the current drive efficiency is found to increase to 0.52 A/W-m\(^2\). Thus, IBWs can significantly stretch out the energetic electron distribution function and enhance CD efficiency.

---

**Figure 9.** The IBW form factor (times the speed of light), \(cD_{\text{IBW}}\), as a function of the magnetic field, \(B_0\). The curves are for three different plasma temperatures of 4, 7, and 10 keV. The plasma is a hydrogen minority in a deuterium plasma with \(\eta = 0.03\). The electron density is \(2 \times 10^{13}\) cm\(^{-3}\), \(c\nu/\omega_{\text{ICRF}} = 1.1\), and the ICRF frequency is 48 MHz. The form factor is shown for tail electrons with \(\rho_e = 0\).

---

Figure 10. The parallel (to the magnetic field) electron distribution function versus normalized parallel momentum ($P_n$). ($P_n = 1$ corresponds to an electron energy of 1 MeV.) The plasma is a hydrogen minority in a deuterium plasma with $\eta = 0.03$. The electron density is $2 \times 10^{13}$ cm$^{-3}$, the magnetic field is 3.7 Tesla, the temperature of the electrons and ions is 4 keV, the LHW frequency is 3.7 GHz, and the ICRF frequency is 48 MHz. The small dashed box (with the right inclined lines) is the parallel momentum extent of the LHW spectrum while the two larger dashed boxes (with the left inclined lines) give the parallel momentum extent of the IBW spectrum. The IBW spectrum is symmetric because the launched ICRF spectrum is symmetric in $k_x$. The narrow (in $P_n$) inside curve is the distribution function due to LHWs only; while the broader outside curve is due to the combination of LHWs and IBWs.

In our previous calculations, we have estimated $D_{ibw}^w$ for the IBW in a very approximate way. In order to explain the JET results, we need a good estimate of $D_{ibw}^w$. For this, we have begun to do ray tracing of the IBWs in order to determine the flux-surface averaged diffusion coefficient along the rays.

1.2.3 Fixed Points in Near Resonance Chaotic Webs

Sponsor

National Science Foundation
Grant ECS 88-22475

This is a continuing study on which we first reported in last year’s RLE Progress Report No. 135 (pp. 184-186). Its aim is to arrive at understanding wave induced chaos in ion dynamics in a magnetic field. The electrostatic wave is assumed to propagate across the magnetic field, and its frequency can be either very near an ion-cyclotron harmonic or in-between two ion-cyclotron harmonics. The basic equations were given in last year's report.

---

Here, we report on the motion of first-order fixed points as the frequency is changing from on resonance to off resonance.

We isolate the m-th harmonic term from the Hamiltonian to obtain $H = \tilde{H}_0 + \alpha \tilde{H}_1$, where

$$
\tilde{H}_0(l, \phi) = (\omega_0 m - 1) l + \alpha J_m \left( \sqrt{\frac{2 m l}{\omega_0}} \right) \cos \phi
$$

(1)

and

$$
\tilde{H}_1 = \sum_{n=-\infty}^{\infty} J_n \left( \sqrt{\frac{2 l}{\omega_0}} \right) \cos(n \psi - \tau)
$$

(2)

Corresponding to (1), the equations of motion are:

$$
\frac{dl}{d\tau} = -\frac{\partial \tilde{H}_0}{\partial \phi} = \alpha J_m \left( \sqrt{\frac{2 m l}{\omega_0}} \right) \sin \phi
$$

(3)

$$
\frac{d\phi}{d\tau} = \frac{\partial \tilde{H}_0}{\partial l} = (\omega_0 - \omega_m) \frac{l}{m} \left( \sqrt{\frac{2 m l}{\omega_0}} \right) \cos \phi
$$

The positions of the first order islands and separatrices are given by the elliptic and hyperbolic fixed points, respectively. To solve for the fixed points, we set both $\frac{dl}{d\tau}$ and $\frac{d\phi}{d\tau}$ to zero. After rearranging terms, we find

$$
\sin \phi J_m \left( \sqrt{\frac{2 m l}{\omega_0}} \right) = 0
$$

(4)

$$
\cos \phi J_m' \left( \sqrt{\frac{2 m l}{\omega_0}} \right) = \frac{(1 - \omega_0 m)}{\alpha} \sqrt{\frac{2 \omega_0}{\omega_0}}
$$

(5)

The set of elliptic fixed points are given by $\sin \phi = 0$ and hyperbolic fixed points by

$$
J_m \left( \sqrt{\frac{2 m l}{\omega_0}} \right) = 0.
$$

When we are exactly on resonance, we have $\omega_m = 1$ and the RHS of equation (5) becomes zero. The elliptic fixed points and the hyperbolic fixed points form a checkered pattern, shown in figure 11.

Figure 11. Fixed points with $1/\omega_0 = 5, \alpha = 0.01$. $x$ - hyperbolic fixed points; $\circ$ - elliptic fixed points.
When $1/\omega_0$ moves away from resonance, say $1/\omega_0 = m + \delta$ ($\delta$ small), the elliptic fixed points move in a vertical path, and the hyperbolic fixed points move in a roughly horizontal path. The result of this movement is that pairs of elliptic points and pairs of hyperbolic points group together into clusters of four. As $1/\omega_0$ departs more and more from resonance, the two pairs of fixed points get closer together until they meet and annihilate each other.

Fixed points with large $\tilde{\omega}$ move faster and are annihilated sooner. This is illustrated in figure 12.

The Hamiltonian $\tilde{H}_0(I, \phi)$ does not display chaos. Instead chaos arises from the perturbation $\alpha H_1(I, \psi, \tau)$. Our interest is in the evolution of chaos along the web when $\alpha$, the magnitude of the wave, increases. However, when increasing $\alpha$, we not only change the width of the web, but also change the position of the fixed points. Since the fixed points define the position of the primary islands and separatrices, changing them will also change the separatrices along which the web evolves. Hence it becomes difficult to study the effects of $\alpha H_1(I, \psi, \tau)$ because they are obscured by changes in $H_0(I, \phi)$ (compare figure 13 with figure 14).

To remedy this difficulty, we notice from the RHS of equation (5) that the positions of fixed points depend on the parameter:

$$\eta = \frac{(1-\omega_0 m)}{\alpha} \sqrt{\frac{\omega_0}{m}} = \frac{\delta}{\alpha}$$

Hence when we increase $\alpha$, we can also increase $\delta$ to keep $\eta$ constant and thereby keep the fixed points in place. This will minimize (although not eliminate) the changes in $H_0(I, \phi)$ so that the onset of chaos due to the perturbation will be more evident (compare figure 13 with figure 15). Work along these lines is continuing.

---

**Figure 12.** Fixed points with $1/\omega_0 = 5.004, \alpha = 0.01$. 

---
Figure 13. Particle orbits (I against $\psi$) with $1/\omega_p = 5.02$, $\alpha = 0.01$.

Figure 14. Particle orbits (I against $\psi$) with $1/\omega_p = 5.02$, $\alpha = 0.05$. 

243
1.2.4 Anomalous Electron Guiding Center Streaming Due to Waves

Sponsors

Magnetic Fusion Science Fellowship Program
National Science Foundation
Grant ECS 88-22475
U.S. Department of Energy
Grant DE-FG02-91-ER-54109

The motion of circulating electrons in a tokamak interacting with a discrete spectrum of waves is given by a Hamiltonian formulation in the guiding center limit and studied by numerical integration.

Recent work by Kupfer\textsuperscript{19} on the chaotic electron dynamics induced by electrostatic waves demonstrated that an electron, on a flux surface with rational safety factor $q$, which is in resonance with two waves with identical parallel phase velocity can have a "streaming" motion perpendicular to the magnetic field. Theoretical analysis of this problem is complicated by the necessity of visualizing dynamics in a four-dimensional phase space including both parallel and radial motion.

We use a Hamiltonian description of the electron guiding center motion\textsuperscript{20} that was derived by making use of several previous guiding center theories.\textsuperscript{21}

The model is based on a tokamak with an MHD equilibrium in the low-beta limit, and a low inverse-aspect ratio $\varepsilon = r/R$. The electron is assumed circulating, with an orbit which remains near its initial flux surface. A small magnetic field shear can also be included in this tokamak geometry. The phase space for the guiding center motion is reduced to four dimensions by gyroaveraging and by assuming that the magnetic moment $\mu$ is a constant: this requires that our study be restricted to waves in the low-frequency, long-wavelength limit. The derivation gives two sets of canonical coordinates for the

---


guiding center, \((z_1, p_1)\) and \((z_2, p_2)\). In terms of the familiar guiding center position \((\psi, \theta, \phi)\) and parallel velocity \(v_1\), and omitting correction terms of order \(\varepsilon\), we find:

\[
\begin{align*}
z_1 &= \phi \\
p_1 &= mR_0 v_1 \\
z_2 &= \phi - q_0 \psi \\
p_2 &= e \psi
\end{align*}
\]  

(1)

Here \(q_0\) is the safety factor. The Hamiltonian in the absence of wave perturbations is

\[
H_0 = \frac{1}{2m_e R_0^2} \left[ p_1^2 + \frac{e^2}{\varepsilon} p_1 p_2^2 \right] + \mu B_0.
\]  

(2)

Here we have included shear in the form of \(s_0 = \frac{1}{q_0} \frac{\partial q_0}{\partial \psi}\), which is assumed to be small.

Drifts are accounted for in the order-\(\varepsilon\) corrections of \((z_1, p_1, z_2, p_2)\), which are not given here.

The simplest perturbation to this Hamiltonian is an electrostatic field, in the form of a few discrete plane waves with identical frequency, whose potential is given by:

\[
\Phi = \sum_{n,m} \Phi_{nm} \cos(n \phi + m \theta + k_\psi \psi - \omega t)
\]  

(3)

where \(n\) and \(m\) are mode numbers. For a given frequency, \(k_\psi\) is found using the dispersion relation, e.g., appropriate to lower-hybrid waves. Using the approximate relations given in equation (1), we obtain

\[
\Phi = \sum_{n,m} \Phi_{nm} \cos(k_1 z_1 + k_2 z_2 + k_\psi p_2 - \omega t)
\]  

(4)

with \(k_1 = n + (m/q_0)\) and \(k_2 = -(m/q_0)\). The dynamics of interest occur when the safety factor \(q_0\) is a rational number, which makes the perturbation periodic in \(z_1\) and \(z_2\) and allows us to choose different integers \(n, m\) giving the same value of the parallel mode number \(k_1\) but different values of \(k_2\) and \(k_\psi\).

The electron is in resonance with a wave when \(d(k_1 z_1 + k_2 z_2 + k_\psi p_2 - \omega t)/dt = 0\). The unperturbed Hamiltonian gives \(z_1 = (m_R R_0)^{-1} p_1\), and the other three coordinates constant, so resonance occurs for \(p_1 = (m_R R_0)^{-1} (\omega/k_1)\).

The equations of motion derived from the perturbed Hamiltonian were used to numerically integrate the motion \(z_1(t), p_1(t), z_2(t), p_2(t)\). We select units so that \(m, e, R_0\), and \(\omega\) are all unity. The safety factor is chosen to be \(q_0 = 2\). To reproduce the radial streaming observed by Kupfer, we choose four electrostatic modes with mode numbers \((n, m) = (1,1)(1,2)(2,0)(2,2)\); in the new coordinates \((k_1, k_2) = (3/2, -1/2)(2, -1)(2,0)(3, -1)\). This case is degenerate: two modes have the same parallel mode number \(k_1 = 2\) with different values of \(k_2\) and \(k_\psi\). The resonance condition in simplified units is \(p_1 = 1/k_1\), so that the resonance surfaces in phase space are at \(p_1 = 2/3\), \(1/2\), and \(1/3\). The amplitude of each wave is the same and was chosen so that the separatrices layers for these three waves barely overlap, which creates a stochastic layer around all three of the resonance surfaces in \((z_1, p_1)\) phase space. The initial condition of the coordinates is chosen to lie in this resonance region. The figures describe the motion observed for two cases with different initial conditions in \((z_2, p_2)\). The two points are started on the same flux surface \((p_{z_2} = 0)\) but at different poloidal angles separated by \(\Delta \theta = n/2\) \((z_{20} = 0, \pi)\).

Figure 16 shows the time series \(p_1(t)\) (which corresponds to parallel velocity) for these two cases. In the \(z_{20} = 0\) case, \(p_1\) is observed to fluctuate rapidly throughout the resonance region, spending a roughly equal amount of time near each of the three resonances. However, in the \(z_{20} = \pi\) case, the electron quickly moves into an orbit close to the \(p_1 = 1/2\) resonance, which is degenerate, and stays there. Figure 17 shows the time series \(p_2(t)\) (corresponding to the flux coordinate) in the two cases. In the first case, motion in the \(p_2\) direction is wildly fluctuating, but with occasional periods of directed, non-chaotic motion. Over long times, these periods of streaming add up to a slow drift in the radial direction. In the second case, this streaming motion is nearly continuous, and the rapid fluctuations are no longer visible on this scale, which is more than an order of magnitude larger than on the previous figure.

The electron streaming appears to be related to the patterns of constructive and destructive interference of the two degenerate waves. Let us explicitly add two such waves:

\[
\Phi_0 \cos(k_1 z_1 + k_2 z_2 + k_{\varphi \psi} p_2 - t) + \Phi_0 \cos(k_1 z_1 + k_{2\theta} z_2 + k_{\varphi \psi} p_2 - t) \tag{5}
\]

\[-2\Phi_0 \cos(\Delta k_2 z_2 + \Delta k_{\varphi \psi} p_2) \cos(k_1 z_1 + \bar{k}_2 z_2 + \bar{k}_{\varphi \psi} p_2 - t),\]

where \(\bar{k}_2 = (k_{2\theta} + k_{2\phi})/2\), \(\Delta k_2 = (k_{2\theta} - k_{2\phi})/2\), \(\bar{k}_{\varphi \psi} = (k_{\varphi \psi} + k_{\varphi \psi})/2\), and \(\Delta k_{\varphi \psi} = (k_{\varphi \psi} - k_{\varphi \psi})/2\). It is easy to show with equation (5) that \(z_{20} = \pi\), \(p_{20} = 0\) is a point where the degenerate waves interfere destructively. For an electron at this point, the \(k_i = 2\) waves appear to vanish, and the regions of phase space affected by the remaining waves no longer overlap.

We would like to see how this interference pattern affects the phase space of the guiding center motion. Unfortunately, a surface of section in this phase space is given by a four-dimensional mapping, which is impossible to visualize. But if a canonical transformation can be found so that two of the four phase space variables are nearly constants of the motion, a plot of this mapping in the phase plane of the other two coordinates is an approximate surface of section. The details of the canonical transformation giving this approximate surface of section are under investigation. This is expected to reveal that there are quantities which are very nearly conserved by the streaming motion.

We are currently working on a similar phase-space analysis of electron guiding center motion in electromagnetic waves, for example fast Alfvén waves. This problem is complicated by the fact that the canonical phase-space coordinates are shifted when a magnetic potential is added to this analysis.
1.2.5 Resonant Fast Alfvén Wave-Particle Interaction in a Tokamak

Sponsor
U.S. Department of Energy
Grant DE-FG02-91-ER-54109

Ion heating in a tokamak by fast Alfvén wave fields has been the subject of numerous investigations both analytical and numerical. In particular, these studies have led to a description of the ion distribution function in the presence of such fields by Fokker-Planck equations in velocity space. However, such Fokker-Planck equations do not include the induced transport due to the fields. For intense heating of tokamak plasmas, it is important to evaluate the spatial transport that may accompany such heating. We have initiated such a study.

Several simplifying assumptions are usually made in order to describe the field-induced diffusion in velocity space. The most serious ones are the neglect of finite banana widths and spatial diffusion induced directly by the wave. Changes in parallel velocity due to Doppler-shifted cyclotron resonances are responsible for spatial transport due to fields. This is immediately seen for trapped particles once one considers the change in toroidal angular momentum due to parallel velocity changes. This effect is mostly relevant at the magnetic axis of a tokamak. In order to study these effects one has to estimate the changes in space and velocity that an ion undergoes through resonance.

We have studied resonance interactions using Hamiltonian formalism in a perturbative way. We assume a small wave amplitude with respect to the equilibrium field; this allows us to write a Hamiltonian of the following form:

\[ H = H_0 + \lambda H_1 \]  

where \( \lambda \) is a parameter that accounts for the smallness of the wave amplitude. The leading order Hamiltonian is the usual guiding center Hamiltonian:

\[ H_0 = \frac{1}{2} m u^2 + \mu B \]  

The wave-particle resonance is contained in \( H_1 \). The perturbed Hamiltonian is Fourier-decomposed as follows:

\[ H_1 = \sum_k H_{1k} e^{i\psi(k)} \]  

where

\[ \psi_1(k) = N_k + M \theta - k_r r - \omega t + \phi + \alpha + \pi/2 \]  

---


Chapter 1. Plasma Dynamics

and

\[ H_{1k} = \frac{iq}{2\omega} \sqrt{2B/m} \left[ J_0 e^{-i\alpha} E_{+k} + J_2 e^{i\alpha} E_{-k} \right] \]  

The equations of motion due only to wave interaction in a tokamak with circular flux surfaces, where the unit vector in the direction of the magnetic field is given by \( b(r) = f(r)e_x + g(r)e_y \), are:

\[ \dot{\chi}_{\alpha 1} = \sum_k \frac{1}{\mu B} b \times \nabla \psi_1(k) H_{1k} e^{i\psi(k)} \]  

\[ \dot{\mu}_1 = -\sum_k \frac{1}{m} \frac{q}{\mu} H_{1k} e^{i\psi(k)} \]  

\[ \dot{\phi}_{\alpha 1} = \sum_k \frac{q}{m} \frac{\partial H_{1k}}{\partial \mu} e^{i\psi(k)} \]  

\[ \dot{v}_{\alpha 1} = -\sum_k \frac{B_r}{mB_r} \cdot \nabla \psi_1(k) H_{1k} e^{i\psi(k)} \]  

The wave-particle interaction is localized in space so that the change in the guiding center variables can be evaluated by integrating equations (6) - (9) in the vicinity of the resonance region. Outside this region, the particle is essentially unperturbed by the wave. The wave-particle resonance condition is given by: \( d\psi_1(k)/dt = 0 \). Thus, the change in guiding center variables is:

\[ \Delta A = \int_{\text{res}} \sum_k \hat{A}_{1k} e^{i\psi(k)} dt \]  

where \( A \) represents any guiding center phase variable. Using the stationary phase approximation:

\[ \Delta A = \sum_k \hat{A}_{1k} |_{\text{res}} \int_{\text{res}} e^{i\psi(k)} dt \]  

We now need to solve for the phase integral in the resonance region.

We introduce the resonance integral defined as

\[ I = \int_{\text{res}} e^{i\psi} = \sum_k \int_{\text{res}} e^{i\psi(k)} \]  

\[ I = \sum_k \int_{\text{res}} e^{i\psi(k)} \]  

\[ I = \frac{e^{i\psi} - e^{-i\psi}}{2} \]  

where this integral has to be evaluated along the unperturbed orbit in the resonance region. Here \( \alpha \) is held to be a constant and \( f = \mu B \), \( \omega_m = \omega_{lb} \), where \( \mu = R_0/\omega_{lb} \).

Resonance integrals for passing particles were evaluated in two simple cases: without Doppler-shift and with Doppler-shift for an asymmetric monochromatic wave. For a resonance without Doppler-shift the resonant point is given by the coordinates that satisfy \( \omega = qB/m \).

The resonant integral at leading order is easily evaluated:

\[ I = \sqrt{\frac{2\pi}{\omega \omega_{lb}}} \left( \frac{e^{i\psi}}{\sqrt{\mu \omega_{lb} \sin \theta}} \right) e^{i\pm \pi/4} \]  

When we include the Doppler-shift for a monochromatic wave we obtain for the resonance condition:

\[ \frac{d\psi}{d\theta} \]  

\[ (\omega_{lb} - \omega_{cb}) \frac{1}{u_{\|} R_0} \frac{a}{g} x - M - N f \frac{a}{gh} R_0 x = 0 \]  

where \( h = 1 + \frac{r}{R_0} \cos \theta \) and \( x = r/a \). The resonance integral can be evaluated asymptotically:

\[ \frac{1}{t} = \left[ \frac{a}{R_0} x \frac{g}{u_{\|}} - \frac{1}{u_{\|} R_0} \right] e^{i\psi} \int_{\text{res}} \sqrt{\frac{2\pi}{\mu \omega_{lb}} \sin \theta} \]  

where

\[ -x^2 \sin \theta \left( \frac{a}{R_0} \right)^2 \left[ \frac{\omega_{cb} f g}{u_{\|}} + \frac{\omega_{cb}}{h g f^2 / \omega_{lb}^2 + 2 u_{\|}^3 N f / g} \right] \]  

Future work will include the evaluation of resonance integrals for trapped particles and the extension to multichromatic spectra.
1.2.6 Publications


1.3 Physics of Thermonuclear Plasmas

Sponsor

U.S. Department of Energy
Contract DE-FGO2-91ER-54109

Project Staff

Professor Bruno Coppi, Neer R. Asherie, Franco Carpiognano, Dr. Paolo Detragiache, David H. Hijirida, Dr. Stefano Migliuolo, Dr. Marco Nassi, Hana Ohkawa, Gregory E. Penn, Caterina Riconda, Todd H. Rider, Dr. Barrett Rogers, Dr. Linda E. Sugiyama, George M. Svolos

As our primary activity in this research program, we study the theory of magnetically confined plasmas in regimes relevant to present-day advanced experiments and to future thermonuclear devices. In the future, these devices will probably ignite the plasma ("thermonuclear fuel") within toroidal magnetic confinement configurations and will involve either first generation fuels, a deuterium-tritium mixture (Ignitor, ITER), or more advanced fuels such as deuterium-deuterium or deuterium-helium mixtures (Candor).

The Ignitor-Ult machine is now in the early stages of construction in Europe. At MIT, the Alcator C-MOD experiment has recently begun operation. Alcator C-MOD combines the favorable features of an elongated plasma cross section with a high magnetic field to produce high plasma currents. This machine is in fact very similar to Megator, an experiment we proposed in the early 1970s as a logical continuation of the Alcator program.

Currently, our research program follows two major avenues. First, we are studying the basic physical
Chapter 1. Plasma Dynamics

processes of thermonuclear plasmas (equilibrium, stability, transport, etc.) as they apply to existing or near-term future systems. In this effort, we closely collaborate with our experimental colleagues and with theorists from other research groups (e.g., Columbia University, JET, Phillips Laboratory, Princeton University, University of Texas, Lawrence Livermore National Laboratories). This work also involves time-dependent simulations of plasma discharges in the Ignitor-Ult experiment. We focus particular attention on the evolution of spatial profiles of plasma current and temperature. Collaboration with our colleagues at the Ente per le Nuove tecnologie l'Energia e l' Ambient (E.N.E.A.) in Italy, as well as inhouse code development, plays a major role in this endeavor.

Second, we explore advanced regimes of thermonuclear burning, including those employing low neutron yield fuels (D-3He, and "catalyzed" D-D). We consider both the design of machines that will contain these very high temperature plasmas as well as the physics that govern their behavior. Below, we discuss some of the salient results of work completed or presently being carried out by members of our research group.

1.3.1 Hyper-Cyclotron Emission (HCE) Induced by Fusion Reaction Products

We have proposed a theoretical interpretation and developed the relevant analysis for the radiation emission above the cyclotron frequency of the fusion products that has been observed in well confined plasma experiments with a significant population of fusion reaction products. The emission spectrum typically shows distinct peaks at multiples (p=1 to 7 in the case of the Deuterium-Tritium plasma experiments carried out by the JET machine) of the alpha particle cyclotron frequency \( \Omega_\alpha \). The frequency values suggest that the emission is coming from a localized region at the outer edge of the plasma column. For \( p \geq 8 \), the spectrum becomes continuous, while below \( \omega = \Omega_\alpha \) some smaller peaks are observed. We have found that by considering the interaction of radially localized compressional Alfvén modes with anisotropic (in velocity space) alpha particle distributions, we can explain the key features of the discrete part of the spectrum. The continuum component of the spectrum is related to the presence of magnetic shear and to the transition from compressional Alfvén modes to whistler modes that occurs as the ratio \( \omega/\Omega_\alpha \) increases. It is the overlapping of a class of these quasi-whistler modes that is proposed to produce the continuum.

We have investigated the radial structure of the Alfvén mode in detail. We find it to be localized in the plasma at a radius which agrees with the experiment results (for the discrete component of the spectrum) obtained by the JET machine. Since we are assuming the plasma to be cold, we use the MHD equations with the addition of the Hall term in Ohm's law to study the mode structure in this frequency range. This work has been performed in a cylindrical geometry, and will be extended to the toroidal case in order to fully understand the poloidal structure of the mode.

The calculation of the mode structure has been carried out assuming that plasma is composed of deuterium and tritium. This brings out the essential features of the problem, but neglects the effects of impurity ions. As the impurities are typically much heavier than the bulk plasma ions, they introduce resonances for \( \omega < \Omega_\alpha \) (the cyclotron frequency of deuterium = \( \Omega_\alpha \)) which may correspond to the smaller peaks measured experimentally. In addition, they introduce a significant correction to the Alfvén speed.

We have explored some of the aspects related to the linear growth rate, \( \gamma_\alpha \), of the HCE instability. In particular, it is the perpendicular velocity gradient, \( \delta v_p/\delta v_\alpha \), that determines whether the mode is unstable or not. Throughout this work it has been assumed that the background plasma is cold. Although this is a good assumption for the electrons, complications arise for the ions. Energy is transferred from the alpha particles to the mode

---

which then transfers its energy to the ions. Eventually the ions have enough energy so that thermal corrections (e.g., resonance broadening) become important. This may affect the growth rate in a significant way.\(^3\)

1.3.2 Stabilization of External MHD Instabilities

An obstacle to reaching the second stability region\(^3\) in toroidal magnetic confinement experiments is the possible onset of ideal MHD external modes driven predominantly by the current density gradient near the edge of the plasma column. This region in parameter space is one where the plasma is stable against pressure gradient driven internal MHD instabilities, while having a relatively high value of \(\beta_\perp = 8 \pi \ell \rho B / B \perp\). A recent series of analyses\(^32\) have shown that these regimes correspond to equilibria with plasma currents that are peaked toward the outer edge of the torus and can excite macroscopic external modes.

By considering a simple cylindrical equilibrium with magnetic field \(B(r) = B_\perp(r) + B_\parallel(r)\) and \(\beta = 8 \pi \ell \rho B / B \perp \) we begin to explore the possible self-healing of the plasma by diamagnetic effects. Since the external modes tend to have toroidal mode number, \(m = 1\), and poloidal mode number, \(m = q_\perp\) (where \(q_\perp\) is the value of the magnetic winding index, \(q = \ell B / B \perp\) at the plasma edge) and proposed second-stability experiments will operate with high values of \(q_\perp\), there exists the possibility that the diamagnetic frequency of the plasma, \(\omega_D = - m (c/\ell n B) \ell \rho / \ell r\), may not be negligible when compared to the growth rate, \(\gamma\), of the ideal mode. When that is the case, the so-called FLR stabilization\(^33\) by \(\omega_D\) can come into play.

The derivation of the dispersion relation, the relevant boundary conditions and related work can be found in Coppi et al.\(^34\) Considering the simple case of a plasma surrounded by vacuum, the following quadratic form can be constructed from the dispersion equation:

\[
\omega^2 - \omega \langle \omega_\perp \rangle - \langle \omega_D^2 \rangle = 0
\]

where to leading order in \(1/m\):

\[
\langle \omega_\perp \rangle = \int_0^a dr \frac{\langle \ell \rho \ell n \ell B \rangle^2}{\int_0^a dr \ell \rho \ell n \ell B^2}
\]

\[
\langle \omega_D^2 \rangle = \frac{- a^2 \ell \rho \ell n \ell B^2 + \int_0^a dr \ell r \ell n \ell B^2 / 4\pi}{\int_0^a dr \ell \rho \ell n \ell B^2}
\]

\[
\Sigma = \frac{B_\perp^2}{4\pi m r^2 q_\perp^2} \quad (m - q_\perp)
\]

\[
F = k \cdot B \quad \text{and} \quad \Lambda^2 = \frac{1}{m^2} |\ell \rho |^2 + |\ell n \ell B|^2.
\]

The only term that can produce a growth rate is the boundary term \(\Sigma\). In particular, we see that \(\Sigma > 0\), or equivalently \(0 < m - q_\perp < 1 + q_\perp / m\), is a necessary condition for instability. Conversely, one can show from equation 1 that a sufficient condition for stability is \(\min(|\omega_\perp|) \geq 2 \gamma_0\), where \(\gamma_0\) is the growth rate for \(\omega_\perp = 0\).

---


In the simplest case of approximately constant $\omega_n$, it is possible to see that $\omega_n$ can stabilize the ideal MHD mode by inspection of equation 1. In particular, the dispersion relation $\omega(\omega - \omega_n) = \omega_0^2$ (where $\omega_0$ is the eigenvalue obtained with $\omega_n = 0$) gives

$$\omega = \frac{1}{2} \left( \omega_n \pm \sqrt{\omega_n^2 + 4 \omega_0^2} \right)$$

with the well-known 33 stability condition:

$$\omega_n^2 = -4 \omega_0^2 = 4 \gamma_0^2.$$  

Note that once $\omega_0$ is known, e.g., from a numerical stability analysis 35 of the ideal MHD problem, diamagnetic stabilization can be investigated directly (i.e., without solving the boundary value problem anew) by just evaluating $\omega_n$.

We further specialize our cylindrical analysis to regimes where the ratio of edge to average value of the plasma current, $J_e/J$, is small and the plasma density (and pressure) do not vanish at the edge. That restricts us to relatively weakly growing ideal modes and benefits the $\omega_n$ effect. The relevant dispersion relation, obtained 36 by solving a boundary value problem, is:

$$\omega(\omega - \omega_n) = 3 \zeta_0 (m - n_q a)^2 \omega_A^2 \Delta$$

where $\omega_A = B_a(a)/(\sqrt{4\pi \rho_e a})$ is the relevant Alfvén transit frequency,

$$\Delta = 1 - \frac{J_a}{(m - n_q a) \langle J \rangle},$$

$$\zeta_0 = \frac{(m - n_q a) \langle J \rangle}{\langle J \rangle} - J_a.$$ 

It is evident that there is a limit to the ratio $J_e/J$ for which stability can be attained, but this limit will have to be estimated only after realistic growth rates are obtained from toroidal geometry calculations.

If a perfectly conducting wall is placed at $r = b > a$ surrounding the plasma, the quantity $\Delta$ in the dispersion relation (8) is replaced by $\Delta + \Delta_w$, where $\Delta_w = 1/[((b/a)^{2m} - 1)]$ is positive. Clearly, this can have an important effect in reducing $\gamma_0$ to values below the indicated stability threshold.

When finite electrical resistivity of the wall is considered, the two oscillatory modes that arise when the diamagnetic frequency is sufficient to overcome the ideal MHD instability drive (see equation 6), become damped, but a new purely growing mode tends to develop with growth time that is proportional to the resistive diffusion time in the wall. This mode appears to be "immune" to FLR stabilization.

1.3.3 Stabilization of Internal Kinks by Circulating Particles

This work is being done in collaboration with F. Porcelli with the JET Joint Undertaking, Abingdon, U.K.

The topic of linear stabilization of internal kinks by energetic particles has received considerable interest since the observation 36 of sawteeth can be prevented in tokamaks by the application of radiofrequency heating at the cyclotron frequency of a minority population of ions. The explanation 37 of this phenomenon resides in the dynamics of the minority ions which are accelerated to high energies (several MeV in the case of the JET experiments): the ions acquire most of their energy when the turning points in their orbits coincide with the plane of resonance with the wave, $\Omega = \omega$ (here $\omega$ is the frequency of the applied RF, while $\Omega(r, \theta)$ is the spatially varying cyclotron frequency of the ions). Thus, ion cyclotron resonance heating (ICRH) preferentially energizes ions that are trapped in the bad curvature region of the tokamak by the magnetic well. When these ions are sufficiently energetic, namely, they have a precession frequency (due to grad-B and magnetic curvature) that is larger than the ideal MHD growth rate of an internal kink, they tend to react in a special manner to the perturbation. Instead of $E \times B$ drifting with the rest of the plasma, they react so as to conserve the magnetic

---


flux contained within the plane defined by the precession of each individual guiding center. An increase in the potential energy of the system ensues, provided certain conditions (e.g., centrally peaked pressure profile) are met.

Here, we dwell anew on the interplay between energetic particles and \( m=1 \) internal kinks, concentrating on the circulating species. As is rather obvious, passing particles tend to spend a larger fraction of their time in the region of good curvature (\( \theta = \pi \) in tokamaks with circular cross-section) and thus their streaming along the field lines tends to negate the radial displacement imposed by the \( m=1 \) perturbation. Thus, this population tends to provide a potential energy "hill" that the mode must overcome in order to become unstable. As we show below, the quantitative effect can be of the same order as that provided by the trapped ions discussed in the more standard descriptions.

We remember the dispersion relation for the \( m=1 \) internal kink in the ideal limit:

\[
\omega(\omega - \omega_A) + \omega_A^2(\lambda_A + \lambda_k)^2 = 0
\]

where \( \omega_A = v_A/\sqrt{3} R \) is the characteristic Alfvén frequency, \( \omega_A = (c/enB) (dp/dr) \) is the ion diamagnetic frequency (computed at \( r = r_i \), the position where the magnetic winding index \( q(r) = rB_\theta / B_\phi \) is equal to unity; here \( \zeta \) and \( \theta \) are the toroidal and poloidal coordinates). The quantity \( \lambda_A \) is a measure of the free energy available from the plasma pressure gradient (it is proportional\(^{39} \) to the square of \( \beta_p = (8\pi / B_\phi) \langle \rho(r) \rangle \) where \( \langle \rho(r) \rangle \) is the mean pressure within \( r \leq r_i \) and must be sufficiently positive for an ideal mode to become unstable in the presence of diamagnetic effects). The term denoted \( \lambda_k \) is the contribution from the energetic particles and is the subject of this discussion. We refer the reader to, e.g., Coppi et al.,\(^{40} \) for a derivation of this term. Here, we limit ourselves to stating that it is composed of two parts: one, which we term "adiabatic" to which all particles (trapped and circulating) contribute, the other, which we call "non-adiabatic" is due to trapped particles only (under the assumption that the mode frequency is much smaller than the transit frequency of the energetic passing ions). Focus on the contribution from the "adiabatic" portion

\[
\lambda_k^{ad} = - \frac{4\pi^2}{R(rqB_\theta r_i^2)} \left( \frac{2}{m n} \right)^{3/2} \int_0^{r_i} dr \int_0^\pi d\theta \cos \theta \int_0^{\varepsilon_{ih}} d\varepsilon \int_0^{2\pi} d\alpha \frac{dE}{\sqrt{1 - \alpha}} \times (2 - \alpha) \frac{\partial F_h}{\partial r}
\]

where \( \varepsilon_{ih} \) is the maximum energy of this population. When the distribution function of the energetic particles, \( F_h \), is isotropic (function of energy but not magnetic moment, \( \mu = m v^2 / B = \alpha v^2 \sqrt{B(r, \theta)} \)) this term vanishes identically (the \( \cos \theta \) factor averages out). Thus, it was found\(^{41} \) that fusion alpha particles would contribute only via their trapped fraction to the linear stability of internal kinks and would do so through the "non-adiabatic" part of \( \lambda_k \). As has been indicated in several works (see, e.g., the review articles\(^{42} \) and references therein) this term is stabilizing (\( \lambda_k^{ad} < 0 \)) for peaked pressure profiles (\( dp/dr < 0 \)).

Consider now the case of a population consisting entirely of deeply trapped ions. It can be shown\(^{43} \) that the stabilizing contribution \( \lambda_k \) is now decreased by a factor \( 1 - q_0 \) (where \( q_0 \) is the value of \( q(r) \) at the magnetic axis), all other factors being

---

equal (i.e., same pressure gradient). Thus, the inference is that the "adiabatic" portion of \( \lambda_k \) provided by trapped particles is positive (destabilizing). As discussed earlier,\(^{44} \) a trapped particle lives in the bad curvature region, on the average, and will contribute in a destabilizing manner from its fluid-like ("adiabatic") part of the response.

It then stands to reason that the average circulating particle contributes a stabilization to the system by counteracting the fluid-like term due to the trapped particles (i.e., summing circulating and trapped contributions leads to an overall cancellation of the adiabatic piece). Thus, the circulating particles are termed\(^{44} \) to play a "sacrificial role."

To put this in a more quantitative context, consider the adiabatic contribution anew, writing it as:

\[
\lambda_k^{ad} = -\frac{8\pi \sqrt{1}}{(qR_0 \theta_1) \sqrt{r_1}} \int_0^{r_1} \left( \frac{r}{R} \right)^{3/2} \left( \int_0^\xi d\xi \right)^{3/2} (G_T + G_C)
\]

where \( G_T \) and \( G_C \) are contributions from the trapped and circulating particles, respectively:

\[
G_T = \int_0^1 d\kappa \left[ 2E(\kappa) - K(\kappa) \right]
\]

\[
\left( \frac{2}{m_h} \right)^{3/2} \frac{\partial F}{\partial r}
\]

\[
G_C = \int_1^{R/R} d\kappa \left\{ \frac{1}{\sqrt{\kappa}} \left[ 2\kappa(1/\kappa) - K(1/\kappa) \right] + K(1/\kappa) \right\}
\]

\[
\left( \frac{2}{m_h} \right)^{3/2} \frac{\partial F}{\partial r}.
\]

Here \( \kappa \) is a pitch angle variable (the range \( 0 < \kappa < 1 \) corresponds to trapped particles), and \( K \) and \( E \) the first two complete elliptic integrals. The terms in curly brackets are simply the orbit average of \( \cos \theta \) (between turning points \( \pm \theta_0 \) for trapped particles and between \( \pi \) for circulating particles). They are the same factors found in the orbit average of the geodesic portion of the precession frequency, \( \omega_0 \). The dependence of these terms on pitch angle can be seen, e.g., in figure 9 of Coppi and Rewoldt,\(^{45} \) (note that our terms correspond to \( 2K(\kappa)G_T(\kappa) \) in their notation). It is obvious from this figure that circulating particles give a stabilizing contribution (\( \lambda_k^{ad} < 0 \)) for centrally peaked pressure profiles. Note that the contribution is maximal for particles that are near the trapping boundary (\( \kappa = 1 \)) and that barely trapped particles on the other side of the boundary also give a stabilizing contribution via their adiabatic term. The mechanism is the same as for the circulating particles: their orbits have turning points that are near the inner midplane of the torus (\( \theta \sim \pi \)) and consequently they spend the greater part of their time in the region of good curvature. However, these barely trapped particles (so-called "sloshing" particles\(^{46} \) will not actually stabilize the mode: unlike the circulating particles (whose entire contribution to the problem is contained within \( \lambda_k^{ad} \)), their response involves the kinetic term, \( \lambda_k^{ad} \). It turns out that these barely trapped particles are also the ones for which drift reversal has occurred, namely \( \omega_{dr} < 0 \) (here \( \omega_{dr} \) is the bounce-averaged precession frequency). The drift reversal point is a function of the magnetic shear at the \( q=1 \) surface and located between the point where \( \omega_{dr} = 0 \) and the trapped-circulating boundary (\( \kappa = 1 \)) for positive shear (cf. figure 9 of Coppi and Rewoldt).\(^{45} \) Thus, a cancellation between the fluid-like and kinetic contributions to \( \lambda_k \) from these sloshing particles tends to occur. Note also that particles that sit in the very near proximity of the trapped-circulating boundary are subject to scattering (by collisions and/or fluctuations).\(^{47} \) Thus, most of the stabilization will come from circulating particles located in the region \( 1 < \kappa < 2 \).

We conclude by noting that the contribution from energetic circulating particles to the problem of \( m=1 \)


mode linear stability can be schematically written as:

\[ \lambda_n = - \left( \frac{r_n}{R} \right)^{3/2} \frac{\beta_{ph}}{(\rho_q)^2} C \]  

(15)

where \( \beta_{ph} \) is the hot particle poloidal beta defined in the same manner as for isotropic distributions (cf. equation (IV-33) of Migliuolo\(^{48} \) ) and \( C \) represents the pitch angle contribution, which is of order unity for "slowly passing" particles. This contribution is of the same order as that due to the trapped portion of an isotropic distribution of energetic particles. One is therefore led to propose that sufficiently energetic neutral beams, injected tangentially to the magnetic field, may be effective in producing sawtooth-free discharges in tokamaks.

1.3.4 Deuterium-Helium(3) Fusion Burning Using ICRH Minority Heating in Ignitor

The design characteristics of theIgnitor experiment\(^{49} \) allow it to satisfy the conditions where the 14.7 MeV protons and the 3.6 MeV \( \alpha \)-particles produced by the D-\(^{3}\)He reactions can supply thermal energy to a well confined plasma. In particular, Ignitor:

- can sustain a plasma current exceeding that required \( I_p \geq 6 \text{ MA} \) to confine the proton orbits at birth;
- can produce sufficiently high plasma densities that the slowing down time of both the protons and the \( \alpha \)-particles is shorter than the electron energy replacement time of the thermal plasma in which they are generated.

In order to boost the D-\(^{3}\)He fusion reaction rate, ion-cyclotron heating of a D plasma with a \(^{3}\)He minority can be used to create a tail in velocity space for the \(^{3}\)He distribution function\(^{50} \) with an optimal mean energy value. We note that the D-\(^{3}\)He fusion cross section has a maximum near \( E_\text{c} \approx 850 \text{ keV} \) for a helium beam incident on a deuterium target.

A preliminary analysis of the fusion power \( P_f \) that may be produced in Ignitor by this method indicates that \( P_f = 1 \text{ MW} \) may be reached, thanks primarily to the high value of the RF power density that can be coupled to the helium nuclei. We recall that in previous experiments, fusion power levels of 1.5 kW and 140 kW have been obtained in PLT\(^{51} \) and in JET\(^{52} \) respectively. However, neither of these experiments had the current and particle densities required for the confinement and the slowing-down of the 14.7 MeV protons.

Numerical simulations for this scenario in Ignitor were performed using the FPPRF\(^{53} \) code, that combines a ray tracing package with the solution of the Fokker-Planck equation for the minority species. Typical RF parameters were \( P_{\text{RF}} = 18 \text{ MW}, \nu = 132 \text{ MHz}, \langle k_{\|} \rangle = 5 \times 10^7 \text{ m}^{-1} \). The parallel wavelength \( \langle k_{\|} \rangle \) is estimated on the basis of the size of the adopted ICRH antennas. The ray tracing routine (SPRUCE) solves the dispersion relation for electromagnetic waves propagating in a hot plasma, while another set of routines evolves the distribution function of the minority ions in time by solving the bounce-averaged Fokker-Planck equation:

\[ \frac{\partial f}{\partial t} = \langle Q \rangle + \langle C \rangle + \langle S \rangle \]  

(16)

where \( \langle Q \rangle \) represents the power deposition by the wave and is a quasilinear operator which computes the (spatially varying) resonant wave-particle energy exchange, \( \langle C \rangle \) is the standard collision operator involving pitch angle scattering, slowing down (by electrons and background ions) and energy exchange with other species (electrons and ions), and \( \langle S \rangle \) represents the combination of sources and sinks (charge exchange, prompt losses of fast ions, etc.).

---


Page 255
The code returns wave dispersion properties (in particular $N(\omega)$, where $\bar{N} = \omega_0 / \omega$ is the index of refraction), the RF power deposition profile as a function of $r$ and $\theta$, the minority ion distribution function $f(r, E, \mu)$, the mean energy of the RF heated minority as a function of $r$, as well as an estimate of the beam-target fusion reaction yield (i.e., number of protons per second, for the case of D-$^3$He where the "beam" is composed of $^3$He ions and the "target" of D ions). The code is able to treat the interaction with the minority species at either the fundamental or the first harmonic of the cyclotron frequency. 

In the minority heating scenario for the fast wave, four characteristic distances can be identified that are important for the damping of the wave. In terms of $x = R - R_{mag}$ (where $R_{mag} = R_0$ is the distance of the magnetic axis from the symmetry axis), these are:

1. $x_c = 0$, the distance of the cyclotron resonance $\omega = \omega_0$ of the helium ions (note that the wave field at this location has mostly a right-handed polarization which is opposite to the one required for optimal coupling with the ion gyromotion);

2. $\Delta x_c = k_2 V_{th}\omega / \Omega_{He}$, the width of the thermal broadening of the cyclotron resonance;

3. $x_l = -(1/2)(n_{He}/n) R_{mag}$, the distance of $N^f - L = \Sigma (\omega_0^f / \Omega_0^f)^2 \Omega_0^f / (\Omega_0^f - \omega)$ cut-off surface $k^2 = 0$ (the other cut-off surface, $N^f = R = \Sigma (\omega_0^f / \Omega_0^f)^2 \Omega_0^f / (\Omega_0^f + \omega)$, is located near the plasma periphery). In the cold plasma approximation, at $x_c$ the wave has a complete left-hand polarization;

4. $x_S = -(7/12)(n_{He}/n) R_{mag}$, the distance of $N^f - S = (R + L)/2$ ion hybrid resonance surface $k^2 \to \infty$. When the effects of finite plasma temperature are taken into account, the (electromagnetic) fast wave can couple to the (electrostatic, short wavelength) ion Bernstein wave (mode conversion) near the ion hybrid surface leading to heating of the bulk plasma.

When ICRH is applied in the minority heating regime it is important that:

1. the evanescent region separating the cutoff and ion hybrid resonance surfaces be as large as possible in order to minimize direct heating of the bulk plasma via mode conversion (we find that the addition of a small concentration of a third ion species—hydrogen or tritium—is helpful in this respect);

2. the width $|\Delta x_l|$ be broad enough to overlap the cutoff surface where optimal conditions of wave field polarization for coupling with helium gyromotion are encountered.

As the helium is heated from the initial plasma temperature to an effective tail temperature of 1 MeV (or more), the width $|\Delta x_l|$ is seen to increase from 0.01 m to 0.1 m, finally overlapping the cutoff region. Accordingly the peak in $\text{Im}(N_f)$, associated with the damping of the wave, moves from its original position at the cyclotron resonance toward the $N_f^f - L$ cutoff surface.

The helium distribution function is obtained, at several times during the simulation, and the results then written in practical units ($T = T/10$ keV, $n = n / 10^{20}$ m$^{-3}$, $\ln \Lambda = \ln \Lambda / 10$). Generally, over 95 percent of the RF power is deposited (promptly) into the helium. A substantial fraction of this power (up to 50 percent) is then transferred to the electrons by collisional processes. The fusion power output is estimated with several simplifying assumptions (e.g., an isotropic velocity distribution function for the fast ions) and the quoted value $P_f = 1$ MW should be considered as preliminary.

Alternatively, we have also computed the D-$^3$He reactivity parameter $<\nu_{f3}>$ for a $^3$He distribution function as given by Stix. The relevant two-dimensional velocity space integral is computed

\[ \xi_{RF} = \frac{\gamma^{1/2}_{\nu}}{\hat{n}_0 \hat{n}_{He} \ln \Lambda} \left( \frac{\langle \rho_{RF} \rangle}{\text{MW/m}^3} \right) \]
numerically. For a given set of temperatures, there is an optimal value of \( \xi_{\text{ref}} \) for which \( \langle \sigma v \rangle_F \) is maximum, and this value can be lowered considerably by increasing \( T_e \) and \( T_i \). Since \( \xi_{\text{ref}} \) is inversely proportional to the square of the electron density (for constant minority concentration), while \( P_F \propto n_{\text{eff}} \langle \sigma v \rangle_{\text{ref}} \), an increase in the bulk plasma temperature allows a higher optimal density at constant values of \( \rho_{\text{ref}} \) and therefore of \( P_F \).

A zero-dimensional (no quantitative radial profile effect included) estimate of the total fusion power can be obtained from \( \langle \sigma v \rangle_F \) in the following way. On the basis of the inferred value of \( \langle \rho_{\text{ref}} \rangle \), we can select, for given temperatures, an optimal density product \( n_{\text{He}} \rho_{\text{He}} \). This is determined in such a way that the corresponding value of \( \xi_{\text{ref}} \) is the one for which \( \langle \sigma v \rangle_{\text{ref}} \) is maximum, i.e., \( \langle \sigma v \rangle_{\text{ref}} \propto 2.0 \times 10^{-22} \text{ m}^3/\text{s} \). Then assuming that these optimal conditions can be produced only in a small fraction \( V_{\text{eff}} \) of the total plasma volume (say \( V_{\text{eff}} = (1/25) V \), we estimate

\[
P_F = 6 \langle \rho_{\text{He}} \rangle_{\text{optimal}} (V_{\text{eff}}/m^3) \text{MW} \tag{18}\]

In the case where \( T_e = 20 \text{ keV} \), \( T_i = 15 \text{ keV} \) for example, and \( \langle \rho_{\text{ref}} \rangle = 65 \text{ MW/m}^3 \) over \( V_{\text{eff}} \), we find \( n_{\text{optimal}} = 2.5 \) and \( \rho_{\text{optimal}} = 0.2 \), so that \( P_F = 1 \text{ MW} \). This value is in reasonable agreement with that obtained from the code FPPRF under similar conditions.

### 1.3.5 Deuterium-Tritium Ignition Experiments with High Magnetic Fields

High magnetic field experiments have been designed to investigate deuterium-tritium fusion ignition conditions, on the basis of known experimental and theoretical understanding of plasma behavior. Our group has long maintained a strong effort in pioneering the practical use of such configurations. We have shown that the most advantageous and least expensive designs incorporate an interlocking set of characteristics—tight aspect ratio, relatively small size with significant vertical elongation, high toroidal and poloidal magnetic fields, large plasma currents, high plasma densities, good plasma purity, strong ohmic heating, good plasma and \( \alpha \)-particle confinement, and robustness against ideal MHD and resistive plasma instabilities. We have investigated the physical basis for these plasma properties, while also constructing the design and engineering solutions for attaining the necessary parameters (in the Ignitor UlT machine presently being built in Italy). Since ignition depends upon many spatially and temporally varying processes, many of our studies are based upon the numerical simulation of a free boundary plasma from the current ramp through ignition, using the MHD evolution and plasma flux surface transport code, TSC.  

The ignition of a 50:50 deuterium-tritium plasma requires a minimum value of the parameter \( n_e T_e = 4 \times 10^{20} \text{ sec/m}^3 \) in order to achieve ignition with \( T_e = T_i = 15 \text{ keV} \), where \( n_e \) is the peak plasma (electron) density, \( T_e \) the peak temperature, and \( \tau_e \) the energy replacement time. Here ignition is defined to be the point when the plasma heating due to fusion \( \alpha \)-particles, \( P_{\alpha} \), equals the plasma thermal losses \( P_l \). Relatively high values of the plasma density, \( n_e = 10^{21} \text{ m}^{-3} \), then require only moderate values of \( \tau_e \), whose magnitude is less easy to predict with certainty. Both these values should be achievable, based on the favorable confinement properties of high density plasmas that have been demonstrated by a series of high field experiments, the Alcator A and C and the latest, the Alcator CMOD, at MIT and the FT/FTU devices at Frascati, Italy. Experimentally, the maximum plasma density \( n_e \) that can be supported correlates with the ratio \( B_T/R_o \), where \( B_T \) is the toroidal magnetic field at the center of the plasma column, at major radius \( R = R_o \). On the basis of the Alcator C machine, where \( n_e = 2 \times 10^{21} \text{ m}^{-3} \) was achieved with \( B_T = 12.5 \text{ T} \) and \( R_o = 0.64 \text{ m} \), and the TFTR machine at Princeton, where even larger ratios of \( n_e R_o/B_T \) were achieved, a configuration with \( R_o = 1.3 \text{ m} \) and \( B_T = 13 \text{ T} \) should be able to sustain reliably densities of \( 10^{22} \text{ m}^{-3} \).

A strong toroidal magnetic field also supports a high poloidal field \( B_p \) and correspondingly large plasma current \( I_p \). A significant vertical elongation, e.g., \( \varepsilon = 1.8 \), substantially increases the plasma current that can be carried for a given \( B_T \) and \( R \). If the density correlates with the (volume) averaged toroidal current density, \( \langle I_p \rangle \), then experimental results suggest that somewhat less than 1kA/cm\(^2\) should offer considerable margin to attain the desired peak density \( n_e \sim 10^{21} \text{ m}^{-3} \). High values of \( B_T \) produce a strong rate of ohmic heating, while large currents \( I_p \) tightly confine the fast \( \alpha \)-particles produced by the fusion reactions, so that they deposit their energy in the center of the plasma.

---


Chapter 1. Plasma Dynamics

The degradation of the plasma energy confinement, commonly observed when injected (nonohmic) heating is applied, is reduced at higher plasma current. In addition, the poloidal plasma beta $\beta_p$ can be kept small at ignition to improve the plasma stability and in particular to stabilize the ideal MHD modes with mode numbers $m=1$, $n=1$ that are associated with sawtooth oscillations. Large plasma density combined with good ohmic heating allows ignition at low plasma temperatures. This reduces the fusion power, and therefore the thermal wall loads. The low beta increases the overall margin of plasma stability.

The plasma purity has been shown to improve with increasing plasma density, that is, the effective charge $Z_{\text{eff}} = \sum n_i Z_i / n_0$, decreases monotonically with $n_0$, in an extensive series of experiments starting with the Alcator A. The major effect of impurities is to dilute the concentration of fusing nuclei, while a secondary effect is the increase of power loss due to bremsstrahlung radiation. If auxiliary heating is not used, $Z_{\text{eff}}$ cannot exceed about 1.6 for D-T ignition in the Ignitor, as indicated by our analyses.

Relatively high plasma edge densities also help to confine impurities to the scrape off layer, where the induced radiation helps to distribute the thermal wall loading more uniformly over the plasma chamber surface. The low ignition temperatures associated with high density further help to keep the plasma clean, by reducing the thermal wall loading, that result in sputtering.

Peaked plasma density profiles should be maintained, by external means such as pellet injectors, if necessary. Peaked profiles maintain stability to ion temperature gradient modes that enhance the ion thermal transport. Since the neoclassical (Ware) inward particle pinch is relatively strong in a tight aspect ratio, high field configuration and an anomalous inflow is also present, pellets that penetrate partway into the plasma can be successfully used to raise the plasma density and produce peaked profiles near ignition.

Plasma configurations, such as X-points, that concentrate the thermal (particle) heat flux on localized areas of the vessel wall, limit the amount of fusion power that can be handled. In addition, a more limited plasma current can be sustained. However, X-points and detached plasmas can be obtained with relatively little sacrifice in the plasma and magnet parameters and they may prove desirable to limit the degradation of the plasma confinement caused by nonohmic heating, by creating the conditions known to produce "H-mode" operation.

Divertors represent a more severe compromise, since they alter the design of the plasma chamber and the toroidal magnet. The major radius must be increased to accommodate a reliable divertor, reducing the ratio $B_t / R$ and therefore the maximum plasma density. The magnetic fields are also reduced, lowering the plasma current and the ohmic heating, and increasing $\beta$. A large injected heating system becomes necessary to replace the ohmic heating and the resulting degradation of the plasma confinement makes low temperature ignition difficult. The divertor plates must then handle large thermal heat fluxes. There is no demonstrated advantage to using divertors in high density plasmas.

1.3.6 Time Dependent Ignition

The transient nature of ignition has important consequences. The initial current ramp, when $I_p$, $n_e$, $B_t$, and the plasma cross section are increased simultaneously, to their maximum values, has important effects on the plasma energy balance and stability at ignition. These effects arise from the relatively slow inward diffusion of the plasma current, which is added at the outer surface of the plasma by the current ramp, compared to the growth of the central temperature due to plasma heating. The current ramp generates a inhomogeneous toroidal electric field in the plasma that is peaked near the plasma edge and allows large values of ohmic heating at high central temperature. The magnetic safety factor $q$ can be easily maintained above unity or held to a very small $q < 1$ region during the current ramp, and a more careful study, shows that it can also be kept small after the ramp, at least until the central temperature reaches high values and fusion $\alpha$-particles begin to appear, both of which are stabilizing effects for $m=1$ modes. Furthermore, small amounts of injected heating (e.g., $P_{\text{inj}} < P_{\text{inj}/2}$ during the current ramp can maintain a very small size (or nonexistence) for the $q < 1$ region until well past ignition, if central temperatures approach 10 keV by the end of the current ramp, through the freezing-in of the central current density at low resistivity. Injected heating also reduces the magnetic flux consumption required to reach ignition, particularly if ignition occurs during the current

ramp. These characteristics are illustrated in table 3 (reference case for the maximum ignitor design parameters, peak $B_T = 13.5$ T, $I_p = 12$ MA, $R = 1.3$ m, $a = 0.48$ m, $\kappa = 1.8$, and $n_{\infty} = 1.1 \times 10^{21}$ m$^{-3}$, at 50:50 D:T ratio and $Z_{\infty} = 1.2$).

It has been shown that it is simultaneously possible to maintain monotonically increasing $q$ profiles without large low shear regions, and with edge values $3 < q_{\infty} < 4$ during the current ramp, to beyond ignition. These conditions should prevent instabilities associated with internal plasma modes (e.g., "locked" or quasistationary modes) that can be triggered during the current ramp and often lead to serious disruptions, in which the confinement of the plasma can be lost. Since hollow q profiles are usually associated with the excitation of internal macroscopic modes and enhanced, "anomalous"

Table 3. Reference Discharge for the Ignitor Ult Maximum parameters at the End of Ramp $I_p = 12$ MA,
$B_T = 13.5$ T, $R = 1.3$ m, $a = 0.48$ m, $b = 0.87$ m, $\delta = 0.4$, $n_{\infty} = 1.1 \times 10^{21}$ m$^{-3}$, $n_{\infty}/<n_\phi> = 2.2$, $Z_{\infty} = 1.2$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>End Ramp</th>
<th>Ignition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>3.0</td>
<td>4.3</td>
</tr>
<tr>
<td>$L/2$</td>
<td>0.32</td>
<td>0.38</td>
</tr>
<tr>
<td>$\beta_p$</td>
<td>0.08</td>
<td>0.13</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.8</td>
<td>1.26</td>
</tr>
<tr>
<td>$q_{\infty}$</td>
<td>3.3</td>
<td>3.6</td>
</tr>
<tr>
<td>$W$</td>
<td>7.5</td>
<td>11.7</td>
</tr>
<tr>
<td>$T_e$</td>
<td>4.0</td>
<td>11.0</td>
</tr>
<tr>
<td>$\tau_e$</td>
<td>710</td>
<td>660</td>
</tr>
<tr>
<td>$P_{oh}$</td>
<td>13.0</td>
<td>9.5</td>
</tr>
<tr>
<td>$P_\alpha$</td>
<td>2.0</td>
<td>17.8</td>
</tr>
<tr>
<td>$n_{\infty}$</td>
<td>1.5</td>
<td>12.0</td>
</tr>
<tr>
<td>$P_B$</td>
<td>3.2</td>
<td>4.1</td>
</tr>
<tr>
<td>$P_{IC}$</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>$V_{q-1}$</td>
<td>1.4</td>
<td>5.8</td>
</tr>
<tr>
<td>$\Delta \Phi$</td>
<td>29.2</td>
<td>31.4</td>
</tr>
<tr>
<td>$I_{BS}$</td>
<td>0.6</td>
<td>1.0</td>
</tr>
</tbody>
</table>

The requirement that the edge $q$ is between 3 and 4 (high plasma current) means that special care must be devoted to maintaining $q > 1$ up to ignition. If only ohmic heating is contemplated, the steadily increasing size of the $q < 1$ region after the end of the current ramp imposes a more severe limit on the time in which the plasma can ignite and on the required energy confinement level, than the

A major question for all ignition experiments is the degree of degradation expected in the plasma energy confinement near ignition, since D-T ignition is easily achieved if the confinement remains at the optimal, ohmic heating level. One strategy for a high field experiment is to maintain a high level of ohmic heating up to ignition, $P_\alpha = 2P_{oh}$, to reduce the degree of degradation. Since $\alpha$-particle heating possesses two important characteristics of ohmic heating that are not shared by any presently available form of injected heating—axisymmetric deposition and generation in the center of the plasma column—we expect that the degradation should not be as severe.
energy balance alone, if it is assumed that sawtooth oscillations large enough to destroy the central peaking of the temperature cannot be avoided. Then ignition in the Ignitor reference case, at $T_e = 11$ keV, requires $\tau_E = 0.66$ sec and must occur within approximately 1.5 sec of the end of the current ramp (table 1). Our theoretical analysis, on the other hand, indicates that Ignitor remains, in all regimes, within the stability limits of the ideal MHD and resistive $m=1$, $n=1$ modes. In addition, moderate amounts of auxiliary heating, $P_{aux} \sim 10 - 15$ MW, started during the current ramp, allow ignition down to the limits predicted by $n_e \tau_E$, i.e., $\tau_E \leq 0.4$ sec, while maintaining very small $q=1$ regions well beyond ignition. Similarly, in the ohmic case, if the requirement of small $q < 1$ region is dropped, either on the basis of the theoretical analysis or by externally stabilizing the sawtooth oscillations, ignition can also occur at these $\tau_E$'s and times of $t_i = 5 - 5.5$ sec.

The importance of ohmic heating during the ignition sequence at high field and density means that a model for the electron thermal transport should, like the one used here, simulate ohmic regimes and reproduce typical toroidal loop voltages in steady state ohmic experiments, that are observed to be an almost "universal" constant. In addition, the total diffusion coefficient should increase with injected heating and reproduce the degraded confinement observed in present experiments that are dominated by injected heating.

A second major question for ignition is the effect of variation in the plasma density and its profile, since pellets injected to raise the density are unlikely to fully penetrate a high density plasma. For a given level of thermal transport, there is an optimum density for fastest ignition. A higher density is more favorable under degraded conditions. Higher density, however, accelerates the toroidal current penetration at a given time by lowering the $T_e$, producing larger $q < 1$ regions earlier than at lower density. This effect also operates in the outer part of the plasma radius when density profiles are broadened. Thus, for $n_{eq} = 1.1 \times 10^{21}$ m$^{-3}$, profile peaking factors $n_{eq}/<n_e> \simeq 1.9$ where $<n_e>$ is the volume average give relatively results similar to the standard case ($n_e = 2 \times 10^{21}$ m$^{-3}$) for ignition. Broader profiles rapidly lead to degraded ignition, e.g., $n_{eq}/<n_e> = 1.5$ requires longer $\tau_i$ and higher $\tau_E$, and yields a significantly larger $q=1$ radius. Lower central density, e.g., $n_{eq} = 6.5 \times 10^{20}$ m$^{-3}$ at the end of the current ramp increasing to $8 - 9 \times 10^{20}$ by ignition, allows broader profiles, $n_{eq}/<n_e> = 1.5$. For related reasons, increasing the plasma density after the current ramp is more advantageous than increasing the density during the ramp. At high density, a region of low magnetic shear develops in the mid-region of the minor radius. This region becomes seriously unstable when its value of $q$ approaches unity, since ideal MHD instabilities with $m=1$ can occur. This is one of the major limits on the broad density profile cases.
Section 3  Electromagnetics

Chapter 1  Electromagnetic Wave Theory and Applications
Chapter 1. Electromagnetic Wave Theory and Applications

Academic and Research Staff
Professor Jin Au Kong, Professor Terry P. Orlando, Dr. Kung Hau Ding, Dr. Robert T. Shin, Dr. Y. Eric Yang, Dr. Yi Yuan, Qizheng Gu

Visiting Scientists and Research Affiliates
Dr. Lars H. Bomholt, Dr. Arthur K. Jordan, Dr. Kevin O'Neill, Dr. Antonio Orlandi, Dr. Francesca Scire-Sicappuzzo, Dr. Jean-Claude Souyris, Dr. Leung Tsang, Dr. Michael Teuk, Dr. Herre S.J. van der Zant, Dr. Jiqing Xia, Masanori Yamaguchi

Graduate Students

Undergraduate Students
Robert D. Bock, Emanuel G. Dutra, William K. Lee

Technical and Support Staff
Charmaine A. Cudjoe-Flanders, Kit-Wah F. Lai, Angela R. Odoardi

1.1 Remote Sensing of Earth Terrain

Sponsors
California Institute of Technology/
Jet Propulsion Laboratory
Agreement 952948
National Aeronautics and Space Administration
Grant NAGW-1617
Agreement 958461
U.S. Navy - Office of Naval Research
Grant N00014-89-J-1107
Grant N00014-92-J-1616
Grant N00014-92-J-4098

Project Staff
Professor Jin Au Kong, Dr. Francesca Scire-Sicappuzzo, Dr. Leung Tsang, Dr. Kung Hau Ding, Dr. Robert T. Shin, Dr. Jean-Claude Souyris, Robert G. Atkins, William Au, Pierre Coutu, Hong Tat Ewe, Chih-Chien Hsu, Joel T. Johnson, Alex X. Mou, Ante Salcedo, Shih-en Shih, Murat E. Veysoglu

1 Swiss Federal Institute of Technology, Lausanne, Switzerland.
2 U.S. Navy, Naval Research Laboratory, Washington, D.C.
3 U.S. Army, Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire.
4 Universita degli Studi de L'Aquila, L'Aquila, Italy.
5 Professor, Istituto di Informatica e Telecomunicazioni, Catania, Italy.
6 Centre D'Etude Spatiales des Rayonnements, Toulouse Cedex, France.
7 Professor, University of Washington, Seattle, Washington.
8 Digital Equipment Corporation, Tewksbury, Massachusetts.
9 Schlumberger-Doll Research, Ridgefield, Connecticut.
10 Hitachi Chemical Company, Tsukuba, Ibaraki, Japan.
In this project, we are investigating the theory of polarimetric passive remote sensing of wind-generated sea surfaces with potential application to ocean surface winds. The small perturbation method (SPM), derived to second order, is applied to the Stokes vectors of the thermal emission from random rough dielectric surfaces described by anisotropic directional spectra. To verify the accuracy of the SPM, a Monte Carlo simulation is performed to calculate the Stokes vectors of the emission from the simulated one-dimensional random rough surfaces with a power-law spectrum for various observation angles and surface parameters. The theoretical results of the SPM for all four Stokes parameters are in excellent agreement with the numerical results obtained from the Monte Carlo simulation. The SPM is then applied to small-scale sea surface described by an empirical surface spectrum. The results indicate that polarimetric radiometry allows us to measure wind vector with only one azimuthal observation angle. This is made possible by the different azimuthal dependence between the third Stokes parameter and the first two Stokes parameters. Radiometers with a single-azimuth-observation design will not only simplify the antenna-beam-scanning mechanism, but will also be free from the complexity associated with the data co-registration required for multiple-azimuth-observation designs.

A numerical study of the polarimetric thermal emission from ocean surfaces randomly rough in one dimension using a Monte Carlo technique is carried out. A set of finite length surface profiles with desired statistics was generated using a spectral method. Each surface was extended periodically to create an infinite rough surface, and the thermal emission was computed using the extended boundary condition method (EBC) and the method of moments (MOM). The results from the set of surfaces were then averaged to obtain the Monte Carlo estimate of polarimetric thermal emission. The surface statistics chosen were intended to model a wind perturbed ocean surface in the X to K band microwave region. The results of the study show that the third Stokes parameter, $U_{\theta}$, is sensitive to the azimuthal angle between the surface periodicity and the looking angle, the rms height of the surface, and the surface power law spectrum slope, and that this parameter is insensitive to variations in polar angle, permittivity, and surface spectrum high frequency content.

Experimental results and theoretical calculations are presented to study the polarimetric emission from water surfaces with directional features. It is observed that the measured Stokes parameters of corrugated fiberglass-covered water surfaces are functions of azimuth angles and agree very well with theoretical calculations. The theory, after being verified using experimental data, was then used to calculate the Stokes parameters of periodic surfaces without fiberglass surface layer and with rms height of the order of wind-generated water ripples. The magnitudes of the azimuthal variation of the calculated emissivities at horizontal and vertical polarizations corresponding to the first two Stokes parameters are found to be comparable to the values measured by airborne radiometer and SSM/I. In addition, the third Stokes parameter not shown in the literature is seen to have approximately twice the magnitude of the azimuth variation of either $T_{\nu}$ or $T_{\sigma}$. The results indicate that passive polarimetry is a potential tool in the remote sensing of ocean wind vector.

To interpret the polarimetric active remote sensing data of saline ice, we have developed a random medium model using the radiative transfer theory. The ice layer is described as a host ice medium embedded with randomly distributed inhomogeneities, and the underlying sea water is considered to be a homogeneous halfspace. The random medium model is characterized by a correlation function described by correlation lengths in both the horizontal and vertical directions. Multiple scattering effects are accounted for by solving the radiative transfer equations numerically. The effects of random roughness at the air-ice and ice-water interfaces are accounted for by modifying the boundary conditions in the radiative transfer equations. Analysis of the model for reconstruction of sea ice parameters is made. An optimization approach is used for inversion. The discrepancy between the data and the results of the forward model is minimized by changing the inversion parameters according to a nonlinear programming scheme. Reconstruction of correlation lengths in the horizontal and vertical dimensions has been accomplished using the polarimetric backscattering coefficients at different angles of incidence as input data. Effects of data diversity and noise on the reconstruction of the physical parameters of sea ice from the backscattering coefficients are being investigated.

Vegetation canopy can be modeled as either a mixture of multiple-species discrete scatterers described by a certain size, shape, and orientation distributions, or a continuous random medium characterized by correlation functions. A layer model is developed and applied to interpret radar backscattering coefficients at 5.3 GHz for a soybean canopy. The canopy is modeled as a random medium containing spheroidal scatterers for the leaves. The data were taken over an extended time period from the early to late stages of vegetation. The theoretical results and the experimental data are in good agreement. The lateral correlation length is
observed to be highly correlated with the canopy fractional volume. With consideration of the interrelations among biophysical parameters, the model is then used to simulate backscattering coefficients under various conditions. The results provide sensitivity domains of radar responses to soil moisture and vegetation biomass for inversion assessment. Furthermore, the polarization signatures of the vegetation canopy are synthesized to illustrate structural information conveyed by polarimetric data.

1.2 Electromagnetic Waves in Multilayer Media

Sponsors

Digital Equipment Corporation
Joint Services Electronics Program
Contract DAAL03-92-C-0001
U.S. Navy - Office of Naval Research
Agreement N00014-90-J-1002
Agreement N00014-89-J-1019

Project Staff

Professor Jin Au Kong, Dr. Lars H. Bomholt, Dr. Antonio Orlandi, Dr. Robert T. Shin, Dr. Michael Tsuk, Dr. Jiqing Xia, Dr. Y. Eric Yang, Dr. Yi Yuan, Robert G. Atkins, Pierre Coutu, Hong Tat Ewe, Chih-Chien Hsu, Gregory T. Huang, Joel T. Johnson, Kevin Li, John H. Oates, M. Ali Tassoudji, Murat E. Veysoglu

The electromagnetic radiation from a VLSI chip package and heatsink structure is analyzed by means of the finite-difference time-domain (FDTD) technique. The dimensions of a typical configuration calls for a multizone gridding scheme in the FDTD algorithm to accommodate fine grid cells in the vicinity of the heatsink and package cavity and sparse gridding in the remainder of the computational domain. The issues pertaining to the effects of the heatsink in influencing the overall radiating capacity of the configuration are addressed. Analyses are facilitated by using simplified heatsink models and by using dipole elements as sources of electromagnetic energy to model the VLSI chip. The potential for enhancement of spurious emissions by the heatsink structure is illustrated. For heatsinks of typical dimensions, resonance is possible within the low gigahertz frequency range. The exploitation of the heatsink as an emissions shield by appropriate implementation schemes is discussed and evaluated.

The use of equipment enclosures to shield against emissions and external energy is a common practice. Efforts range from metal impregnated plastics to thick steel panels incorporating waveguide design in air vents. The choice of complexity depends on system performance and application. The question has often been raised as to whether high field intensities within such enclosures will affect equipment performance. Clearly, this is increasingly pertinent in cases of heavy shielding where resonator-like structures with high quality factors result. Moreover, in these cases, any energy leakage may be highly frequency selective with substantial associated field strength. The primary motivation for studying such structures is that computers are now operating at faster speeds and consuming more power, resulting in significantly increased levels of power at higher frequencies. The finite-difference time-domain (FD-TD) technique is used in analyzing this problem and the metrics are total radiated power and field patterns within the enclosure. The excitation sources are a line current source for the two-dimensional models and dipoles for the three-dimensional cases. A Gaussian amplitude is assumed in order to obtain multifrequency characteristics and to exhibit the resonance properties of the model. In addition to the radiated power through the apertures, the power absorbed in the resistive material is also calculated.

In high-speed digital circuits, high frequency phenomena affects the characteristics of the interconnections. Physical discontinuities or non-uniformities in the connections may cause severe reflections when they no longer act like conducting wires, but behave like transmission lines and/or waveguides. In multilayered digital circuits, vias constitute one of the most commonly-used class of interconnects. Vias are not good carriers of high-speed signals because they cause signal distortion and reflections as well as severe degradation in the high frequency components. Previously, analysis and modeling of a single via and some quasi-static or quasi-TEM analyses of single via configurations have been carried out.

To date, analysis of coupled noise between adjacent vias has received very little attention. This is mainly because of the complexity of multi-via structures and the difficulty of modeling and analyzing them accurately over a broad frequency range. The electromagnetic coupling between two adjacent vias in a multilayered integrated circuit is analyzed by means of equivalent magnetic frill array models incorporated with the even- and odd-mode approach. Closed-form expressions for the coupled noise on the passive via are derived. The coupling responses in the frequency domain and crosstalk waveforms in the time domain for some multilayered via structures are calculated based on these formulas. A four-layer experimental model is constructed, and measurements are taken for the transmission, reflection, and coupling responses.
Measurements show good agreement with calculated results over a frequency range of up to 18 GHz.

1.3 Simulation of Electromagnetic Wave Scattering

Sponsors

DEMACO
U.S. Army Cold Regions Research and Engineering Laboratory
Contract DACA89-93-K-0009
U.S. Department of Transportation
Agreement DTRS-57-92-C-00054TTD1

Project Staff

Professor Jin Au Kong, Dr. Kevin O'Neill, Dr. Y. Eric Yang, Philippe Berisset, Emanuel G. Dutra, Hong-Tat Ewe, Qizheng Gu, Chih-Chien Hsu, William K. Lee, M. Ali Tassoudji, Alberto Ugarte, Li-Fang Wang

This is a long-term project for the development of a comprehensive theoretical model for computer simulation of three-dimensional micro- and millimeter wave scattering and emission phenomena. Our major tasks include (1) a simulation program—development of a baseline simulation that includes the effect of environment, hard targets, atmospheric absorption and scattering, and ground reflection; (2) validation and documentation—generation of a synthetic scene for comparison with real data and production of user manuals and validation reports; and (3) upgrade assessment—studies of models that can be added to the baseline simulation and various ways to integrate the models.

Our immediate goals in this research project are to (1) assemble any physical models available that satisfy the minimum requirements defined above and (2) build a preliminary simulation package using these models. The simulation package should demonstrate the process of generating scenes from standard input data sets and the possibility of direct or indirect comparison with recorded images.

We have developed a baseline simulation package with X-window/Motif graphical user interface. This program (1) first accesses the GIS feature map and elevation map and creates displays of them, (2) allows a user to identify and highlight particular terrain features in the map region from a menu, and (3) interfaces with the background terrain scattering coefficient database created by the EMSARS program and the clutter statistics model to generate simulated radar images. A simple terrain shadowing effect is included. The current database is created using the random medium volume scattering model and composite rough surface scattering model.

1.3.1 Publications


Chapter 1. Electromagnetic Wave Theory and Applications

Meeting Papers


Chapter 1. Electromagnetic Wave Theory and Applications


Chapter 1. Electromagnetic Wave Theory and Applications


1.4 Superconducting Transmission Lines

1.4.1 Simulations of Vortices in Arrays of Josephson Junctions

Sponsors
Advanced Research Projects Agency/Consortium for Superconducting Electronics
Contract MDA972-90-C-0021
National Science Foundation
Fellowship MIP 88-58764

Project Staff
Joel R. Phillips, Dr. Herre S.J. van der Zant, Robert D. Bock, Professor Terry P. Orlando, Professor Jacob K. White

Vortices play a central role in determining the static and dynamic properties of two-dimensional (2D) superconductors. Artificially fabricated networks of superconducting islands weakly coupled by Josephson junctions are model systems to study the behavior of vortices. Studies of vortices in Josephson junction arrays generally neglect the magnetic fields induced by currents flowing in the array; i.e., it is assumed that the penetration depth for flux \( \lambda \) is much larger than the size of the array. With the present niobium technology, arrays have been made with \( \lambda \) of the order of the cell size; therefore, effect of induced fields must be considered for an accurate description of these systems.

We use numerical simulation to investigate how a variety of static vortex properties are affected by finite penetration depth \( \lambda \), and we calculate for the first time the self-consistent current and magnetic fields from a vortex in a 2D array. We find that to calculate the correct current and field distributions, we must include nearly all mutual inductance terms to account for the full 3D behavior of the magnetic fields. However, to calculate the energy barrier for cell to cell vortex motion, which was first shown by Lobb, Abraham, and Tinkham (LAT) to be

\[ 0.2 E_J \]

where \( E_J \) is the Josephson coupling energy, including only self and nearest neighbor inductance's is sufficient. The LAT calculation neglected the presence of induced magnetic fields. We show that induced fields may increase the energy barrier substantially above 0.2 \( E_J \). Our calculations also show that the thermodynamic lower critical field of the array is enhanced when the computation self-consistently accounts for induced magnetic fields and that by using only a self-inductance term to model the induced fields, the lower critical field is overestimated.

There has been much recent interest in the properties of Josephson-junction arrays driven by DC and AC currents, particularly in the study of fractional and subharmonic Shapiro step structure. While the occurrence of fractional steps in a magnetic field seems to be explained by a phenomenological model of a moving vortex superlattice, the origin of subharmonic structure, particularly in zero applied field, is less clear. Several explanations have been proposed, including inductive (self-field) effects. There has also been long-standing interest in the use of Josephson arrays as coherent microwave sources, and recently the first observation of coherent emission from 2D Josephson-junction arrays was reported. The arrays used in the experiments had high critical currents, indicating that the strength of the magnetic fields induced in these arrays by circulating supercurrents is fairly large. However, there have been relatively few dynamic studies of Josephson arrays that consider the effect of induced magnetic fields; and most of the previous treatments have severely approximated the form of the induced fields.

We have made dynamic simulations of 2D Josephson arrays where inductive effects are important. We have included the mutual inductance interactions between all cell pairs in the arrays when necessary to calculate the correct screening properties of the induced field. We find that induced magnetic fields can produce subharmonic Shapiro step structure in zero applied magnetic field, and we have developed a qualitative
description of the movement of vortices on the sub-
harmonic Shapiro steps. Although the symmetry-
breaking effect of the self-fields can generate
subharmonic Shapiro steps, the steps do not
appear to be due to the formation of an induced
commensurate vortex state at the edges of the
array. Another effect of the self-fields is that in
arrays with penetration depths small relative to the
size of the array, the screening effects due to
induced fields may be strong enough to destroy
phase-locking over some parameter range, signifi-
cantly reducing the width of the giant Shapiro
steps. Thus, for coherent emission from 2-D
arrays, self-field effects must be reduced.

1.4.2 Superconducting Transmission Lines

Sponsor
Advanced Research Projects Agency/
Consortium for Superconducting Electronics
Contract MDA972-90-C-0021

Project Staff
Laurence H. Lee, Professor Jacob K. White, and
Professor Terry P. Orlando in collaboration with W.
Gregory Lyons

A full-wave spectral-domain volume-integral-
equation method has been developed to analyze
various configurations of superconducting trans-
mission lines: microstrip lines, striplines, coplanar
strips, and coplanar waveguides. In the formu-
lation, a spectral-domain Green's function for
isotropic, layered media is used to set up an inte-
gral equation for the electric field inside the super-
conducting strips. Galerkin's method with roof-top
basis functions is employed to solve for the
complex propagation constant and current distribu-
tion. The characteristic impedance of the structures
is then obtained from transmission line theory. This
method rigorously accounts for the anisotropy and
the finite thickness of the superconducting films,
yielding accurate characterization for the loss and
kinetic effect of the superconductors. However, this
technique is computationally inefficient.

To implement an efficient method, an equivalent
surface impedance is used to transform the super-
conducting strip with finite thickness to an infinitely
thin strip. This equivalent surface impedance
accounts for the loss and kinetic inductive of the
superconductors. An empirical formula for the
current distribution in a thin superconducting film is
determined for the derivation of the equivalent
surface impedance. To include effects of aniso-
tropic substrates, a 2-D dyadic Green's function for
anisotropic, layered media is used to formulate an
integral equation for the surface current. Galerkin's
method with entire-domain basis functions is used
to solve for the complex propagation constant and
the surface current. The characteristic impedance
is then calculated using the power-current definition.
This method has been used to analyze supercon-
ducting single and coupled microstrip lines on
anisotropic substrates.

A closed-form expression was found for the current
distribution in an isolated strip made from a super-
conducting thin film with thickness less than a few
penetration depths. This equation gives a good
approximation within certain limits for the resistance
and inductance per unit length of superconducting
striplines and microstrip lines. Both the ac resis-
tance of an isolated superconducting strip and the
resistance per unit length of a superconducting
stripe calculated using this closed-form
expression agree well with the numerical results
generated by the modified Weeks' method. The
possibility now exists of using this expression along
with the integral equation method to develop an effi-
cient full-wave method for analyzing supercon-
ducting planar transmission lines.

1.4.3 Discrete Superconducting Vortex
Flow Transistors

Sponsor
Advanced Research Projects Agency/
Consortium for Superconducting Electronics
Contract MDA972-90-C-0021

Project Staff
David Berman, Dr. Herre S.J. van der Zant, and
Professor Terry P. Orlando in collaboration with Dr.
K.A. Delin

Vortex Flow transistors (VFT) are of renewed
interest after a discrete flux-flow device made of
weak links in a high-temperature film has demon-
strated promising results. Current gains higher than
one, transresistances of the order of 10 Ω, a
maximum output voltage of about 20 mV, and
speeds of the order of 50 ps were obtained. Flux-
flow devices were proposed more than 10 years
ago and were fabricated in a long continuous
Josephson junction. There are two versions of a

11 MIT Lincoln Laboratory, Lexington, Massachusetts.
VFT in long Josephson junctions; one is a device made of a long overdamped junction (the McCumber parameter $\beta_2 < 1$) and the other is made with an underdamped junction ($\beta_2 > 1$). The disadvantages of long Josephson junctions are that the output resistance is low (much smaller than 1 $\Omega$) and that the output voltage is generally much smaller than 1 mV. Moreover, it is difficult to manipulate the parameters and improve the response time of the device. Discrete vortex flow transistors (DVFT) overcome these difficulties.

Vortex Flow Transistors (VFT) are three-terminal devices: a gate current regulates the number of vortices in the device and a bias current causes the vortices to move which leads to a voltage proportional to the number of vortices. A DVFT consists of a parallel array of Josephson junctions which are coupled to each other by superconducting wires. Each junction is separated from its neighbor by the lattice constant $p$ of the cell. An important parameter in discrete 1D arrays is the penetration depth $\lambda_D$ in units of $p$, which is a measure of the discreteness of the system. The penetration depth $\lambda_D$ is defined as $\lambda_D = \sqrt{L_s/L_i}$. Here, $L_s$ is the self-inductance of a cell in the array and $L_i$ is the Josephson inductance, $L_i = \Phi_0/(2\pi L)$, where $\Phi_0$ is the flux quantum and $L$ is the critical current of a single junction. For $\lambda_D < 1$, vortices are well localized objects; whereas, for $\lambda_D > 1$ vortices are spread out over several cells. From measurements of Fiske modes in underdamped arrays, we have determined that in our DVFT geometry, $L_s = 1.1 \mu_0 p$. We also have shown that for overdamped DVFTs one would like to have $\lambda_D < 1$.

We have fabricated discrete overdamped vortex flow transistors (DVFTs) made of short niobium Josephson tunnel junctions connected in parallel. The results of our DVFT compare with the high-$T_c$ devices and long continuous Josephson junctions. We have also done model calculations on our DVFT which are in good agreement with our experimental results. We calculated the current gain, transresistance, the output voltage and output resistance. Our model is also applicable to other 1D discrete systems and can also be used to model long Josephson junctions. Our analysis shows that, when disregarding the lower temperatures and voltage levels in low-$T_c$ materials, DVFTs made of niobium tunnel junctions can perform at the level comparable to the present high-$T_c$ flux-flow devices.

1.4.4 Fiske Modes in One-Dimensional Parallel Josephson-Junction Arrays

Sponsor
National Science Foundation

Project Staff
Dr. Herre S.J. van der Zant, David Berman, Professor Terry P. Orlando in collaboration with Dr. K.A. Delin

A one-dimensional array of Josephson junctions connected in parallel by superconducting wires may qualitatively be viewed as a discrete version of a single continuous Josephson junction. In a magnetic field, a continuous underdamped Josephson junction shows currents steps in the current-voltage characteristic which are known as Fiske modes. The voltage differences between adjacent Fiske-steps map out the dispersion relation for electromagnetic waves of 1D systems. In a magnetic field, the dispersion relation $\omega(k)$ is linear for a continuous system with some damping, but should bend and exhibit standing wave behavior near the Brillouin zone for a discrete system. Moreover, in discrete arrays the flux through one cell couples to neighboring cells through the mutual inductances. This effect is generally neglected in long Josephson junctions where the magnetic field is assumed to be only parallel to the junction surface.

In arrays, the discreteness parameter is defined as the ratio of the Josephson inductance $L_i$ to the self-inductance of the individual loops $L_s$. The square root of the discreteness parameter is the penetration depth ($\lambda_D$) for 1D discrete system expressed in units of the lattice spacing, $\lambda_D = \sqrt{L_s/L_i}$. For $\lambda_D > 1$, vortices extend over several array cells and the phase of the vortices can be approximated by a sinusoidal wave. This is the linear regime where for a long continuous Josephson junction the Fiske mode analysis has been done. There has not been a systematic analysis of discrete arrays in this linear regime. In the linear regime long Josephson junctions may be used as flux-flow oscillators and parametric amplifiers in the millimeter and sub-millimeter wave range.

We have studied resonances in underdamped, discrete one-dimensional Josephson junction arrays with $A\lambda >> 1$, so that a linear approximation can be used. We have shown that in general the position of the resonances is determined by the dispersion relation of a transmission line of inductors $L_s$ connected in parallel with capacitors $C$ alone, which replace the Josephson junctions. In calculating the dispersion relation, we have taken into account the inductances between adjacent cells $L_{nn}$ and
observed good agreement with the experimental data. Consequently, from measuring the voltages where the steps occur, one can construct the dispersion relation of the 1D discrete system. We have also measured resonances in a 1D array where the primitive cell consists of two different lattice cells. We have observed an acoustic branch in the dispersion relation for low frequencies and an optical branch for higher frequencies which are in agreement with numerical simulations of the full dynamics.

Our Fiske mode analysis on 1D arrays may also be important from a more applied point of view. With current injected at one of the edges of a 1D array, the array critical current varies linearly with applied field and the slope in this region depends on \( \Lambda_0 \) and \( L_2/L_4 \). From measuring Fiske modes in a 1D array, one can obtain \( L_2/L_4 \) and \( L_2/C \) so that this critical current measurement yields \( L_2/(L_4 \) is known because the junction critical current is known) which in turn yields the junction capacitance. Note, that in this procedure all quantities are measured and that no additional information about the barrier or the dielectrics is needed. We have used this method to determine the specific junction capacitance as a function of the critical current density for the PARTS process. Moreover, the array critical current as a function of applied field determines the current gain of a discrete vortex flow transistor and mutual inductance effects can be taken into account through \( L_2/L_4 \). Thus, a good understanding of the dynamics of 1D arrays with inclusion of mutual inductances can be used to determine circuit parameters accurately which is essential when designing SFQ devices and circuits with discrete vortex flow transistors.

### 1.4.5 Publications


Theses


Section 4  Radio Astronomy

Chapter 1  Radio Astronomy
Chapter 1. Radio Astronomy

Academic and Research Staff
Professor Bernard F. Burke, Professor David H. Staelin, Professor Jacqueline N. Hewitt, Dr. Philip W. Rosenkranz, John W. Barrett, Dr. Pauline M. McMahon

Visiting Scientists and Research Affiliates
Dr. Alan Wright

Graduate Students

Technical and Support Staff
Anne E. Conklin, Clare F. Smith

1.1 Extragalactic Radio Source Studies

Sponsor
National Science Foundation
Grant AST 90-22501

Project Staff
Professor Bernard F. Burke, John W. Barrett, Dr. Pauline M. McMahon, Ian M. Avruch, Deborah J. Becker, Samuel R. Conner, Sheperd S. Doeleman, André B. Fletcher, Mark R. Griffith, Lori K. Herold

There were four principal objectives carried out during 1993. These were: (1) an investigation of possible means of detecting exoplanets (planets orbiting other stars); (2) searches for gravitational lenses; (3) a general study of extragalactic radio sources; and (4) an experiment to detect primitive hydrogen concentrations at the epoch before galaxy formation began.

1.1.1 Exoplanets

The investigation of exoplanets is a continuation of work carried out during the previous four years for the Office of Solar System Exploration of the National Aeronautics and Space Administration (NASA). Professor Burke served as chairman of a Science Working Group which studied how NASA could formulate a program to search for and study the formation and evolution of exoplanets. Professor Burke was the Editor-in-Chief of a report, Toward Other Planetary Systems. Since that report was issued by NASA, work has continued on the problem of direct detection of exoplanets. Interferometric methods using aperture-synthesis methods that are commonly used in radio astronomy offer a promising approach to this difficult problem, and preliminary designs for such systems have been studied.

1.1.2 The Search for Gravitational Lenses

The search for gravitational lenses is being carried out by making "snapshot" radio maps with the Very Large Array (VLA) of the National Radio Astronomy Observatory (NRAO). The core source list is the MG (MIT-Greenbank) Survey, and the work is referred to as the MG mapping project. Nearly 10,000 maps of radio sources have been gathered, and approximately 4,000 of these have been reduced by a new procedure that greatly improves the accuracy and signal-to-noise ratios of the final maps. These are then examined to see if morphological structures are evident that are characteristic of gravitational lensing—the distortion of radio signals from distant radio galaxies by foreground galaxies, resulting in the production of multiple images. A new example of the type called an "Einstein Quad" was discovered during the current year and is being readied for publication. Several

1 Australian National Telescope Facility, Parkes, Australia.
candidates have also been identified as "Einstein Rings," and they are now being scrutinized more closely to verify that they are examples of gravitational lensing. To date, our group has discovered six well-established examples of lensing; this is the largest number found by any group in the world.

1.1.3 The Study of Extragalactic Radio Sources

The collection of radio maps that have been produced are useful for many other purposes. A collection of close double sources, with angular separations of under an arc-second, was studied with the British aperture-synthesis array known as MERLIN. This improves upon the angular resolution of the VLA by a factor of five, and from that work a collection of about 15 "compact steep-spectrum radio galaxies" has been identified. A sample of about 150 close double sources is being observed optically, in collaboration with H. Spinrad of the University of California at Berkeley, to determine the redshifts of the parent object, and about 80 redshifts have been obtained so far. Finally, a collaboration with the Australia Telescope of the Australian CSIRO has resulted in the preparation of the Parkes-MIT-NRAO (PMN) Survey of radio sources at 6 cm wavelength. This survey covers the southern sky, from the equator to the South Celestial Pole, and increases the number of known radio sources in the Southern Celestial Hemisphere by an order of magnitude. Two sections have already been published, the Southern Survey (declinations -36 to -87 degrees) and the Tropical Survey (declinations -9 to -30 degrees).

1.1.4 Detection of Primitive Hydrogen Concentrations

A new observational experiment was started this year, in collaboration with P. Horowitz of Harvard University. The universe is known to us on the largest scale through studies of quasars and galaxies and by the cosmic microwave background (CMB). The CMB is at a redshift of about 1,000 and represents the greatest distance that can be studied by present methods. Galaxies and quasars are more local, having been detected in a limited way out to redshifts of about 4.8. The formation of galaxies and quasars captures most of the primeval hydrogen, which was initially the most common constituent of the universe. It appears that beyond a redshift of about 5, the galaxies were still being formed, and there should be large concentrations of neutral hydrogen still in the process of contracting to form galaxies. An experiment has been mounted to look for these primeval concentrations, using the Arecibo Observatory in Puerto Rico, where the 1,000-foot radio telescope is ideally suited to this search. A radio spectrometer is now under construction, and a prototype section of the spectrometer has been built and is currently undergoing tests at the Arecibo Observatory.

1.1.5 Publications


Invited Talks


1.2 Studies of Gravitational Lenses

Sponsors
Alfred P. Sloan Fellowship
David and Lucile Packard Fellowship
Award for Science and Engineering
National Aeronautics and Space Administration
National Science Foundation
Presidential Young Investigator Award

Project Staff
Professor Jacqueline N. Hewitt, Grace H. Chen, John D. Ellithorpe, Charles A. Katz, Mark D. Messier, Christopher B. Moore

Gravitational lenses provide radio astronomers with the opportunity to re-examine many long-standing problems in astrophysics. Two of these problems, addressed by the work described here, are (1) the nature and distribution of luminous and dark matter in galaxies and (2) the calibration of the age and scale of the universe (i.e., the determination of the cosmological Hubble parameter).

These applications require that the distribution of gravitating matter be inferred from observed imaging properties of the lens. This, in turn, requires that we make detailed measurements of the brightness distribution of the image. Detailed radio maps of two gravitationally lensed extragalactic sources, MG1131+0456 and MG0414+0534, are being compiled from data taken with the National Radio Astronomy Observatory (NRAO) Very Large Array (VLA) radio telescope and the recently completed telescopes of the NRAO Very Long Baseline Array (VLBA). Figure 1 presents a VLA image of MG1131 that shows the polarization properties of the object. The unusual morphology is due to the gravitational lens that distorts the background radio source into a system of arcs that form a ring.

Figure 2 shows MG0414; the four components are believed to be images of the same background quasar. The polarization information of figure 1 and the high-resolution imaging of figure 2 are being exploited to model the mass distribution of the lens. This project is being carried out in collaboration with Professor Christopher Kochanek of Harvard University.

In a gravitational lens system, there are differences in the propagation time in the order of days to months for the different ray paths associated with the images of the background source. Monitoring programs aimed at measuring these time delays are being carried out with the VLA. Analysis of these light curves has shown that their variations are correlated, providing further evidence that the MG0414 system is gravitationally lensed and indicating that, with increased signal-to-noise, time delays will be measurable. Measurements of the time delays would provide a gauge of the distances involved in the geometry of the gravitational lens system. Such an independent measure of distance would provide a calibration of the cosmological Hubble parameter. The monitoring is continuing.

1.2.1 Publications


Invited Talks

Hewitt, J.N. "Gravitational Lenses, Time Delays, and Hubble's Constant." University of Maryland, Department of Astronomy Colloquium, April 21, 1993.

Chapter 1. Radio Astronomy


1.3 Radio Interferometry of Nearby dMe Stars

Sponsors
David and Lucile Packard Fellowship
Award for Science and Engineering
National Aeronautics and Space Administration
Grant NAGW-2310

Project Staff
Professor Jacqueline N. Hewitt, John D. Ellithorpe, Christopher B. Moore

dMe stars are dwarf M stars that show evidence of unusual surface activity. For some time, these dwarf stars have been known to flare strongly at optical and radio wavelengths. More recently, it has been demonstrated that many dMe stars exhibit low-level quiescent emission that is detectable by Very Long Baseline Interferometry (VLBI). The detection of dMe stars on VLBI baselines makes possible measurement of their position with high precision and shows that astrometric detection of planetary companions may be feasible. The astrometric monitoring program has continued with the analysis of data at three epochs spanning three years.

1.4 The Detection of Astronomical Radio Transients

Sponsor
David and Lucile Packard Fellowship
Award for Science and Engineering

Project Staff
Professor Jacqueline N. Hewitt, John D. Ellithorpe, Charles A. Katz, Christopher B. Moore

The detection of radio transients of astrophysical origin is of interest for a number of reasons. First, some already well studied sources are known to exhibit radio transients, and further characterizing their nature would increase our understanding of the emission processes involved. Examples include the sun, flare stars, and interacting stellar systems. Second, there have been serendipitous discoveries of radio transients, apparently of astrophysical origin, but with no known counterpart at other wavelengths. Since these may represent a new class of object, we would like to establish whether they are natural phenomena and ascertain their frequency of occurrence. Third, transients at other wavelengths are known to exist. The NASA Compton Gamma Ray Observatory is detecting gamma ray bursts at a rate of approximately one per day. X-ray transients have also been detected; some of these are associated with transient radio sources. In the near future, an all-sky monitor will be launched on the X-ray Transient Explorer. An all-sky radio monitor is needed to detect any radio counterparts to the x-ray and gamma ray transients and to put upper limits on their fluence. Finally, because of the confusion with man-made interference, the investigation of radio phenomena on short-time scales is an area that has not been fully explored. As such, there is potential for the discovery of new phenomena.

To address the scientific questions outlined above, we are proposing construction of a network of geographically separated correlating arrays. The technical feasibility of the project depends critically on the levels of man-made radio interference at the sites being considered. Current efforts are geared toward measuring the interference environment, as well as developing a design for the arrays.
Figure 1. A VLA image of the polarized radio emission in the gravitational lens MG113+0456.
Figure 2. Radio images of the gravitational lens MG0414. The low-resolution image is constructed from data taken with the VLA. The insets show high resolution images constructed from data taken with the newly completed VLBA antennas and correlated at Haystack Observatory.

1.5 Algorithms for Advanced Microwave Sounding Unit Operational Use

Sponsor

MIT Lincoln Laboratory
Agreement BX-4975

Project Staff

Dr. Philip W. Rosenkranz, Michael J. Schwartz

This project provides scientific support to the National Oceanic and Atmospheric Administration (NOAA), for the Advanced Microwave Sounding Unit (AMSU), which is scheduled for launch on NOAA polar-orbiting weather satellites starting in 1995, and also for follow-on instruments. The effort emphasizes development of atmospheric transmittance algorithms, estimation of surface emissivity and precipitation, issues related to instrument design and specification, and general retrieval methods.

Measurements from the MIT aircraft-based Microwave Temperature Sounder and from present satellite radiometers, such as the Microwave Sounding Unit, are used to develop and test prototype algorithms.

A rapid algorithm for computing transmittance in the mesosphere, where the oxygen lines are split by the terrestrial magnetic field, has been developed. This algorithm should prove useful for processing data from future generations of satellite instruments (beyond AMSU).

1.6 Earth Observing System: Advanced Microwave Sounding Unit

Sponsor

National Aeronautics and Space Administration/
Goddard Space Flight Center
Contract NAS 5-31276
The AMSU will provide microwave-band measurements to the Atmospheric Infrared Sounder (AIRS), which will be a facility instrument on NASA's planned Earth Observing System (EOS). This effort is directed toward developing algorithms for the retrieval of temperature and humidity profiles, precipitation, sea ice, land snow cover, and other parameters. The temperature and humidity profiles obtained from AMSU are intended to provide initial conditions for the AIRS infrared cloud-correction and profile retrieval algorithms. Work to date has focused on simulation efforts. A rapid (as distinguished from line-by-line) microwave transmittance algorithm for lower-atmospheric oxygen, water vapor and liquid water was developed.

Artificial neural network techniques have been applied to retrieval of humidity profiles from simulated microwave spectral observations. At pressures greater than 700 mb over land, an improvement of several percent in relative humidity was obtained in comparison to the statistical iterative technique. Preliminary work indicates that neural networks can provide more compact representations of radiometric data than linear methods can.

1.6.1 Publications


Thesis


1.7 High-Resolution Passive Microwave Imaging of Atmospheric Structure

Sponsor

National Aeronautics and Space Administration/Goddard Space Flight Center
Grant NAG 5-10

Project Staff

Professor David H. Staelin, John W. Barrett, Dr. Philip W. Rosenkranz, Michael J. Schwartz, Paul W. Fieguth

The MIT Microwave Temperature Sounder (MTS) is a dual-band radiometer, incorporating an imaging spectrometer at 118 GHz and a tunable fixed-beam radiometer at 52-54 GHz. This instrument is flown on the NASA ER-2 high-altitude aircraft. In February 1993, we operated the MTS on ER-2 flights based at Townsville, Australia. These flights were an adjunct to the international TOGA-COARE experiment. Two overpasses of tropical cyclone Oliver were made, yielding unique high-resolution images at 118 GHz.

In September and October 1993, we participated in the CAMEX experiment, in which the ER-2 was based at Wallops Island, Virginia. It flew underneath a DMSP satellite and over convective activity near Florida. Besides the MTS, other instruments similar to those that will be on EOS were on board.

On other flights during these two experiments, upward-observing profiles of atmospheric brightness temperature were obtained. These will be used to test atmospheric transmittance algorithms.

1.7.1 Publications


---


Chapter 1. Radio Astronomy


1.8 Rapid Precision Net-Form Manufacturing

Sponsor
MIT Leaders for Manufacturing Program

Project Staff
Professor David H. Staelin, Paul W. Fieguth

An active CCD-based imaging system developed in 1992 was further documented. It can measure shapes optically within a scannable volume of ~20 cubic inches with rms accuracies of as little as 50 nm for mirror-like surfaces, and ~1-2 micrometers for surfaces like chalk, plastic, and copper.

1.9 Conformal Experiment Design

Sponsor
MIT Leaders for Manufacturing Program

Project Staff
Professor David H. Staelin, Paul W. Fieguth, Rizwan R. Koita, Michelle S. Spina

We continued our work on the development of a user-friendly program to assist engineers with both experiment design and analysis of results. The program accepts a list of control parameters and interactions between them in their a priori estimated order of importance so that orthogonal experiment arrays can be designed to minimize confusion when experimental results are interpreted.

This program and its experiment design were applied to the design of automotive fuel pumps to improve their pressure, flow, efficiency, and other characteristics. From eight trial pumps with three control variables, it was found that simple modifications to a minor part could profitably increase the manufacturing yield of certain products.

1.10 Recognition of Natural Nearly Repetitive Signals

Sponsor
MIT Leaders for Manufacturing Program

Project Staff
Professor David H. Staelin, Kevin G. Christian

Dolphins (and some machinery) emit nearly repetitive whistles resembling bird song. A database of 1169 dolphin whistles from 22 animals was compressed by converting them to single frequency-versus-time traces and then to a 16-dimensional coding space where each dimension is a linear or nonlinear operator on the trace. Consideration of the accuracy with which dolphins could repeat their own whistles within this 16-D space suggests their vocabulary could potentially exceed a billion distinguishable whistles. However, when the accuracy with which one dolphin could reproduce the signature whistles of another was considered, the potential shared distinguishable signature-whistle vocabulary of dolphins appeared to be limited to less than a few hundred whistles. This result is based on sufficiently limited copying data that extending these experiments is important if definitive results are to be obtained.

This database of 1169 whistles was also used in whistle-retrieval experiments in which the whistle in the database nearest to a new input whistle was to be retrieved in minimum time. It was found that ~35-53 msec was the average retrieval time for a non-optimized program on a SparcStation 2 when only the six dimensions with the highest signal-to-noise ratios were employed for retrievals. The algorithm employed was slower if other than six dimensions were used for the initial addressing step.

Thesis


Part III  Systems and Signals

Section 1  Computer-Aided Design

Section 2  Digital Signal Processing
Section 1 Computer-Aided Design

Chapter 1 Custom Integrated Circuits

Chapter 2 Computer-Integrated Design and Manufacture of Integrated Circuits
Chapter 1. Custom Integrated Circuits

Academic and Research Staff
Professor Jonathan Allen, Professor John L. Wyatt, Jr., Professor Jacob K. White, Professor Srinivas Devadas, Professor Dimitri A. Antoniadis, Professor Berthold K.P. Horn, Professor Hae-Seung Lee, Professor Steven B. Leeb, Professor Terry P. Orlando, Professor Martin A. Schmidt, Professor Stephen D. Senturia, Professor Charles G. Sodini, Dr. Robert C. Armstrong, Dr. Craig L. Keast

Visiting Scientists and Research Affiliates
Dr. Andrew Lumsdaine, Dr. Ichiro Masaki, Dr. Herre S.J. van der Zant, Dr. He Yie

Graduate Students

Technical and Support Staff
Susan E. Chafe, Dorothy A. Fleischer

1.1 Custom Integrated Circuits

Sponsors
IBM Corporation
National Science Foundation

Project Staff
Professor Jonathan Allen, Dr. Robert C. Armstrong, Donald G. Baltus, Robert G. McDonald, Chin Hwee Tan

The overall goal of VLSI computer-aided design (CAD) research is to provide the means to produce custom integrated circuits quickly, correctly, and economically. Traditionally, correctness has been verified at several representational levels of abstraction, such as required topologies for given circuit families (e.g., static CMOS) and stipulated logic formalisms. These techniques for checking correctness are usually local to the particular representational level involved and, while they are important components of the overall design testing procedure, they do not attempt to provide for the alignment and consistency checks between the different abstract representational levels and an input behavioral specification. In addition, they do not characterize the set of possible designs at each representational level corresponding to the initial functional specification in a way that ranges over the wide variety of possible performance levels. For this reason, there is a strong need to provide CAD tools that serve as a framework for design exploration when performance is a parameter, thus providing the desired performance together with consistently aligned representations at all levels.

This research group studies a variety of topics with an emphasis on performance-directed synthesis of custom VLSI designs. An overview of the viewpoint that motivates these projects is provided in a major survey paper in which the need for coordinating the design optimization process over the several levels of representation is emphasized. Since design exploration is so central to the production of

---

1 Notre Dame University, Notre Dame, Indiana.
2 Computer Science Department, General Motors Research Laboratory, Warren, Michigan.
3 Microelectronics Center, Southeast University, Nanjing, China.
4 Mathworks, Natick, Massachusetts.
high-performance designs, emphasis is placed on how performance can be characterized at the several levels of representation and how overall optimal performance can be achieved in an integrated way.

The conversion of a high-level functional specification to a detailed mask layout for a substantial processor is a major task that confronts modern CAD synthesis systems. When the target architecture for such a system is restricted, however, it is often possible to introduce metrics that facilitate optimized performance. With this in mind, graduate student Donald G. Baltus has focused on the optimal design of array logic using systolic representations that allow only nearest neighbor communication between processors in the array. Given this architectural restriction, it is possible to introduce realistic communication constraints into the synthesis process so that meaningful optimization of total delay for a calculation can be achieved. An input language called FLASH has been introduced to represent algorithms functionally. These algorithms are then converted to data-dependence graphs that are augmented to provide indexed representations, thus allowing compact representation of calculations that do not grow with the size of the array.

A major achievement of this work is the introduction of several constraints that allow for substantial reduction of the design space, yet without the possibility of rejecting candidates for the optimal design. One such constraint is due to locality, which requires communication only between nearest neighbor processors. While this constraint may seem restrictive, it introduces the benefit of design compaction through a set of regular transformations that increase the performance of the resultant configuration. There is a strong mathematical basis for this work, and it has been shown that the entire space of all candidate designs is comprehensively searched in an efficient way. The thesis that represents this work is now complete, and its results were recently presented at a meeting on array architectures. It is believed that this is the first time guaranteed delay optimal designs for a broad class of systolic architectures is available. Experience with the system has indicated that many nonintuitive designs are revealed. Current extensions include the coupling of this program (called Descartes) to an augmented layout synthesis system so that it will be possible to generate high-quality layouts directly from a high-level functional specification for this important class of architectures.

These results are an important contribution to systems synthesis and architectural exploration. They are fundamentally and mathematically based to scientifically facilitate high-quality synthesis, rather than being based on heuristic strategies, which are often motivated from a limited understanding of the basic design problem.

In an attempt to minimize power while preserving speed performance, graduate student Chin Hwee Tan has developed a design system based on a standard cell library that seeks to provide the minimum power implementation subject to a given delay constraint. The motivation for this research is the observation that it may be possible to increase the size of devices driving the node in order to gain speed without an excessive power penalty, if indeed the node can be demonstrated to switch at a sufficiently low frequency.

These studies take advantage of new capabilities to estimate the switching frequency of nodes in combinational circuits, and hence, they represent a new class of studies based on realistic power estimates. Ms. Tan has created a standard cell library, including detailed layouts where sizing has been introduced, to achieve a range of trade-offs between delay and power. Specifically, three designs for each gate have been introduced in the library and are available to the design system. Once the design has been mapped to a set of gates, and a delay specification has been introduced, power minimization is achieved by initially finding the critical path through the network. This is the longest delay path in the network, and hence, the first candidate for the introduction of faster cells in order to meet the delay specification while potentially minimizing added power.

A variety of strategies were introduced to increase performance, including the identification of nodes that present a heavy load to the previous stage, while also having a low switching frequency. These nodes identify the previous stage as a candidate for the use of cell versions that have wider transistors and are capable of driving the large load at a higher speed. Reordering of inputs to gates (such as for a three-input NAND gate) can be shown to have a

---


substantial effect on performance, while not disturbing the power dissipated by the gate. The algorithms in this system provide a rigorous capability to exploit this potential. Techniques were also introduced to relax and decrease the size of transistors that are not on and do not influence the critical path, and hence, these transistors may be made smaller. This is a desirable direction for several reasons, including the fact that the algorithm tends to equalize the delay of paths to the network, reducing the possibility of glitches and providing a more optimal design from the viewpoint of wave pipelining. Experience with a wide class of combinational circuits has shown that it is possible to reduce delay up to 43 percent with only an 8 percent increase in power.

Possible extensions to the algorithm include the development of a similar yet separate algorithm to optimize delay under a power constraint. In addition, consideration of false paths will be introduced, and the study of basic techniques to optimize convex systems that use discrete variables will be made. Although this study has focused only on static CMOS circuit topologies, other circuit styles may also be investigated.

As reported in previous RLE Progress Reports, Armstrong has developed a comprehensive system for the incremental consistency maintenance in a multirepresentational structural VLSI database. Since the original presentation of this research, several extensions have been made. These were motivated by the overall complexity of the code needed for such a system and requirements for flexibility and interfaces to a variety of other CAD systems.

The most significant accomplishment was the generation of the entire VLSI CAD database system from high-level schema descriptions. This capability includes the potential for custom graphic user interfaces when specified as schemas. Using such schemas, a few hundred lines of high-level code can be used to generate many thousands of lines of complex low-level code that is extremely difficult to develop and is highly prone to errors. For example, new techniques for high-speed constraint propagation were introduced through the compilation process from high-level constraint descriptions to on-line propagation code. This provides the designer with greater flexibility, and also allows the system to be custom-made and generated quickly. In addition, the levels of representation were extended to include physical mask layout, whereas previously, only symbolic layout was represented.

A client interface to the original system was developed and used with SPICE for a variety of teaching applications. For example, several different integrated circuit inverter configurations can be introduced and manipulated while almost instantaneously comparing the results available from the system. The SPICE interface is arranged so that whenever any aspect of the circuit representation is changed, a new SPICE simulation is automatically triggered and displayed. This kind of capability is extremely useful for design, and it can exploit the fact that the design can be manipulated at several different levels of representation in such a way that the database system automatically propagates these changes to the circuit level representation, where they are utilized by SPICE. Thus, for example, the user can manipulate a layout in such a way that changes in the circuit representation will immediately yield new SPICE waveform representations. The user is also able to attach both current and voltage probes to any point in the network, and multiple probes are easily accommodated. Experience has been gained from classroom use, and it has been shown that a greater understanding of circuit behavior can be acquired from the interactive use of this system.

Attention has also been directed to the utilization of this database in relation to existing CAD systems. Feedback from industrial users has indicated the need for a client interface to existing CAD systems, as well as adherence to today's evolving industry standards for database capabilities and interface styles, such as those introduced by the CAD Framework Initiative. Adherence to these specifications increases the portability of the system and its overall usefulness. Consideration is also being given to the possibility of using a commercial object-oriented database in order to provide long-term data storage functionality. Nevertheless, the original database design for this project has demonstrated its ability to provide satisfactory performance in the interactive mode, which has not always been possible with commercially based systems. There are existing languages to express high-level

---


Chapter 1. Custom Integrated Circuits

schemas, such as Express, and the utilization of these languages is currently being studied.

In the past, true single-phase clock circuits were studied in Ms. Lisa A. Pickelsimer’s thesis. In terms of understanding the high-speed behavior of these circuits and generating a design methodology to systematically provide optimal designs in this circuit style. Now, an additional circuit style is being investigated that shows potential for extremely high performance. Self-resetting CMOS circuits are an interesting class of circuits that can provide for very high speed and utilization within a wave pipelining discipline. These circuits do not utilize level signals (except at clock transitions) and instead generate pulses; the leading edge of which is generated by the normal circuit transition. The falling edge is generated by a rapid feedback loop that automatically limits the length of these pulses. While interesting examples of these self-resetting circuits have been demonstrated in the literature, a detailed investigation of their general properties, configurations, design methodology, and potential applications is currently underway in graduate student Robert G. McDonald’s thesis project.

High-performance computing performance for CAD continues to be provided through a variety of configurations. These include the networking of several processors together, using the parallel virtual machine (PVM) methodology, as well as new scalable parallel architectures that provide for a tighter coherence between the existing (workstation-like) processors. Work is underway to acquire such a processor, which will be useful for many CAD algorithms that exhibit substantial parallelism, as well as the support of multiple client utilization of CAD databases with extremely fast communication between processors.

1.2 Analog VLSI Systems for Integrated Image Acquisition and Early Vision Processing

Sponsor
National Science Foundation
Grant MIP 91-17724

Project Staff
Professor Berthold K.P. Horn, Professor Hae-Seung Lee, Professor Charles G. Sodini, Professor Jacob K. White, Professor John L. Wyatt, Jr., Dr. Craig L. Keast, Dr. Ichiro Masaki, Steven J. Decker, Lisa G. Dron, Ibrahim M. Elfadel, J. Mikko Hakkarainen, Frederick P. Herrmann, David R. Martin, Ignacio S. McQuirk, Mark W. Reichelt, Mark N. Seidel, Luis M. Silveira, Christopher B. Umminger, Paul C. Yu

1.2.1 Project Summary

In real-time machine vision the sheer volume of image data to be acquired, managed and processed leads to (1) communications bottlenecks between imagers, memory, and processors, and (2) very high computational demands. We are designing and testing experimental analog VLSI systems to overcome these problems. Our goal is to determine how the advantages of analog VLSI—high speed, low power and small area—can be exploited and its disadvantages—limited accuracy, inflexibility and lack of storage capacity—can be minimized. The work is concentrated on early vision tasks, i.e., tasks that occur early in the signal flow path of an animal or machine.

The next section details substantial progress on seven different designs, including systems for camera motion estimation, automatic alignment to calibration marks, brightness adaptive imaging, and other tasks. Perhaps the most exciting news concerns the stereo vision system, the initial portion of which was completed by Mikko Hakkarainen in his doctoral work with Professor Hae-Seung Lee last May.

This project has attracted the attention of General Motors, which sent a vision systems engineer, Dr. Ichiro Masaki, to work with us for a year to investigate the potential applications of this type of design.

---


to future use in automobiles. In directing our attention to practical automotive applications, Dr. Masaki has made a significant contribution to this project.

1.2.2 Vision Chip Designs

**BICMOS/CCD Focus-of-Expansion Chip for Camera Motion**

Graduate student Ignacio McQuirk, working with Professors Berthold Horn and Hae-Seung Lee, has continued development and testing of a chip that determines the direction of camera translation directly from a time-varying image sequence with a real-time analog VLSI system. Mr. McQuirk's approach assumes that a camera moving through a fixed world has translational velocity only. The focus of expansion (FOE) is the projection of the camera translation vector onto the image plane and hence gives the direction of camera motion when the camera parameters are known. The FOE is the image of the point towards which the camera is moving and the point from which other image points appear to be receding. Various possible algorithms for estimating the FOE have been compared in terms of accuracy, robustness, and feasibility for single-chip analog VLSI implementation with on-chip imaging circuitry. An algorithm based on minimizing the sum of squares of the differences at stationary points between the observed time variation of brightness and the predicted variation was chosen for implementation. Mr. McQuirk designed a BICMOS/CCD chip for estimating the FOE using on-chip CCD imaging circuitry and a row-parallel processing scheme. A complete 64x64 version of this chip has been fabricated through MOSIS.

Mr. McQuirk is currently in the process of building a test setup for the FOE chip. From an electrical standpoint, this involves the completed circuit design and subsequent fabrication of a printed circuit board with the support components required for operation of the FOE system in all its various testing and operational modes. This board consists of pattern generators, CCD clock drivers, FOE chip bias circuitry, I/O drivers, a Feedback Control Processor, and an interface to the Vision Chip Test System. The Vision Chip Test System itself will be upgraded to handle the multiple channels necessary for the FOE chip I/O. This will permit us to test the FOE chip with electrical data, allowing for algorithmic, system, and circuit level characterization. For operation using optical input through the on-chip imager, Mr. McQuirk is also in the process of completing a setup which will optically present the FOE chip with a real-time image sequence due to 3-D motion. This will be accomplished using a flexible fiber optic image guide to transmit the image from a lens system mounted on a precision linear track to the FOE test board. This additionally requires the imager on the FOE chip to be calibrated to ascertain accurately the projection from the 3-D motion to the FOE location.

**Brightness Adaptive CCD Imager with Analog and Digital Output**

Graduate student Steven Decker, working with Professor Charles Sodini, has been designing an imager to act as the front end of a modular vision system. It will feature a 64x64 array of pixels, large fill factor, operating speed on the order of 1000 frames/second, column parallel output, brightness adaptation capability, and user selectable analog or 10-bit digital output.

The cyclic A/D converter is based on an algorithm which requires only addition, division by two, and comparison of two signals. The first two operations are performed in the charge domain because the input analog signal is a charge, and CCDs can add and evenly divide charge packets with good precision using little area. Charge packet splitting accuracy is greatly increased by an algorithm which passes each half-packet through the charge splitter again and appropriately recombines the resultant quarter-packets. Analog output can be obtained by shifting the signal charge onto an integrating capacitor in a feedback loop around an operational amplifier. The linearity of the analog output is much better than for a conventional CCD output because the poly-poly capacitor in the feedback loop is much more linear than the diffusion capacitance of a conventional CCD output.

The brightness adaptive imaging element is based on a design that was previously patented elsewhere. Each pixel is composed of an MOS capacitor biased in deep depletion and an overflow gate. The overflow gate establishes a barrier to the flow of imaged charge from the main potential well to a sink. During the integration period, the height of the barrier \( \phi(t) \) is varied as a function of time. A relationship between the light intensity and final imaged charge can be derived for a given \( \phi(t) \). In particular, it is possible to obtain a logarithmic compression curve so that, ideally, any light intensity can be sensed without saturating the imager. The advantages of this technique over other schemes which have been considered are simplicity, flexibility, and small size. Professor Wyatt has developed a method to obtain the required \( \phi(t) \) to achieve a given compression curve.

Efforts have been directed towards designing a practical realization of the brightness adaptive
imager. In the original patent, \( \phi(t) \) is assumed to be a continuous, infinitely precise function of time. The implications of approximating \( \phi(t) \) with a digital generated staircase function have been examined. Using the digital generation scheme and digital output, reconstruction of the original light intensity from the imaged charge is easy since this can be done digitally. Digital generation of \( \phi(t) \) also makes it possible to alter the compression function in real time based on previously imaged frames.

A test chip has been designed and layed out. This chip has a 4x4 array of wide dynamic range pixels, four adjacent A/D converters, and four adjacent analog output circuits. The chip is currently being fabricated in the MIT CCD/CMOS 1.75 micron process. A test setup is being designed concurrently.

**Single-Chip Alignment Sensor**

Graduate student Christopher Umminger, working with Professors Charles Sodini and Berthold Horn, has been researching implementations of single-chip sensors to measure alignment error for tasks in automated manufacturing. These sensors will optically acquire an image of the alignment marking and output a measure of the misalignment. Currently, they have settled on two methods for the task of aligning a mark to the sensor. The first is based on the surveyor's mark. The sensor consists of a four quadrant photodiode combined with an edge detection circuit to provide extra alignment accuracy in the face of nonuniform illumination. The second sensor design uses a square grating mark. A position sensitive moiré fringe pattern formed between the sensor and the marking is used to achieve accurate alignment.

The first sensor design using the surveyor's mark was fabricated in 1992. Testing of the imaging photodiodes and associated amplification circuits has shown that they function as expected. In July 1993, a paper was presented on this chip at the SPIE Annual Meeting in the session on optical alignment. A second version of the sensor which significantly improves its speed was submitted for fabrication in August 1993. Design and layout of the moiré fringe sensor was finished in April and the fabricated chip has just been received. Testing on this chip will occur in the fall. A mock alignment system is being constructed to characterize these sensors in an alignment operation.

**Content-Addressable Parallel Digital Image Processor**

Massively parallel associative processors may be well suited as coprocessors for accelerating machine vision applications. They achieve very fine granularity, as every word of memory functions as a simple processing element. A dense, dynamic, content-addressable memory cell supports fully-parallel operation, and pitch-matched word logic improves arithmetic performance with minimal area cost. An asynchronous reconfigurable mesh network handles interprocessor communication and image input/output, and an area-efficient pass-transistor circuit counts and prioritizes responders.

Graduate student Frederick P. Herrmann has completed the design of a 16K-trit associative processing chip, under the supervision of Professor Charles Sodini. The chip integrates 256 processing elements, each with 64 trits of memory. A 16x16 mesh network connects the processing elements, and larger meshes may be constructed using several chips. Fabrication in MIT's 1.5 um CCD/CMOS process was completed in February. Some subsystems were tested and their functionality was verified. High leakage current on that run prevented complete testing. The design has been resubmitted to MIT and to MOSIS, with chips expected in September.

Mr. Herrmann has also been working with Jeffrey Gealow and Lawrence Hsu (both with the MIT Associative Processing Project) on a software development system for the associative processor. Simulators for the parallel processor array and its sequential controller were demonstrated in July, and the system is now being used to prepare test vectors for the associative processing chip.

**Video Rate Analog-to-Digital Conversion**

Graduate student Paul Yu, working with Professor Hae-Seung Lee, has been working on an operational amplifier that uses a replica-amplifier gain-enhancement technique. The op-amp is a key component of a pipelined A/D converter. The replica-amp technique can be generally applied to any conventional op-amp topologies such as folded-cascode and class-AB, single-ended or differential. Unlike conventional techniques such as cascoding which increase gain by increasing the output resistance, the replica-amp technique increases gain by matching main and replica amps. Among the advantages of the replica-amp technique are low supply, high swing, and effectiveness in driving resistive loads. A two-stage topology without any cascode was chosen to maximize the available output swing. The experimental result of this low-
supply, high swing, CMOS op-amp was presented at the 1993 International Solid-State Circuit Conference.

An analytical study has been conducted with regard to the settling time of an op-amp using this replica-amp gain enhancement technique. It was shown that this technique has only a small effect on the settling time. This analytical result is confirmed by a SPICE simulation as well as by the experimental results. The results of this study were recently submitted to IEEE Circuits and Systems II.

In view of the trend toward ever decreasing supply voltages for power minimization and the fact that conventional techniques do not work well at low supply, the replica-amp technique offers a promising alternative.

Our attention has shifted toward the capacitor mismatch problem which contributes to gain-error in a pipelined A/D converter. A simulation program has been written in C programming language to investigate a new technique of minimizing differential non-linearity (DNL). A Monte Carlo simulation has been carried out to determine the minimum capacitor size necessary under this new technique. It appears that the technique can lead to a low-power pipelined A/D converter operating at video speeds.

**Design of Specialized Analog and Digital VLSI Processors for Computing the Relative Orientation Between Two-Camera Systems**

Computing the relative orientation between the coordinate systems of two cameras is an important problem for calculating depth from binocular stereo and for determining general camera motion. Graduate student Lisa Dron, working with Professor Berthold Horn, has continued work on the design of a complete, autonomous system for computing relative orientation in real-time with specialized analog and digital VLSI hardware. Such a system would be suitable for mounting on mobile or remote platforms that cannot be tethered to a computer and for which the size, weight and power consumption of the components are critical factors. The problem of computing relative orientation is difficult for two reasons: one is the nonlinearity of the motion equations which must be solved; and the other is the difficulty of determining point correspondences between two images. It is the latter problem which entails the largest I/O and computational requirements and where specialized processors will offer the greatest benefit. Methods already exist for finding point correspondences in software which can be implemented on available general purpose digital image processors. However, the resulting systems are still too large, slow and power-hungry for many applications.

This research consists of two parts. The first, which is now essentially completed, was to develop the theoretical foundations of the relative orientation system. This involved developing and adapting algorithms for determining point correspondences and solving the motion equations which are both robust as well as simple enough to be easily implemented in hardware. It also involved determining the limits of system operation as a function of error in the correspondences, field of view, and image resolution. The second part of the research, which is currently underway, is to design, fabricate and test prototype chips for the specialized processors which will be used to find the point correspondences. Two separate processors are needed: one which computes a binary edge map from the input image data, and the other which determines translational offsets between patches of the edge maps from two different images. Fabrication of the prototype circuits will be done through MOSIS.

The design for the edge detection chip is based on a variation of the CCD circuit developed by Dr. Craig Keast for the focal-plane imaging and segmentation chip, which has been modified to implement the multiscale veto algorithm described in the previous progress report. An important problem to be solved, which was one of the major difficulties encountered by Dr. Keast, is to find an effective method for computing on chip small differences between neighboring pixel values with at least 6 bits of dynamic range. Three test chips containing one- and two-dimensional arrays, as well as their isolated subcircuits, have been fabricated and tested based on an initial design. These circuits did not reach the desired performance level of 6 bits of dynamic range, however, largely due to offsets in the absolute-value-of-difference circuits. An improved design which should reduce the effect of offsets is currently being studied, and a new set of test chips will soon be sent to MOSIS for fabrication. If the tests of these new chips are satisfactory, a full 32x32 array will be sent out for fabrication by the end of the year.

The second processor receives the binary edge maps produced by the edge detection chips and determines point matches using a block correlation method. Because the input is binary, the matching circuits require relatively little hardware and can be implemented with purely digital logic. On the other hand, because the computations are simple and do not require high precision, there may be a net benefit in terms of silicon area in using analog processing for parts of the circuit. In order to test this idea, a design incorporating analog elements has been formulated for the core of the matching circuit.
Chapter 1. Custom Integrated Circuits

A test chip based on this design will be fabricated and tested and the result will be compared with an entirely digital implementation.

Switched-Capacitor System for Merging Depth and Slope Estimates

Graduate student Mark Seidel, working with Professors John Wyatt and Thomas Knight, has been designing, fabricating, and testing a switched-capacitor chip to robustly compute the depth of a scene by merging depth estimates (possibly from a stereo algorithm) with surface slope estimates (possibly from a shape-from-shading algorithm). This multisensor image fusion design uses the least-squares properties of certain switched-capacitor systems to compute a dense depth map given noisy or sparse depth and slope input. A MOSIS Tinychip containing some subcircuits of a 1-D version was fabricated and tested; all of the subcircuits were operational. A final 1-D version of the network was designed, fabricated, and partially tested. The protection circuits test positively, as does the connectivity on-chip. The remainder of the testing will be completed before the end of the year.

1.2.3 Mixed Circuit/Device Simulation for CMOS/CCD Systems

Graduate student Mark Reichelt, under the supervision of Professor Jacob K. White, and in collaboration with Professor Andrew Lumsdaine of the University of Notre Dame, is working on algorithms for accelerating mixed circuit/device simulation, as this kind of analysis is required to properly simulate the performance of CCD-based analog vision circuits. Initially, they have focused on developing parallelizable algorithms for semiconductor device transient simulation, and are investigating waveform relaxation (WR), a natural candidate (This work is also partly funded by the ARPA VLSI contract N00014-91-J-1698).

Part of this work is focused on more theoretical aspects of applying parallel WR to device transient simulation. As the WR algorithm decomposes systems of differential equations into subsystems which can be solved independently, the algorithm has easily exploited parallelism. And when using WR on a parallel computer, it is likely that the different subsystems will be assigned to different processors. In that case, communication can be avoided, and overall computational efficiency improved, if each of the separated subsystems can be solved with independently determined timesteps. Using different timesteps for different subsystems implies that a multirate integration method has been used to solve the system, and there is no guarantee that this multirate method is stable. In the past year, we have been able to prove that certain multirate methods are A-stable for several model problems associated with semiconductor device simulation.

The more practical part of this work was directed toward accelerated waveform relaxation techniques for parallel semiconductor device simulation. We adapted both the generalized conjugate-direction method and a convolution successive over-relaxation (CSOR) algorithm to performing parallel semiconductor device simulation, and compared the two approaches using both a cluster of workstations and a 32-node i860-based hypercube. Both algorithms are competitive with standard algorithms on a single processor, but achieve nearly linear speed-up on the hypercube. Mark Reichelt finished his doctoral dissertation on the subject of convolution SOR and also received a best student paper award at the SIAM sponsored Copper Mountain Conference on Multigrid methods.

1.2.4 Hybrid Analog/Digital Vision Chips for Intelligent Vehicle Applications

Dr. Ichiro Masaki, a visiting scientist from General Motors, is studying the feasibility of applying hybrid analog/digital vision chip technologies to intelligent vehicle systems in collaboration with Professors Hae-Seung Lee, Charles Sodini, John Wyatt, Berthold Horn, and their students, David Martin and Steve Decker. The demand is increasing rapidly for developing intelligent vehicles that are integral to safe, efficient, and environmentally friendly traffic systems. The technical characteristics of the intelligent vehicle applications match well with the natures of hybrid analog/digital vision chip technologies. Intelligent vehicles need high-speed processing, low mass-production system costs, low power consumption, and small physical sizes. Large development costs will be compensated by the large market size and limited accuracy can be accepted. Examples of applications include on-vehicle systems such as a collision avoidance system and on-road systems including a traffic monitoring system.

A stereo vision system is being developed as the first case study. Installed in a vehicle, the stereo vision system will measure the distance from the vehicle to a vehicle in front of it. The goal is to keep a reasonable distance between the vehicles by controlling a vehicle’s speed automatically. The system includes two custom chips: a wide dynamic range imager chip and a hybrid analog/digital array processor chip. The intensity dynamic range of
typical conventional television cameras is large enough for in-door applications but does not cover various weather conditions on highways. Our imager chip will have enough dynamic range for vehicle applications. The array processor chip will carry out edge detection and stereo matching operations using both analog and digital circuits. The chip is expected to require less silicon area, lower power consumption, and lower mass-production cost compared to conventional digital microprocessors. Small-scale test chips are being developed for feasibility studies.

**Analog versus Digital Approaches to Machine Vision Hardware**

Recent technical progress in analog vision chips has stimulated comparisons between analog and digital processing schemes through case studies. The analog vision chips have demonstrated their small cost in silicon area and high speed characteristics for some applications. The goal of this project is to identify the capabilities and limitations of the analog chips and to find what application areas would be suitable for analog, digital, and hybrid chips, respectively.

In collaboration with Professors Charles Sodini, Berthold Horn, and John Wyatt, Dr. Ichiro Masaki and graduate student David R. Martin, under the supervision of Professor Hae-Seung Lee, are studying stereo vision algorithms. These algorithms are frequently used in various applications including intelligent vehicles and flexible manufacturing systems.

The chip implementation level comparison is underway. Currently a cell in a Multi-Instruction Multi-Data Array Processor architecture, implemented in both analog and digital, is being designed. Each processing element includes an arithmetic unit, a data flow control unit, and a system control unit which includes program memory. Several analog circuit issues for the accuracy of the cell's arithmetic unit have been resolved.

Our next goals include finalization of the chip design, layout using the Cadence design tools, fabrication of the chips, and performance comparisons based on both simulated and actual chips. The criteria for the comparisons will include cost in silicon area, processing speed, manufacturing and development costs, testability, and other related issues.

**1.2.5 Solutions of Theoretical Problems in Analog Hardware and Machine Vision**

**Neural Network Theory**

Working under the guidance of Professor John Wyatt, graduate student Ibrahim Elfadel has investigated a generalization of the effective energy function, first introduced by Peterson and Soderberg, and used it to solve constrained optimization problems within the ansatz of saddle-point methods in statistical physics. Mr. Elfadel has shown that within this paradigm, a smooth convex potential function is associated with each closed bounded constraint set. He has also shown that using the conjugate of the convex potentials, it is possible to derive, from the effective energy, a cost function that is a natural generalization of the analog Hopfield energy. Descent dynamics and deterministic annealing can then be used to find the global minimum of the original minimization problem. When the conjugate is hard to compute explicitly, we show that minimax dynamics can be used to find the extrema of the effective energy. This general effective-energy framework has allowed Mr. Elfadel to tie up different cost functions that were used previously to solve optimization problems involving local competition of the winner-take-all type. Moreover, he has used the same general framework to propose continuous dynamical systems for finding the fixed points of "rotor" neural networks.

**Reduced-Order Models for Interconnect Simulation**

To analyze the effects of packaging parasitics on circuit performance, these parasitics are frequently modeled with equivalent circuits which are then included in a circuit simulation. Complicated packages or interconnects can generate extremely large equivalent circuits which must then be simplified, or reduced, for efficient use in a circuit simulator. The most commonly used Padé, or moment-matching, approaches to performing this model order reduction can generate unstable equivalent circuits. To overcome the instability problems, Graduate Students Luis Miguel Silveira and Ibrahim Elfadel, under the guidance of Professor Jacob White, have been investigating a state-space algorithm for model-order reduction based on truncating a balanced system realization. This algorithm has the important characteristic that the reduced order model resulting from the truncation is guaranteed to be stable. The simulation of simple RLC circuits modeling transmission lines has demonstrated that the balanced-realization approach generates accu-
rate, stable reduced-order models, whereas the Padé method does not.

**Switched-Capacitor Network Theory**

Graduate student Mark Seidel, working with Professor John Wyatt, has been studying general networks of switches, capacitors, and sources. These networks are important in the implementation of computational and image processing problems formulated as minimization algorithms. Discrete-time system stability was proved for a large subclass of these networks using contraction mappings; this subclass includes many useful image processing networks, including the SC implementation of the coupled depth and slope network. The explanation of derived and inverted bounding expressions and estimates, along with the important limiting case of large N parameters, was accepted and presented at the European Conference on Circuit Theory and Design.

Under constant input conditions, the switched-capacitor networks can be shown to be generally equivalent to networks comprised of nonreciprocal linear passive resistive multiports and multiterminal elements. Each such element is associated with exactly one of the capacitors in the network. The transformation to these resistive networks results in a resistive network equivalent to the original SC network in the sense that all node voltages are the same, and all currents are equal to the corresponding average transferred charges per period. This equivalence indicates certain rules for the reverse problem of synthesizing a switched capacitor network from a resistive network. This synthesis can always be accomplished for networks of two-terminal resistors and lossy two-port transformers. It has also been shown that even if the resistive network to be implemented includes only the nonreciprocal passive resistive elements discussed earlier, clock phase assignment is sometimes impossible. A conference paper explaining this linear analysis and synthesis of SC networks was accepted and presented at the European Conference on Circuit Theory and Design.

Constant-input switched-capacitor network steady-state analysis and synthesis has been extended to two classes of networks containing certain nonlinear SC elements. This theoretical extension now allows for the analysis of realistic SC implementations, including the reverse-biased-diodes nonlinear parasitic capacitances associated with the MOS switches. More importantly, this extension also allows for implementing arbitrary networks of two-terminal voltage-controlled conductances, two-terminal positive linear resistors, and lossy two-port transformers. An example of this kind of network is the one- and two-dimensional resistive fuse grid. The new networks explain certain existing nonlinear SC networks, as well as offering a synthesis method leading to a resistive fuse network with significant performance improvement. Furthermore, the nonlinear elements are dynamically controlled by two global wires, thereby making continuation method implementation feasible. A conference paper presenting the nonlinear synthesis results was presented at the Midwest Symposium on Circuits and Systems.

First and second drafts of the majority of Mr. Seidel's thesis document were completed. This portion of the thesis covers bounds and estimates theory and develops linear and nonlinear analysis and synthesis theory. The thesis will be completed before December 1994.

### 1.2.6 Publications

#### Journal Articles


- **Elfadel, I.** "Convex Potentials and their Conjugates in Analog Mean-Field Optimization." Submitted to *Neural Computation*.


Chapter 1. Custom Integrated Circuits


Book Chapters


Conference Papers


Elfadel, I.M. "Integration of Shape Deformation and Motion." Submitted to the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '94).


Chapter 1. Custom Integrated Circuits


Patents


Theses


Publicity


Internal Publications


Abstracts for Talks


1.3 Parallel Algorithms for Device Simulation

Sponsors

Advanced Research Projects Agency/ National Science Foundation
Grant MIP 91-17724
IBM Corporation
U.S. Navy - Office of Naval Research
Grant N00014-91-J-1698

Project Staff

Dr. Andrew Lumsdaine, Khalid Rahmat, Mark W. Reichelt, Professor Jacob K. White, Professor Jonathan Allen

The growing importance of mixed circuit/device simulation, its enormous computational expense, as well as the increasing availability of parallel computers, have made the development of specialized, easily parallelized, algorithms for transient simulation of MOS devices necessary. In earlier work on the Waveform Overrelaxation Device Simulator
Chapter 1. Custom Integrated Circuits

(WOMS) program, the easily parallelized waveform relaxation (WR) algorithm was shown to be a computationally efficient approach to device transient simulation even on a serial machine. The WR algorithm, though, typically requires hundreds of iterations to achieve an accurate solution.

To use WORDS in a mixed circuit/device simulator, we have been investigating ways of making the program more robust and efficient. We determined how to compute the terminal currents accurately using different timepoints at different mesh nodes. We also improved the timestep selection procedure by determining how to refine the timesteps as WR iterations proceed and reducing the total computation by as much as a factor of two by using only a few coarse timesteps in early iterations. The more accurate, electric field dependent mobility model was also implemented. Recent work on theoretical aspects of these methods have answered several long-standing questions about multirate stability.12

We found experimental evidence that WR using standard overrelaxation acceleration can produce oscillatory results, and are investigating methods for eliminating this phenomenon. A frequency-dependent overrelaxation algorithm using lowpass filtering was developed,13 as well as a waveform conjugate-direction approach.14 Experimental results indicate that both approaches reduces the number of waveform iterations required by more than a factor of seven. Finally, experimental results using a parallel computer show that although the accelerated WR methods are only competitive with the best of the standard algorithms for device transient simulation on a serial machine, the WR algorithms are substantially faster on a parallel machine.15

1.4 Numerical Simulation of Short Channel MOS Devices

Sponsors
IBM Corporation
National Science Foundation
Fellowship MIP 89-58764
U.S. Navy
Contract N00174-92-Q-0133

Project Staff
Khalid Rahmat, Professor Dimitri A. Antoniadas,
Professor Jacob K. White

The model used in conventional device simulation programs is based on the drift-diffusion model of electron transport, and this model does not accurately predict the field distribution near the drain in small geometry devices. This is of particular importance for predicting oxide breakdown due to penetration by "hot" electrons. There are two approaches for more accurately computing the electric fields in MOS devices, one is based on adding an energy equation to the drift-diffusion model and the second is based on direct solution of Boltzman's equation. In the first approach, an energy balance equation is solved along with the drift-diffusion equations so that the electron temperatures are computed accurately. This combined system is numerically less tame than the standard approach, and must be solved carefully. We have developed a two-dimensional device simulator in which an energy balance equation is solved for electron temperature along with the usual drift-diffusion equations. The program avoids temperature instabilities produced by previous discretization approaches through the use of a careful application of exponential-fitting to the energy equation. Drain currents for silicon MOSFETs predicted by the simulator, using one set


Chapter 1. Custom Integrated Circuits

of model parameters, match well with experimental data for devices over a range of channel lengths from 90 μm to 0.16μm. Also, a method to compute substrate current has been derived which uses the electron temperature provided by the simulator. The computed substrate currents match well with measured data, for the regime above subthreshold, for MOSFET's with channel lengths as short as 0.16μm.16

The limited ability to accurately predict hot-electron effects from an energy-balance based simulator has led to our current investigation. We are now trying to solve the full Boltzmann equation, using a spherical harmonics based approach, but only in portions of the MOS device which are likely to create hot carriers.17

1.5 Coupled Simulation Algorithms for Microelectromechanical CAD (MEMCAD)

Sponsors
Federal Bureau of Investigation
Contract J-FBI-92-196
National Science Foundation
Fellowship MIP 88-58764

Project Staff
Xuejun Cai, Mattan Kamon, Dr. He Yie, Professor Martin A. Schmidt, Professor Stephen D. Senturia, Professor Jacob K. White

High fabrication costs and increasing microsensor complexity is making computer simulation of the realistic geometries necessary, both to investigate design alternatives and to perform verification before fabrication. At MIT, we are developing a microelectromechanical computer-aided design (MEMCAD) system to make it possible for microsensor designers to easily perform realistic simulations. Carefully selected commercial software packages have been linked with specialized database and numerical programs to allow a designer to easily enter a three-dimensional microsensor geometry and quickly perform both mechanical and electrical analysis. The system currently performs electromechanical analyses, such as calculating the capacitance versus pressure (or force) curve for both a square diaphragm deformed by a differential pressure and can be used to calculate levitation forces in structures as complicated as a comb drive.18

To support design of electromechanical structures, we are currently investigating two approaches to combining finite-element mechanical analysis with multipole-accelerated electrostatic analysis. The first method is the obvious relaxation algorithm19 and the second method is a more sophisticated surface/Newton generalized conjugate-residual scheme.20 By comparing the two methods, we have demonstrated both theoretically and by example that our surface/Newton-GCR algorithm is faster and more robust than the simpler relaxation scheme.21

1.6 Numerical Techniques for Simulating Josephson Junction Arrays

Sponsors

Advanced Research Projects Agency
Consortium for Superconducting Electronics
Contract MDA972-90-C-0021
National Defense Science and Engineering
Graduate Fellowship
National Science Foundation
Fellowship MIP 88-58764

Project Staff

Joel R. Phillips, Dr. Herre S.J. van der Zant, Professor Terry P. Orlando, Professor Jacob K. White

Vortices play a central role in determining the static and dynamic properties of two-dimensional (2D) superconductors. Artificially fabricated networks of superconducting islands weakly coupled by Josephson junctions are model systems for studying the behavior of vortices. Through simulation, we have discovered that the static properties of vortices in an array of Josephson junctions can be significantly influenced by magnetic fields induced by the vortex currents. The energy barrier for vortex motion is enhanced, nearly doubling for penetration depths on the order of a cell size. Moreover, we have found that correct calculation of the vortex current distribution, magnetic moment, and lower critical field require modeling mutual inductance interactions between all cell pairs in the array. To make numerical simulation of the system with all inductive effects computationally feasible, a novel FFT-accelerated integral equation solver was derived. This algorithm is sufficiently efficient to allow study of large (500 x 500 cells) arrays.22

The development of the FFT-accelerated algorithm for inductance effects in Josephson Junction arrays led to a more general result useful for packaging analysis. For the important special case of conductors over non-ideal ground planes, extremely fast FFT-accelerated inductance extraction methods have been developed. The FFT technique is so fast compared to its direct rival method that it reduces analysis time from hours to seconds even for relatively simple problems.23

1.7 Efficient 3-D Interconnect Analysis

Sponsors

Digital Equipment Corporation
IBM Corporation
National Science Foundation
Fellowship MIP 88-58764
U.S. Navy - Office of Naval Research
Contract N00014-91-J-1698

Project Staff

Mattan Kamon, Songmin Kim, Keith S. Nabor, Joel R. Phillips, Professor Jacob K. White

We have developed multipole-accelerated algorithms for computing capacitances and inductances of complicated 3-D geometries and implemented these algorithms in the programs FASTCAP and FASTHENRY. The methods are accelerations of the boundary-element or method-of-moments techniques for solving the integral equations associated with the multiconductor capacitance or inductance extraction problem. Boundary-element methods become slow when a large number of elements are used because they lead to dense matrix problems which are typically solved with some form of Gaussian elimination. This implies that the computation grows as $n^2$, where $n$ is the number of panels or tiles needed to accurately discretize the conductor surface charges. Our new algorithms, which use generalized conjugate residual iterative algorithms with a multipole approximation to compute the iterates, reduces the complexity so that accurate multiconductor capacitance and inductance calculations grow nearly as $nm$ where $m$ is the number of conductors. For practical problems which require as many as 10,000 panels or filaments, FASTCAP and FASTHENRY are more than two orders of magnitude faster than standard boundary-element based programs.24 Manuals and source code for


Chapter 1. Custom Integrated Circuits

FASTCAP and FASTHENRY are available directly from MIT.

1.8 Adaptive Gridding Techniques for Multipole-Accelerated Solution of Integral Equations

Sponsors
National Science Foundation
Semiconductor Research Corporation
Contract SRC 93-SJ-360

Project Staff
Michael Chou, F. Thomas Korsmeyer, Professor Jacob White

Finding computationally efficient numerical techniques for simulation of three dimensional structures has been an important research topic in almost every engineering domain. Surprisingly, the most numerically intractable problem across these various disciplines can be reduced to the problem of solving a three-dimensional Laplace problem. Such problems are often referred to as potential problems, and application examples include electrostatic analysis of sensors and actuators; electro- and magneto-quasistatic analysis of integrated circuit interconnect and packaging; and potential flow based analysis of wave-ocean structure interaction.

The recent development of extremely fast multipole-accelerated iterative algorithms for solving potential problems has renewed interest in integral equation formulations. In such methods, the fast multipole algorithm is used to implicitly construct a sparse representation of the dense matrix associated with a standard integral equation formulation, and this implicit representation is used to quickly compute the matrix-vector products required in a Krylov subspace based iterative method like GMRES. That the multipole-accelerated approach leads to practical engineering analysis tools has been demonstrated for electrostatic analysis of integrated circuit interconnect and packaging problems, as well as general Laplace problems.

We are extending our own work on multipole-accelerated iterative methods for potential problems to address automatic error control, which is the key difficulty which prevents such techniques from finding broader use in engineering applications. In particular, we are interested in first investigating the straightforward problem of how best to include higher-order boundary or volume elements. We are then planning to use such a result, along with other error estimation procedures, to perform automatic element refinement. The spatial hierarchy already introduced by the multipole algorithm and our use of a locally preconditioned iterative scheme make it possible to perform very efficient local element refinement and may allow for an efficient variable element order scheme. Because we have been using realistic examples from a wide variety of engineering disciplines, we can determine the robustness of our procedures.

1.9 Circuit Simulation Algorithms

Sponsors
IBM Corporation
National Science Foundation Fellowship MIP 88-58764
Grant MIP 91-17724

Project Staff
Luis M. Silveira, Steven B. Leeb, Professor Jacob K. White

A challenging problem in the area of analog circuits is the simulation of clocked analog circuits like switching filters, switching power supplies, and phase-locked loops. These circuits are computationally expensive to simulate using conventional techniques because these kinds of circuits are all clocked at a frequency whose period is orders of magnitude smaller than the time interval of interest to the designer. To construct such a long time solution, a program like SPICE or ASTAP must calculate the behavior of the circuit for many high frequency clock cycles. The basic approach to making simulation of these circuits more efficient is...
to exploit only the property that the behavior of such a circuit in a given high frequency clock cycle is similar, but not identical, to the behavior in the preceding and following cycles. Therefore, by accurately computing the solution over a few selected cycles, an accurate long time solution can be constructed. Such approaches are known as "envelope-following" algorithms.

In our recent work, we have been trying to make the envelope-following algorithm more robust and efficient by exploiting the fact that the envelope of "quasi-algebraic" components in the solution vector do not need to be computed. An automatic method for determining the quasi-algebraic solution components has been derived; experimental results demonstrate that this modified method reduces the number of computed clock cycles when applied to simulating closed-loop switching power converters.28

1.10 Techniques for Embedded System Design, Formal Verification, and Synthesis for Low Power Dissipation

Sponsors

Advanced Research Projects Agency/U.S. Navy - Office of Naval Research
Grant N00014-91-J-1698
National Science Foundation
Fellowship MIP 92-58376

Project Staff
Professor Srinivas Devadas, Mazhar M. Alidina, Vishal L. Bhagwati, Stan Y. Liao, José C. Monteiro, W. John Rinderknecht, Amelia H. Shen

1.10.1 Introduction

Complex digital systems have always combined application-specific integrated circuits (ASICs) and reprogrammable components, such as off-the-shelf microprocessors, where the application-specific component is often referred to as the hardware component, and the program running on the reprogrammable component is referred to as the software component. More recently, however, to achieve higher performance and density, as well as lower power dissipation and cost, designers have turned to more complicated packaging and are much more careful about distributing functionality between ASICs and reprogrammable components. Examples of such systems include embedded control systems for printing and plotting, telecommunications systems, and other signal processing systems.

While most digital functions can be implemented by software programs, a major reason for building dedicated ASICs is to satisfy performance constraints. Specialized integrated circuits are needed to assist the reprogrammable component for certain performance-critical tasks. In some cases, reprogrammable processors can meet performance constraints. However, a second important reason for using ASICs is that ASICs have significantly lower power dissipation than reprogrammable processors. Thus, system-wide power dissipation constraints may be met by using dedicated ASICs. However, when using standard printed circuit boards to


305
provide the interconnect between ASICs and programmable components, some of the performance and power dissipation advantages of using ASICs may be eliminated. For this reason, the use of novel interconnect structures like multichip modules (MSMs) and three-dimensional packaging is increasing.

Our focus is on solving several critical problems in the design and verification of these large-scale digital systems and targeting maximal performance and minimal power dissipation. Currently, large-scale digital systems exploit sophisticated packaging methods. However, in the near future most of the hardware for such systems will be integrated on a single chip. In this context, we are addressing the problems described in the following sections.

1.10.2 Embedded System Design Aids

Currently, we lack the CAD tools that would allow for the quick and efficient design of a high level mixed hardware/software system or an embedded system. Most ASICs that correspond to application-specific processors in mixed hardware/software systems contain a ROM which occupies a large fraction of the area of the chip and contributes significantly to the power dissipation. The size of the ROM is determined by the amount of code that it must store, which, in turn, is generated by a compilation process from a higher level software specification. Separate ROM chips may also be associated with the reprogrammable processors in the system.

Compilation methods play a very important part in embedded system design. While compilation techniques are well-developed for purely software systems, a direct usage of these techniques to generate code in mixed hardware/software systems is not recommended. Typically, these techniques attempt to maximize system performance, i.e., the speed at which the threads of code are executed without regard to code density. In a mixed hardware/software system during compilation, one should maximize code density and minimize ROM size while meeting a system-level performance constraint. Maximizing code density will result in improved area and power dissipation.

At present, global, efficient methods to maximize code density during compilation have not been developed. It is possible to formulate this problem as one of data compression allowing the use of modified text compression algorithms to minimize code size. Graduate student Stan Liao is developing global and efficient optimization methods that will result in minimally-sized ROMs after compilation, improving the area and power dissipation characteristics of the system.

1.10.3 Formal Verification

It is critical to verify that the circuits obtained after application of sophisticated synthesis and optimization methods to the initial specification are error-free. Formal verification methods can guarantee correctness of an implementation with respect to a specification, even if the specification and implementation are described at vastly different levels of abstraction.

Most verification problems in VLSI design require the use of Boolean manipulation. Graduate student Amelia Shen has developed a representation of Boolean functions termed the Free Boolean Diagram (FBD) that has proven to be superior to other graph representations of Boolean functions. A FBD package capable of manipulating complex Boolean functions has been developed. This package serves as an equivalence checking engine for combinational circuits and is used in our behavioral verification methodology.

Our work on the behavioral verification of ASICs has been successful in checking the equivalence of descriptions at vastly different abstraction levels, such as a dependency graph of computations and a logic-level circuit. In the past year, graduate student Vishal Bhagwati has succeeded in developing a formal methodology for checking the equivalence of pipelined microprocessors against unpipelined specifications. The specification is the instruction set of the microprocessor with respect to which the correctness property is to be verified. A relationship, namely the $\beta$-relation, is established between the input/output behavior of the implementation and specification. The relationship corresponds to changes in the input/output behavior that result from pipelining, and takes into account data hazards and control transfer instructions that modify pipelined execution. The correctness requirement is that the $\beta$-relation hold between the implementation and specification. Symbolic simulation of the specification and implementation is used to verify their functional equivalence. The pipelined and unpipelined microprocessor can be characterized as definite machines (i.e., a machine in which for some constant $k$, the output of the machine depends only on the last $k$ inputs) for verification purposes. It can be shown that only a small number of cycles, rather than exhaustive state transition graph traversal and state enumeration, must be simulated for each machine to verify whether the implementation is in $\beta$-relation with the specification.
Chapter 1. Custom Integrated Circuits

The verification problems associated with mixed hardware/software systems are much more difficult to solve than those encountered in ASICs. While simulation is an integral part of verification methodology, it can be very expensive and time-consuming. Formal verification methods for such systems are thus very attractive. The next challenge is to develop a verification methodology for checking the equivalence of a software description of the functionality of the entire system against a mixed hardware/software implementation.

1.10.4 Logic Synthesis for Low Power

For many consumer electronic applications, low average power dissipation over a system-wide basis is desirable, and, for certain special applications, low power dissipation is of critical importance. For applications such as personal communication systems like hand-held mobile telephones, low-power dissipation may be the tightest constraint in the design. More generally, with the increasing scale of integration, we believe that power dissipation will assume greater importance, especially in multi-chip modules (MCMs) where heat dissipation is one of the biggest problems.

The average power dissipation of a circuit, like its area or speed, may be significantly improved by changing its architecture or technology. But once these architectural or technological improvements have been made, the switching of the logic and leakage currents will ultimately determine its power dissipation. Traditionally, logic synthesis has been applied to improve the area or speed of a circuit. We are developing methods for fast, accurate logic-level power estimation methods that will help development of architectural and logic-level synthesis strategies that target low power. We believe that the development of logic optimization algorithms that directly target low power dissipation poses the next major challenge for logic synthesis.

Switching activity is the primary cause of power dissipation in CMOS circuits. Accurate, average switching activity estimation for sequential circuits is considerably more difficult than for combinational circuits, because the probability of the circuit being in each of its possible states has to be calculated. The Chapman-Kolmogorov equations can be used to accurately estimate the power dissipation of sequential circuits by computing the exact state probabilities in steady state. However, the Chapman-Kolmogorov method requires the solution of a linear system of equations of size \(2N\), where \(N\) is the number of flip-flops in the machine.

Graduate student José Monteiro has developed a computationally efficient scheme to approximate average switching activity which requires the solution of a non-linear system of equations of size \(N\), where the variables correspond to state line probabilities. The approximation method is within three percent of the exact method, but is up to 100 times faster for large circuits.

Graduate student Mazhar M. Alidina has developed a powerful sequential logic optimization method that is based on (1) selectively precomputing the output logic values of the circuit one clock cycle before the values are required and (2) using the precomputed values to reduce internal switching activity in the succeeding clock cycle.

The primary optimization step is the synthesis of the precomputation logic, which computes the output values for a subset of input conditions. If the output values can be precomputed, the original logic circuit can be "turned off" in the next clock cycle and will not have any switching activity. The size of the precomputation logic determines the power dissipation reduction, area increase, and delay increase relative to the original circuit.

Given a logic-level sequential circuit, we are developing automatic methods of synthesizing precomputation logic to achieve maximal reductions in power dissipation.

1.10.5 Publications

Journal Articles


Chapter 1. Custom Integrated Circuits


Conference Papers


Chapter 2. Computer-Integrated Design and Manufacture of Integrated Circuits

2.1 Introduction

Sponsor
U.S. Navy
Contract N00174-92-Q-0133

Computer-Aided Fabrication Environment (CAFE) is a software system being developed at MIT for use in the fabrication of integrated circuits and microstructures. The distinguishing feature of CAFE is that it can be used in all phases of process design, development, planning, and manufacturing of integrated circuit wafers. CAFE presently provides day to day support to research and production facilities at MIT with both flexible and standard product capabilities. This manufacturing software system is unique in its development of a process flow representation which is integrated into actual fabrication operations. CAFE provides a platform for work in several active research areas, including "technology" (process and device) computer-aided design (TCAD), process modeling, manufacturing quality control, and scheduling.

2.1.1 Architecture

The CAFE architecture is a computer integrated manufacturing (CIM) framework for the deployment and integration of integrated circuits (IC) circuit and process design and manufacturing software. CAFE uses an object-oriented database model which is implemented in a layered manner on top of a relational database. Our database schema is based on GESTALT, an object-oriented, extensible data model. GESTALT is a layer of abstraction which maps user defined objects onto existing database systems (e.g., a relational DBMS) and shields application programs from the details of the underlying database. The integration architecture includes the conceptual schema and models used to represent the IC manufacturing domain in CAFE and the user and programmatic interfaces to the various applications. Two important CAFE applications relate to process simulation and actual wafer fabrication to the same process flow representation.

2.1.2 CAFE Applications

The fabrication of wafers with a process represented as a process flow representation (PFR) involves several steps. A suitable PFR for the specific process must be created and installed. Wafer lots must be created and associated with this specific PFR. These lots must then be "started" to create a task data structure which is isomorphic to the hierarchical structure of the PFR.

At this point, actual machine operations can be scheduled and reservations made for both machines and operators. Finally, the machine operations can be performed, instructions given to the operator and machines, and data collected from the operator or machine and entered into the database.

Current work involves the implementation of hierarchical scheduling, development of factory design tools, real time process control using the PFR, and integration of a TCAD system using the PFR.

2.2 Principal Objectives

There were two principal objectives: (1) installation of CAFE at MIT Lincoln Laboratory and (2) development of a microsystems factory representation (MFR).
Chapter 2. Computer-Integrated Design and Manufacture of Integrated Circuits

2.2.1 MIT Lincoln Laboratory Installation

MIT Lincoln Laboratory has selected CAFE as the CIM system for use in a new IC processing facility. CAFE is now in operation at Lincoln Laboratory, and multiple lots are being processed daily via PFR-based fabrication. Lincoln Laboratory is committed to having all the wafers processed in this facility with PFR-based fabrication.

2.2.2 Microsystems Factory Representation

We have embarked on a new research initiative to develop a microsystems factory representation (MFR). The goal of the MFR is to be a powerful, common language with which semiconductor factories can be described unambiguously. MFR will be used to organize static data pertaining to resources of production systems. That is, it will describe all attributes of the factory that do not change very often. The MFR, the state (and its evolution), the PFR, the demand, and the scheduler will provide all data necessary for simulations, analytical calculations of performance measures, and for the implementation of factory controllers. A system created from these items can represent either a real factory or a simulation of a factory. The only difference is in the state and its dynamics. In a real factory, events occur either as a result of decisions made by the scheduler (controllable events such as loading lots) or at random (uncontrollable events such as machine failures). In a simulation, the random events are synthesized by means of a random number generator.

2.3 Principal Accomplishments

2.3.1 CAFE Installation at MIT Lincoln Laboratory and Expanded PFR-Based IC Fabrication

The CAFE architectural framework at Lincoln Laboratory supports a wide variety of software modules including both development tools and on-line applications. The key components of the CAFE architecture are the data model and database schema, process flow representation, user interface, and application programming and database interfaces. The approach of the MIT PFR is to use the same representation for multiple purposes including simulation, scheduling, process control, and fabrication. The PFR provides an extensible framework for knowledge about process steps including instructions to operators and equipment, scheduling requirements, changes effected to the wafer product, and physical process model parameters.

A primary reason for this selection is that CAFE supports PFR-based fabrication and has demonstrated process simulation driven by the same PFR. Other reasons include CAFE's use of a robust commercial database, provision of an ASCII process definition, flexible process descriptions, handling of splits and joins, and satisfactory user interface support. In conjunction with this selection, there have been substantial improvements in CAFE. The speed of operation has been dramatically increased, and a number of significant new applications have been developed. An important enhancement was the ability to execute applications in the background and at slack times in the evening or early morning. An important new application developed displays the queue of lots available for processing at a machine or set of machines.

A major task was to write a large number of unit processes into the PFR. A new application was developed to enable production of detailed reports from a PFR-based encoding without actually installing the process. Changes and enhancements to the PFR and utility functions which facilitated the expression of the required processing will be described. In addition, a scheme for sharing of installed process flows has been provided.

2.3.2 Microsystems Factory Representation

The microsystems factory representation (MFR) is conceived to be a powerful, common language with which semiconductor factories can be described unambiguously. The MFR will specify the organization of static data pertaining to resources of production systems; that is, it will describe all attributes of the factory that do not change very often. The MFR will also specify the static data itself and the organization of the dynamic data.

The MFR is designed to be used in conjunction with a set of representations of other semiconductor factory data. They include:

1. The state of the factory, a list of all relevant dynamic information. The evolution of the state provides a history of the factory's activities.

2. The process flow representation (PFR), a representation of the process that is used to produce a given product.

3. The demand, all information available about future requirements including known specific orders, long range forecasts, and any statistical measures of uncertainty.

4. The scheduler and the set of rules by which the factory is run. The scheduler takes infor-
Chapter 2. Computer-Integrated Design and Manufacture of Integrated Circuits

Information from the current state, MFR, PFR, and demand, and makes a decision about which operation to do next, where to move lots, etc. This decision, when implemented, determines the next value of the state.

The MFR, state (and its evolution), PFR, demand, and scheduler will provide all necessary data for simulations, analytical calculations of performance measures, and implementation of factory controllers. When a system is created from these items, the state evolves; reports are generated that summarize the behavior of the state over time. These reports include such performance measures as average in-process inventory, mean cycle time, machine utilization, etc.

The system created in this manner can represent either a real factory or a simulation of a factory. The only difference is in the state and its dynamics. In a real factory, events occur either as a result of decisions made by the scheduler (controllable events such as loading lots) or at random (uncontrollable events such as machine failures). In a simulation, the random events are synthesized by means of a random number generator.

One important goal of our approach to scheduling and simulation is that the MFR and PFR, and their use, are designed to be the same, whether they are part of a real-time factory management system, or a simulation of a factory. Future demand information can also be used in the same way, whereas current orders may have to be synthesized in a simulation.

2.4 Process Flow Representation

Project Staff

Professor Donald E. Troxel, Michael B. McIlrath

The objective of this project is the development of a process flow representation (PFR) for the integration of technology (process and device) CAD (TCAD) into IC semiconductor computer integrated manufacturing (CIM). The effort has proceeded on three major fronts: fundamental modeling, application to process design and fabrication at MIT, and cooperative work with industry to define and test a process representation standard.

For high performance computing systems and other advanced technology, concurrency in the design of the product, manufacturing process, and factory is crucial. The goal is to achieve fully integrated design and manufacture, in which the boundary between design and manufacturing domains is eliminated. In particular, information from the manufacturing floor is continuously available from the earliest stages of process and device design onward. Conversely, the manufacturing process, developed concurrently with the product, continues to undergo design via improvement and modification while in production. Computer integrated design and manufacture (CIDM), therefore, requires a coherent manufacturing process representation which is capable of storing information from a variety of different knowledge domains and disciplines as well as supporting access to this information in a consistent manner. We believe that our general semiconductor process modeling framework organizes the complexity of this interrelated information and puts our PFR on a sound footing by giving it clear semantics.

A high level conceptual model for describing and understanding semiconductor manufacturing processing is a crucial element of both the CIDM research program, and of software frameworks for TCAD and CIM, including the MIT Computer Aided Fabrication Environment (CAFE). Initially a "two-stage" generic process step model was used, which described processing steps in terms of two independent components: (1) an equipment-dependent, wafer-independent stage, which maps equipment settings to physical processing environments, and (2) an equipment-independent, wafer-dependent stage, which relates physical environments to changes in the input wafers. Driven by the needs of process control and optimization research, including sophisticated modeling, design, and experimental model verification, our fundamental conceptual process model has evolved from the two-stage generic process model into one which is part of a more general process modeling framework, in which the earlier two-stage model is a special case. Our approach to process representation for both TCAD and CIM is based on this general modeling framework for semiconductor processing. In this framework, state information (e.g., wafer, environment, and equipment state), and models, or transformations, that describe relationships between state descriptions, are formally identified and described. The purpose of this comprehensive framework is to enable an effective representation that can be used throughout the IC semiconductor process life-cycle, from early conception and design phases through fabrication and maintenance.

In the MIT CAFE system, the PFR is expressed in a textual (ASCII) format and then converted into Gestalt objects and loaded into the CAFE database. The textual language of the PFR is extensible, so that it can flexibly accommodate changes and extensions to both the underlying modeling methodology and the requirements of specific applications. The object-oriented nature of
the Gestalt database interface enables the convenient evolutionary development of CAFE software applications built around the PFR. The PFR allows process step descriptions to be "underdetermined" for example, by expressing only the wafer-state change, making it possible to develop a process incrementally with increasing degrees of detail. In addition to expressing the fundamental concepts of wafer transformation within individual process steps, the PFR supports both hierarchical and parameter abstraction and embedded computation, thereby providing support for modular process design and development. Processes expressed in the PFR can be simulated using a variety of technology CAD tools; PFR extensibility allows the incorporation of both simulator-dependent and simulator-independent information. A simulation manager application uses the appropriate information in the PFR along with knowledge of specific simulators to invoke simulation tools and maintain a simulation state.

Extensions to the textual PFR and to the corresponding database schema support specification of individual wafers or wafersets (sublots, or splits) within a lot for processing, and specification of engineering holds at a particular step in the process sequence. An engineering hold might be planned, for example, as part of the fabrication specification for a process which is currently under experimental development. In conjunction with developments in CAFE that support dynamic modification of processes in the fabrication line, the "hold" capability in the PFR can also be used in response to unexpected events that occur during fabrication. With these manufacturing extensions, the PFR is used in the CAFE system at the MIT Integrated Circuits Laboratory (ICL) for the fabrication processes supported by the ICL technical staff, with all processing operations driven by CAFE from the PFR stored in the database.

Work has also proceeded on the definition of a standard semiconductor process representation (SPR) for use in both industry and research fabs, in association with the industry CAD Framework Initiative (CFI) Technical Subcommittee on Technology CAD in defining broadly applicable TCAD framework standards. Michael McIlrath is currently the chair of the working group on SPR. Based largely on the conceptual process model described above, a high level information model of semiconductor processes is being developed in conjunction with Motorola, IBM, and Texas Instruments, and other industry and university representatives. A prototype SPR implementation and programming interface has been developed and employed in an experimental TCAD framework, which included an implementation of the semiconductor wafer representation (SWR) and tools developed by industry and other university researchers.

2.5 Technology CAD Framework

Project Staff
Michael B. McIlrath

Contemporary large-scale software system engineering emphasizes frameworks, wherein common structure and interface specifications enable both current and future software components to be integrated in a flexible and modular way. Software may be roughly divided into tools, such as a simulator, which perform some part of an application task, and services, such as a database, which provide some necessary support capability used by various tools. With framework standards, reusable, interchangeable software components from various suppliers may be deployed in systems which comply with the standard. In the broad sense, a framework standard specifies:

1. data representations for the objects of discourse in the application domain and their semantics, and programmatic interfaces to those representations, and

2. architecture; that is, interactions among software components (tools and services), and how tools fit together to perform application tasks for the user.

Frameworks have been particularly successful in the development of electronic circuit CAD systems. Standards for CAD frameworks are currently being established by the CAD Framework Initiative (CFI), a broad organization of vendor and user companies which has now expanded its scope to include technology (process and device) CAD. A technology CAD (TCAD) framework standard includes programming representations for the fundamental objects of process and device CAD: the physical structures on the wafer, the manufacturing process, and the structure and behavior of the resulting devices. A TCAD framework standard also specifies how application software is structured to use these representations and the underlying software services in process and device design and simulation activities.

At present, development of TCAD framework architectural standards has been limited to (1) specifications of implementation and interfaces for the wafer and process representations and (2) use of domain-independent framework standards, such as intertool communication (ITC), user interface (UI), and extension language specifications. We are investi-
gating higher-level architectural issues, such as the interrelationships between the framework data representations and the connection of compliant tools to achieve end-user design objectives. We are also looking at the larger questions of the relationship between frameworks for different related domains (e.g., circuit CAD and TCAD) and integration of design frameworks into frameworks for CIM.

2.6 Publications

Internal Publications


Meetings and Presentations

McIlrath, M.B. Chair, SPR Working Group Meeting, Waltham, Massachusetts, April 1993.


McIlrath, M.B. Session Chair at the SRC/ARPA CIM-IC Workshop, Carnegie Mellon University, Pittsburgh, Pennsylvania, August 24-26, 1993.


Thesis

Section 2 Digital Signal Processing

Chapter 1 Digital Signal Processing Research Program

Chapter 2 Advanced Telecommunications and Signal Processing Program

Chapter 3 Combined Source and Channel Coding for High-Definition Television
Chapter 1. Digital Signal Processing Research Program

1.1 Introduction

The field of digital signal processing grew out of the flexibility afforded by the use of digital computers in implementing signal processing algorithms and systems. It has since broadened into the use of a variety of both digital and analog technologies, spanning a broad range of applications, bandwidths, and realizations. The Digital Signal Processing Group carries out research on algorithms for signal processing and their applications. Current application areas of interest include signal enhancement, active noise cancellation, processing of speech, music and underwater acoustic signals, advanced beam forming for radar and sonar systems, and signal coding and transmission.

In some of our recent work, we have developed new methods for signal enhancement and noise cancellation with single or multisensor measurements. We have also been developing new methods for representing and analyzing fractal signals. This class of signals arises in a wide variety of physical environments and also has potential application in problems involving signal design. We are also exploring potential uses of nonlinear dynamics and chaos theory for signal design and analysis.

In other research, we are investigating the application of signal and array processing to ocean and structural acoustics and geophysics. These problems require the combination of digital signal processing tools with a knowledge of wave propagation to develop systems for short-time spectral analysis, wave number spectrum estimation, source localization, and matched field processing. We emphasize the use of real-world data from laboratory and field experiments such as the Heard Island Experiment for Acoustic Monitoring of Global Warming and several Arctic acoustic experiments conducted on the polar ice cap.

Our group is also involved in research on new signal processing techniques for multiuser wireless communication systems. The intended applications range from mobile radio networks to indoor personal wireless systems to digital audio and television broadcast systems. Of special interest are broadband spread spectrum systems and code-division multiple access networks both for conventional and secure transmission.

Much of our work involves close collaboration with the Woods Hole Oceanographic Institution, MIT Lincoln Laboratory, and a number of high technology companies in the Boston area.

---

1 Bolt, Beranek, and Newman, Inc., Cambridge, Massachusetts.
2 Associate Professor, Boston University, College of Engineering, Boston, Massachusetts.
3 Department of Electrical Engineering, Systems Division, Faculty of Engineering, Tel-Aviv University, Israel; adjunct scientist, Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts.
1.2 Active Noise Cancellation in Automobiles

**Sponsor**
Bose Corporation

**Project Staff**
Stanley W. Brown, Professor Alan V. Oppenheim

A fundamental goal of this research is to determine the feasibility of real-time active noise cancellation within the interior of an automobile. Active noise cancellation in a volume entails measuring the sound field at a finite number of microphone locations and using predictable characteristics of the acoustic noise to generate and transmit canceling signals from one or more speakers. The goal is to attenuate the noise at the sensors through destructive interference.

Using computer simulations on recorded road noise, we will assess the effectiveness of known algorithms, such as the estimate-maximize (EM), least-mean square (LMS), and optimal control approaches, in attenuating acoustic noise at a point within the car. An autoregressive-moving average (ARMA) model of the measured transfer function from car speaker to noise sensor represents the effect of car acoustics on the canceling signal.

1.3 Single Mode Excitation in the Shallow Water Acoustic Channel

**Sponsor**
MIT-Woods Hole Oceanographic Institution
Joint Graduate Program in Oceanographic Engineering

**Project Staff**
John R. Buck, Professor Alan V. Oppenheim, Dr. Josko Catipovic

Under appropriate conditions, the ocean acts as an acoustic waveguide, with sound propagation occurring almost entirely through a discrete set of normal modes. Shallow water is one such condition. The goal of this research is to use a vertical array to excite the underwater channel so that as nearly as possible, only one of these discrete modes is propagating into the far field. Knowing that the propagating field consisted originally of only a single mode, tomographic techniques can characterize and locate ocean features by observing the distribution of modes received downrange.

The problem of exciting only a single mode is easily solved analytically for the time-invariant case with a uniform sound speed profile and bottom topography. However, it is much more challenging in a realistic ocean environment due to propagation losses and coupling between modes caused by time-varying inhomogeneities in the medium and range-dependent irregularities in the sea floor and water column. Our strategy is to use a reference array of receivers at the start of the far field to obtain information about the coupling in the channel, and then adaptively update the source array weights to obtain a single mode at this reference array.

Currently, we are focusing our attention on developing and investigating strategies for estimating the coupling between the sources and the reference array, and tracking the variations in this coupling over time due to a variety of ocean processes. We are using detailed ocean acoustic modeling programs such as SAFARI to simulate a realistic ocean environment for the work.

1.4 Self-Synchronization of Chaotic Systems: Analysis, Synthesis, and Applications

**Sponsors**
Advanced Research Projects Agency/
U.S. Navy - Office of Naval Research
Grant N00014-93-1-0686
Lockheed Sanders, Inc./
U.S. Navy - Office of Naval Research
Contract N00014-91-C-0125
U.S. Air Force - Office of Scientific Research
Grant AFOSR-91-0034

**Project Staff**
Kevin M. Cuomo, Professor Alan V. Oppenheim

Chaotic systems provide a rich mechanism for signal design and generation, with potential applications to communications and signal processing. Because chaotic signals are typically broadband, noise-like, and difficult to predict, they can be used in various contexts for masking information-bearing waveforms and as modulating waveforms in spread spectrum systems. A particularly useful class of chaotic systems are those which possess a self-synchronization property. Specifically, two identical chaotic systems may synchronize when the second
system is driven by the first. This property enables chaotic systems to be exploited for private communications. A potential difficulty, however, is that the analysis and synthesis of chaotic systems is not well understood because of the highly nonlinear nature of these systems.

In this research, we develop new methods for analyzing and synthesizing a class of high-dimensional dissipative chaotic systems which possess the self-synchronization property. For the class of chaotic systems that we consider, the synchronization is global and highly robust to perturbations in the drive signal. Synchronization error models are derived which provide a deeper understanding of the inherent robustness of these systems. We also develop several methods for embedding an information-bearing waveform in a chaotic carrier signal and for recovering the information at the synchronizing receiver. In addition, the practical aspects of synchronized chaotic systems are demonstrated using a transmitter and receiver circuit with dynamics governed by the chaotic Lorenz system.

1.5 Algebraic and Probabilistic Structure in Fault-Tolerant Computation

Sponsor
Advanced Research Projects Agency/ U.S. Navy - Office of Naval Research
Grant N00014-93-1-0686

Project Staff
Christoforos N. Hadjicostis, Professor George C. Verghese, Professor Alan V. Oppenheim

The traditional approach towards fault-tolerant computation has been modular redundancy. Although universal and simple, modular redundancy is inherently expensive and inefficient in its use of resources. Recently developed algorithm based fault tolerance (ABFT) techniques offer more efficient fault coverage, but their design is specific to each application. A particular class of ABFT techniques involves the design of arithmetic codes that protect elementary computations. For the case of computations that can be represented as operations in a group, the recent doctoral thesis by Beckmann has shown how to obtain a variety of useful results and systematic constructive procedures.

We are exploring the extension of this work to other algebraic structures. Our current studies indicate that much of Beckmann's framework can be generalized to the case of semigroups. In the other direction, we are examining refinements for the case of computations occurring in rings, modules, fields, or vector spaces. Beckmann's thesis shows how to take advantage of the underlying group structure in all these instances, but we expect that the additional structure that they possess can lead to stronger results and more efficient implementations of fault-tolerant schemes. Another objective of our research is to fold probabilistic models for failures and errors into the design and analysis of arithmetic codes. This will allow us to better characterize and design fault-tolerant systems.

1.6 Signal Processing Applications of Chaotic Dynamical Systems

Sponsors
AT&T Laboratories Doctoral Support Program
Advanced Research Projects Agency/ U.S. Navy - Office of Naval Research
Grant N00014-89-J-1489
Grant N00014-93-1-0686
Lockheed Sanders, Inc./ U.S. Navy - Office of Naval Research
Contract N00014-91-C-0125
U.S. Air Force - Office of Scientific Research
Grant AFOSR-91-0034

Project Staff
Steven H. Isabelle, Professor Alan V. Oppenheim

Researchers in areas ranging from animal behavior and medicine to economics and geophysics have found evidence of chaotic behavior in an enormous

---


number of empirically gathered time series. Indeed, the sheer volume of apparently random phenomena which appear to have a deterministic explanation underscores the need for signal processing techniques specifically tailored to the unique characteristics of chaotic signals. In particular, because chaotic signals can generally be observed only indirectly, e.g., through some propagation channel or nonideal laboratory instrumentation, a signal's chaotic structure may be partially obscured by additive noise and convolutional distortion. Consequently, algorithms for reducing these distortions are an important component of signal processing systems for chaotic signals. This research explores effects of convolutional distortion on chaotic signals along with techniques for reducing such distortions.

In general, the limiting trajectory of a chaotic system will be a highly structured set in the state space, while the scalar output will appear erratic and unstructured. It is this "hidden" structure that makes the signal interesting and allows for a simple description. One measure of structure which has been used to characterize a chaotic signal is the fractal dimension of its strange attractor. We are examining the effect of convolution on fractal dimension and using these results to develop deconvolution algorithms. The major challenge here is developing optimal computationally efficient techniques which are uniformly applicable to a broad class of chaotic signals.

1.7 Wavelet-Based Representation and Algorithms for Generalized Fractal Signals

Sponsors
AT&T Laboratories Doctoral Support Program
Advanced Research Projects Agency/
U.S. Navy - Office of Naval Research
Grant N00014-89-J-1489
Grant N00014-93-1-0666
Lockheed Sanders, Inc./
U.S. Navy - Office of Naval Research
Contract N00014-91-C-0125
U.S. Air Force - Office of Scientific Research
Grant AFOSR-91-0034

Project Staff
Warren M. Lam, Professor Gregory W. Womell

While the 1/f family of fractal processes have become increasingly appealing for modeling statistically scale-invariant phenomena, we study a generalization of this signal model to account for more general scaling behavior found in a wide variety of natural phenomena. While many phenomena exhibit scaling behavior, a great number of them do so only over a finite range of scales. For example, while sea floor morphology is observed to be self-similar at fine scales, such scaling behavior is absent in long length scales due to the lack of correlation among points far apart. On the other hand, many phenomena exhibit scaling behavior which varies over scales. For instance, such varying scaling behavior is encountered in the study of diluted gels and colloidal aggregates in the field of materials science.

In this work, we focus on the development of a class of generalized fractal processes for capturing such nonuniform scaling behavior. Exploiting the role of the wavelet transformation as a whitening filter for such processes, we formulate algorithms for addressing a number of practical estimation problems involving such signals. Adopting a maximum-likelihood criterion and invoking an estimate-maximize algorithm, we derive consistent, computationally-efficient spectral parameter estimators which are useful for the classification of generalized fractals. We also formulate a Bayesian minimum mean-squares error signal estimation scheme which is directly applicable in a variety of signal separation and signal recovery scenarios.

1.8 Approximate Signal Processing

Sponsor
U.S. Navy - Office of Naval Research
Grant N00014-93-1-0686

Project Staff
Jeffrey T. Ludwig, Dr. S. Hamid Nawab

In any given problem-solving domain, an algorithm, A, may be considered an approximation to another algorithm, B, provided A is more computational...
efficient than B by virtue of A being designed to yield a lower quality answer than B. Such approximate algorithms are said to carry out approximate processing with respect to the problem-solving domain under consideration. The need for such algorithms arises in applications with real-time constraints.

Within the context of real-time systems, a signal processing task may have to be performed within a time interval whose duration may not be determined prior to run time. In the case of a predetermined time allocation, an approximate processing algorithm may be used to obtain the highest quality answer possible within the allotted time. In cases where the time allocation is not predetermined, it is desirable to use an approximate processing algorithm which produces an answer of improving quality as a function of time. This allows the algorithm to be terminated whenever desired, and the quality of the resulting answer is directly proportional to the actual execution time. These types of approximate processing algorithms are said to carry out incremental refinement of their answers.

In previous research, it has been demonstrated that approximate processing algorithms can be effective in the real-time analysis of music, speech, and synthetic signals. For example, approximations to the short-time Fourier transform (STFT) have been shown to retain prominent time-frequency characteristics of such signals while improving computational efficiency by an order of magnitude over FFT-based calculations of the exact STFT.

Our aim is to develop a formal structure for using approximate processing concepts in designing novel signal processing algorithms in areas such as time-frequency analysis, cepstral analysis, multi-pass analysis, adaptive filtering, and adaptive beam forming.

1.9 Code Division Multiple Access for Digital Storage

Sponsor
U.S. Air Force - Office of Scientific Research
Grant AFOSR-91-0034

Project Staff
James M. Ooi, Professor Gregory W. Wornell

Code division multiple access (CDMA) is an area of communications research in which multiple users superimpose their transmissions on top of one another throughout an entire symbol period and throughout a larger frequency bandwidth than required. Although the technique was once primarily associated with military spread spectrum systems, designers today are finding that CDMA has properties which are desirable outside of military use. These properties include automatic interference reduction when other users are not transmitting, ease of adding new users, graceful degradation as channel capacity is exceeded, and robustness against intentional jamming and multipath interference.

The goal of this research is to extend the CDMA concept to computer storage devices in order to reap some of the above advantages. The analogous behavior on which the work will focus is graceful degradation as capacity is exceeded, tolerance to faulty memory, ease of adding new users, and an automatic improvement in fidelity as the storage requirements decrease. It is also hoped that this may also lead to a more general theory of distributed coding for the storage channel.

1.10 Estimation and Detection of a Class of Chaotic Signals

Sponsors
Advanced Research Projects Agency/
U.S. Navy - Office of Naval Research
Grant N00014-89-J-1489
Grant N00014-93-1-0686
U.S. Air Force - Office of Scientific Research
Grant AFOSR-91-0034

Project Staff
Haralabos C. Papadopoulos, Professor Gregory W. Wornell

Chaos arises in a broad range of physical and man-made systems. Naturally, there is a need for efficient and robust algorithms for estimation and detection of chaotic signals in the presence of noise. We have obtained maximum likelihood algorithms for signal filtering, smoothing, and prediction for a class of chaotic signals, which although nonlinear, possess very convenient recursive implementations that are efficient both in terms of computation and storage. Generalized estimation algorithms are suggested for a much broader class of chaotic signals.

The pseudorandom, broadband characteristics, and ease of generation of chaotic signals make them appealing candidates for signaling waveforms in secure and low probability of intercept communication systems. We explore the viability of spread spectrum communication schemes using chaotic signals and consider generalized antipodal coding.
schemes which possess attractive properties for secure communications. Optimal discrimination strategies for a class of these chaotic signaling schemes and robust, computationally efficient demodulators have been obtained. Monte Carlo simulations reveal that chaotic signaling is potentially useful for the particular communication applications mentioned above.

1.11 Real-Time Active Noise Cancellation

**Sponsor**
National Science Foundation Fellowship

**Project Staff**
Brian M. Perreault, Professor Alan V. Oppenheim

In many industrial and consumer settings, undesirable acoustic noise often exists. In many cases, the noise is so obtrusive that an effort to reduce its effect is warranted. When the noise cannot be reduced using physical, passive methods, active noise cancellation (ANC) techniques are a suitable alternative. Many types of noise have certain predictable characteristics; active noise cancellation exploits these characteristics, using destructive interference from a secondary acoustic source to eradicate the noise.

One algorithm for active noise cancellation results from approaching the problem from a control perspective. If the noise is measured at a single sensor, the essence of the technique is to use a control algorithm to give feedback through an actuator, constraining the measurement at the sensor to zero.

In many applications, such as airplane cockpits, high levels of narrowband noise are present. This type of algorithm may be useful for the cancellation of these narrow bands of noise. This algorithm is being investigated to determine its effectiveness in such an environment. A set of headphones, utilizing a microphone and speaker in each ear cup, is being developed to cancel the cockpit noise. Such a set of headphones is very desirable for a pilot, especially considering that passive attenuation of the headphones is not enough to fully compensate for the noise.

Digital predictive techniques are being used in conjunction with analog to digital (A/D) and digital to analog (D/A) converters in order to implement the active noise cancellation techniques. A digital signal processor is employed to execute the algorithms, due to its high computational ability. A flexible set of hardware has been developed in order to evaluate different ANC algorithms in real time.

1.12 State and Parameter Estimation with Chaotic Systems

**Sponsors**
Advanced Research Projects Agency/U.S. Navy - Office of Naval Research
Grant N00014-89-J-1489

U.S. Air Force - Office of Scientific Research
Grant AFOSR-91-0034

**Project Staff**
Michael D. Richard, Professor Alan V. Oppenheim

Chaotic systems are nonlinear dynamical systems characterized by extreme sensitivity to initial conditions. A signal generated by a chaotic system may appear random, despite its having been generated by a low-order, deterministic dynamical system. Both random and chaotic signals lack long-term predictability; but, in contrast to truly random signals, chaotic signals exhibit short-term predictability. Evidence of chaotic behavior has been reported in many diverse disciplines including physics, biology, engineering, and economics.

We are exploring techniques for state and parameter estimation with chaotic systems. We have implemented the extended Kalman filter,¹ a recursive state estimator for nonlinear systems, and several related algorithms¹⁰ and have evaluated their effectiveness as state estimators for chaotic systems. Preliminary results have shown these algorithms to perform reasonably well. But results have also shown that these algorithms have potentially unacceptable deficiencies when applied to chaotic systems.

More recently, we have developed and begun testing several related, novel, state-estimation techniques loosely motivated by maximum likelihood

---


state estimation. The techniques exploit a distinguishing property of all chaotic systems—the simultaneous existence of stable and unstable manifolds. The combination of these techniques with an estimate-maximize (EM) algorithm is also being considered. Finally, we plan to ascertain the value of various state-estimation techniques in improving the short-term predictability of chaotic signals.

1.13 Model-Based Analysis of Music

Sponsors
Advanced Research Projects Agency/
U.S. Navy - Office of Naval Research
Grant N00014-89-J-1489
National Science Foundation Fellowship

Project Staff
Stephen F. Scherock, Dr. Bernard Gold

Many digital audio signals can be broadly classified as speech or music. Speech signals have been studied extensively since the 1950s for the purpose of automatic speech recognition, production, and data compression. Work has advanced toward each of these goals in part because of the development of models for speech production and recognition. While physically based models exist for music production, applying these models for synthetic music production or data compression has not been fully exploited. Our research involves model-based compression of music from a single instrument, the trumpet.

1.14 Nonlinear Models for Signal Processing

Sponsors
Advanced Research Projects Agency/
U.S. Navy - Office of Naval Research
Grant N00014-89-J-1489
Lockheed Sanders, Inc./
U.S. Navy - Office of Naval Research
Contract N00014-91-C-0125

Project Staff
Andrew C. Singer, Professor Alan V. Oppenheim

This research involves the use of nonlinear system models in signal processing. Predominantly, linear models and algorithms have been employed in this field due to the tractability of their analysis and the richness of the class of signals for which these models are well-suited. However, by trading complexity for tractability, we can further broaden the boundaries of signal processing. For example, we have recently shown that nonlinear signal modeling techniques are both useful and practical for modeling a variety of signals for which linear techniques have proven inadequate. Further, we have begun to find practical applications for nonlinear dynamic systems in chaos, which until recently were thought of as more paradoxical than practical. Another interesting phenomenon whose curious behavior has crossed many disciplines of science, is the theory of solitons. It is our goal to exploit the behavior of these solitary waves and other nonlinear phenomena in search of new paradigms and new directions in signal processing.

1.15 Environmental Robustness in Automatic Speech Recognition

Sponsor
Maryland Procurement Office
Contract MDA904-93-C-4180

Project Staff
Shawn M. Verbout, Professor A.V. Oppenheim

As the technology in automatic speech recognition (ASR) becomes increasingly advanced, greater consideration is given to speech as a viable means of interaction between humans and machines. However, as more ASR systems become incorporated into real-world applications, greater attention must be devoted to making these systems robust with respect to changes in their operating environments. The need for environmental robustness is clear when we consider the difficulties presently encountered in attempting to make ASR systems yield consistently high performance in practical settings. In particular, even the most sophisticated ASR systems today are extremely sensitive to variations in recording conditions and must be carefully retrained for each new environment in which they are deployed.

---


In the context of ASR, environmental variability is a broad notion encompassing the various changes that can occur in any factor affecting acoustic measurements. For example, environmental changes can be due to physiological differences between different speakers; variations in loudness for a single speaker; differences in the physical relationship between the speaker and the microphone; changes in the ambient noise level; or differences in recording equipment and room acoustics. In addition to these factors, which may be termed "physical," there are other environmental factors that may be viewed as "situational." For example, consider the operation of an ASR system in the relatively benign setting of a typical business office. In this environment, a speech signal can be corrupted during a recording session by unanticipated acoustic events such as sudden activation of a manifold system, intermittent keyboard clicking, background conversations, slamming doors and file cabinet drawers, and reverberation due to acoustic reflections from walls, floors, and various objects in the room.

The long-term objective of our research in this area is to determine the signal processing technology required to make the performance of an ASR system relatively insensitive to these kinds of changes in the recording environment, while keeping the performance at a consistently high level. One viable method for making an ASR system more robust is to replace the single sensor in a typical recording environment with multiple sensors located in various parts of the room. This multisensor approach affords some redundancy in the measurement of the primary speech signal, and simultaneously provides a diverse set of reference measurements for noise events throughout the room. The resulting multiple signal measurements may then be processed to extract the primary speech signal from the interfering noise sources. Important intermediate goals in our investigation of the multisensor robustness approach will include quantifying the performance of an existing ASR system with and without preprocessing of the sensor measurements using algorithms such as those developed in an RLE technical report.14

1.16 Active Noise Cancellation

Sponsors

Advanced Research Projects Agency/
U.S. Navy - Office of Naval Research
Grant N00014-89-J-1489
U.S. Air Force - Office of Scientific Research
Grant AFOSR-91-0034

Project Staff

Kambiz Zangil, Professor Alan V. Oppenheim

Unwanted acoustic noise is a by-product of many industrial processes and systems. With active noise cancellation (ANC), one introduces secondary noise sources to generate an acoustic field that interferes destructively with the unwanted noise, and thereby eliminating it. Examples of such unwanted noise include machinery noise, aircraft cabin noise, and fan noise.

Traditional active noise cancellation systems assume that the statistical characteristic of the primary noise is known a priori. Furthermore, almost all of the existing systems use two microphones, and as result suffer from acoustic feedback between the canceling speaker and the input microphone.15

We have developed an adaptive active noise cancellation system which only uses one microphone, and therefore has no feedback problem. This system uses the estimate maximize (EM) algorithm to simultaneously estimate noise statistics and the transfer function between the canceling speaker and the microphone. An optimal canceling signal is then generated based on these estimates.16

We have also developed a two-microphone version of the above system which does not suffer from the


feedback problem and is more intelligent in using the outputs of the microphones.

We are currently studying the problem of noise cancellation in a volume. A topic of fundamental interest is to find analytically simple ways to describe the sound field over a volume from measurements made at a finite set of points in that volume. Similarly, we would like to find ways to alter the sound field in a desired manner using only a finite number of sources.

1.17 Oceanographic Signal Processing

Sponsor
U.S. Navy - Office of Naval Research
Grant N00014-91-J-1628

Project Staff
Professor Arthur B. Baggeroer

Our research programs involve the application of signal and array processing to problems in ocean acoustics and geophysics. Our research requires an understanding of digital signal processing and wave propagation; moreover, most research projects use data from laboratory or field experiments, so an appreciation of real world issues such as noise, sensor calibrations and modeling errors is needed. Several of the topics provide opportunities to participate in oceanographic cruises.

1.17.1 Acoustic Thermometry of Ocean Climate

The Heard Island Feasibility Test demonstrated that coded acoustic signal can be transmitted over 10,000 km ranges. This has led to the Acoustic Thermometry of Ocean Climate (ATOC) program by the Advanced Research Projects Agency (ARPA). A network of acoustic sources and receivers is being deployed in the Pacific Ocean this spring to monitor it by measuring changes in acoustic travel times. These changes will be used to infer temperature changes.

1.17.2 Directional Spectrum Estimation: a Maximum Likelihood Method

Project Staff
John P. Mann, Professor Arthur B. Baggeroer

Antenna arrays can be used to estimate the direction of arrival of signals incident upon the array. Capon's minimum variance method is commonly used to estimate the frequency spectrum arriving from a desired direction. However, this method has some limitations. A finite set of samples from the array is used to form a sample covariance, which must then be inverted to calculate the minimum variance estimate of the spectrum. In order for the sample covariance to provide reasonable results, it has been shown that the number of samples necessary is approximately twice the number of sensors in the array.17 In cases where the signal environment changes fast enough, this requirement becomes a problem. Estimates need to be updated faster than is possible given this constraint.

In this research, we will investigate methods of directional spectrum estimation when the number of samples available is limited. In addition to some commonly used algorithms, a maximum likelihood method for estimating frequency wavenumber spectra has been derived18 and will be evaluated and refined. The performance of the various methods will be compared through simulations on a series of different array geometries and signal environments. The effect of errors in the gain and position of sensor elements will also be investigated.

1.17.3 Implementation of Single Mode Excitation in the Shallow Water Acoustic Channel

Project Staff
Denis J. Peregrym, Professor Arthur Baggeroer

Acoustic signals can be described as a sum of normal modes or eigenfunctions of a waveguide. Under certain conditions in shallow water, a small set of discrete modes carries almost all of the acoustic energy. Our goal is to design a source which excites only a single mode in the waveguide at a reference receiver some distance down the guide from the source. The signal at this receiver


is used to adaptively update the source weights and converge to a single mode. Our current research involves investigating the field shape in the waveguide and finding an optimum position for the reference array in the near field/far field boundary. In addition, we are investigating how the mode shapes or weights at the source array required to generate a single mode can be used to characterize the channel.

1.17.4 Estimation of Normal Mode Amplitudes for Underwater Acoustics

Project Staff
Kathleen E. Wage, Professor Arthur B. Baggeroer

In underwater acoustics, signals are often efficiently represented as weighted sums of the normal modes or eigenfunctions associated with the waveguide. The weights are usually estimated from data obtained with a vertical array of hydrophones. The purpose of this research is to investigate methods for estimating the modal amplitudes (i.e., weights). In addition to studying modal amplitude estimation techniques previously developed, this work also applies several bearing estimation algorithms to the modal amplitude problem. A least squares formulation is the most common in the literature, and several variations of this generalized approach are under consideration. The three bearing estimation algorithms which will be implemented were developed in the context of direction-of-arrival location for diversely-polarized antennas. Our research considers the effects of noise, modeling assumptions, mode coherence, and multiple modal sources on the behavior of each estimator.

The motivation for this work is twofold. First, modal decomposition is a natural way to approach solutions to the acoustic wave equation, and a knowledge of the modal structure is often helpful in exploring the effects of a particular channel on a transmitted signal. Specifically, mode attenuation is important in many acoustic and oceanographic processes, e.g., Kuperman and Ingenito have explored the attenuation of modes due to scattering at rough boundaries. A second motivating factor is that the modal amplitudes are useful in matched mode processing and in matched field tomography. Recent papers indicate that source range and depth information may be obtained by performing matched field beamforming using the modal coefficients. Shang has also used adiabatic normal mode theory to develop a method for acoustic tomography. Obviously, the ability to estimate the mode coefficients accurately and efficiently is essential to the success of these types of algorithms.

1.18 Publications

Journal Articles


Meeting Papers


Theses


Chapter 2. Advanced Telecommunications and Signal Processing Program

Academic and Research Staff
Professor Jae S. Lim

Visiting Scientists and Research Affiliates
Dr. Hae-Mook Jung

Graduate Students

Technical and Support Staff
Cindy LeBlanc, Debra L. Harring, Denise M. Rossetti

2.1 Introduction
The present television system was designed nearly 40 years ago. Since then, there have been significant developments in technology, which are highly relevant to the television industries. For example, advances in the very large scale integration (VLSI) technology and signal processing theories make it feasible to incorporate frame-store memory and sophisticated signal processing capabilities in a television receiver at a reasonable cost. To exploit this new technology in developing future television systems, Japan and Europe established large laboratories, funded by government or industry-wide consortia. The lack of this type of organization in the U.S. was considered detrimental to the broadcasting and equipment manufacturing industries, and in 1983 the Advanced Television Research Program (ATRP) was established at MIT by a consortium of U.S. Companies.

The major objectives of the ATRP are:

- To develop the theoretical and empirical basis for the improvement of existing television systems, as well as the design of future television systems.
- To educate students through television-related research and development and to motivate them to undertake careers in television-related industries.
- To facilitate continuing education of scientists and engineers already working in the industry.
- To establish a resource center to which problems and proposals can be brought for discussion and detailed study.
- To transfer the technology developed from this program to the industries.

The research areas of the program include the design of receiver-compatible ATV system and digital ATV system and development of transcoding methods. Significant advances have already been made in some of these research areas. A digital ATV system has been designed and tested in the fall of 1992 by the FCC for its possible adoption as the U.S. HDTV standard for terrestrial broadcasting. Some elements of this system are likely to be included in the U.S. HDTV standard.

In addition to research on advanced television systems, the research program also includes research on speech processing. Current research topics include development of a new speech model and development of algorithms to enhance speech degraded by background noise.

2.2 ATRP Facilities
The ATRP facilities are currently based on a network of eight Sun-4 and three DecStation 5000 workstations. There is approximately 14.0 GB of disk space distributed among the various machines. Attached to one of the Sun-4s is a VTE display system with 256 MB of RAM. This display system is capable of driving the Sun-4 monitors or a 29-inch Conrac monitor in the lab at rates up to 60 frames/sec. In addition to displaying high-resolution real-time sequences, the ATRP facilities include a Methues frame buffer which drives a Sony 2kx2k monitor. For hard copy output, the lab uses a Kodak XL7700 thermal imaging printer which can
produce 2k×2k color or black and white images on 11"×11" photographic paper.

Other peripherals include an Exabyte 8 mm tape drive, a 16-bit digital audio interface with two channels and sampling rates up to 48 kHz per channel, and an "audio workstation" with power amplifier, speakers, CD player, tape deck, etc. Additionally, the lab has a 650 MB optical disk drive, a CD-ROM drive, and two laser printers. For preparing presentations, the ATRP facilities also include a Macintosh SE30 microcomputer, a Mac IIX, and an Apple LaserWriter.

Obtaining a fast network (FDDI) is under consideration to augment the current 10 Mbps Ethernet. The new network would enable much faster data transfer to display devices, and it would support large NFS transfers more easily.

2.3 Very-low-bit-rate Video Representations

Sponsor
Advanced Telecommunications Research Program

Project Staff
John G. Apostolopoulos

Video plays an important role in many of today's applications, and it is expected to gain even greater importance in the near future with the advent of multimedia and personal communication devices. The large raw data rate of a video signal, together with the limited available transmission capacity in many applications, necessitates compression of the video signal. A number of video compression algorithms have been developed for different applications from video-phone to high-definition television. These algorithms perform reasonably well at respective bit rates of 64 kb/s to tens of Mb/s. However, many applications in the near future, particularly those associated with portable or wireless devices, will most likely be required to operate at considerably lower bit rates, possibly as low as 10 kb/s. The video compression methodologies developed thus far are not applicable at such low bit rates. The goal of this research is to create a video compression approach which can produce acceptable video quality at approximately 10 kb/s.

Conventional video compression algorithms may be described as block-based coding schemes. They partition each frame into square blocks and then independently process each block. These redundancy-reduction techniques include block-based temporal motion-compensated prediction and spatial block discrete-cosine transform. Block-based processing is the basis for virtually all video compression systems today because it is a simple and practical approach to achieve acceptable video quality at the required bit rates. However, block-based coding schemes cannot effectively represent a video signal at very low bit rates. This is because the source model used is extremely limited. Block-based schemes inherently assume a source model of (translational) moving square blocks. However, a typical video scene is not composed of translated square blocks. In effect, block-based schemes impose an artificial structure on the video signal and then try to encode this structure, instead of recognizing the structure inherent to a particular video scene and attempting to exploit it.

In this research a conceptual framework is developed for identifying the structure that exists within a video scene. This is one of the most important perceptual aspects of the information within the scene. Therefore, by identifying and efficiently representing this structure, it may be possible to produce acceptable video quality at very low bit rates. Since real scenes contain objects, a promising source model includes two- or three-dimensional moving objects. A sizable amount of current research involves image segmentation, however, most of this research has focused primarily on image analysis and extremely little on the possibility of compression. Very importantly, significant statistical dependencies exist in the image regions belonging to each object, and this must be exploited. This point has not been addressed; only the previously mentioned block-based schemes have been discussed in the literature.

An efficient compression algorithm must therefore consist of a powerful and compact framework for accurately representing possible objects or structure in a video signal, as well as an efficient approach to represent the information that exists within each object.

The goal of this research is to create a video compression approach which can produce acceptable video quality at approximately 10 kb/s. This will allow video applications in the near future, particularly those associated with portable or wireless devices, to operate at performance levels not achievable by current video compression schemes. In order to attain this goal, we are developing a conceptual and mathematical framework for identifying and exploiting the structure that exists within a video signal.
Chapter 2. Telecommunications and Signal Processing Research

2.3.1 Publications


2.4 Audio Compression Using Hierarchical Nonuniform Filterbanks

Sponsor
Advanced Telecommunications Research Program

Project Staff
Shiufun Cheung, Peter A. Monta

Interest in coding high-fidelity audio has increased in recent years; audio source coding algorithms are finding use in transmission applications such as digital audio broadcast (DAB) and high-definition television (HDTV) and in commercial products such as MiniDisc and Digital Compact Cassette. The common objective is to achieve high quality at a rate significantly smaller than the 16 bits/sample used in CD and DAT systems. We have been considering applications to HDTV, and an earlier implementation, the MIT Audio Coder (MIT-AC) has successfully completed testing at the Advanced Television Test Center. In this research, we study an improved signal representation scheme which uses a hierarchical nonuniform filterbank.

In many waveform coders, the first step in the audio coding process is a short-time spectral decomposition of the signal. Until recently, the representations used by many audio coders have uniform filterbank bandwidths. This results in an undesirable trade-off between time and frequency resolution. For example, if the analysis bandwidth is chosen to be narrow enough to resolve the critical bands in the low frequencies, the resulting poor temporal resolution can result in temporal artifacts, such as the "pre-echo" effect.

There have been various efforts to correct this deficiency. For example, some coders include an adaptive mechanism for improving temporal resolution by adjusting the block transform size when a transient is detected. Another more fundamental approach is the use of a nonuniform filterbank, which can simultaneously satisfy the requirements of good pre-echo control and frequency resolution consistent with critical-band analysis. It also allows a more uniform architecture without an explicit adaptation step.

Previous development along these lines has considered the use of wavelet representations. In our prototype implementation, we use a hierarchical filterbank structure based on M-band perfect-reconstruction cosine-modulated filterbanks (for various values of M). These filterbanks have been developed by Malvar and in an equivalent formulation by Vaidyanathan.

Testing of the new representation has shown improvement over an uniform filterbank scheme. While pre-distortion still exists in transient signals, listening sessions show that pre-echo is inaudible with the new filterbank.

One potential disadvantage of a fixed filterbank containing some wide subbands is that the system may not realize the theoretical coding gain if the signal spectrum is very peaked. Further research is now being performed to devise schemes which can adaptively subdivide the filterbank tree based on spectral flatness measurements.

2.5 Transform Coding for High-Definition Television

Sponsor
Advanced Telecommunications Research Program

Project Staff
Ibrahim A. Hajjahmad

The field of image coding is useful for many areas. For of these areas is the reduction of channel bandwidth needed for image transmission systems, such

---


Chapter 2. Telecommunications and Signal Processing Research

as HDTV, video conferencing, and facsimile. Another area is the reduction of storage requirements. One class of image coders is known as the transform image coder. In transform image coding, an image is transformed to another domain, more suitable for coding than the spatial domain. The obtained transform coefficients are quantized and then coded. At the receiver, the coded coefficients are decoded and then inversely transformed to obtain the reconstructed image.

One transform which has shown promising results is the discrete cosine transform (DCT). The DCT is a real transform with two important properties that make it very useful in image coding. One is the energy compaction property, where large amount of energy is concentrated in a small fraction of the transform coefficients (typically low frequency components). This property allows us to code a small fraction of the transform coefficients with a small sacrifice in quality and intelligibility of the coded images. Another is the correlation reduction property. In the spatial domain there is a high correlation among image pixel intensities. The DCT reduces this correlation and redundant information does not need to be coded.

Current research is investigating the use of the DCT for bandwidth compression. New adaptive techniques are also being studied for quantization and bit allocation that can further reduce the bit rate without reducing image quality and intelligibility.

2.6 Pre-Echo Detection and Reduction

Sponsor
Advanced Telecommunications Research Program

Project Staff
Kyle K. Iwai

In recent years, there has been an increasing interest in data compression for storage and data transmission. In the field of audio processing, various kinds of transform coders have successfully demonstrated reduced bit rates while maintaining high audio quality. However, there are certain coding artifacts which are associated with transform coding. The pre-echo is one such artifact. Pre-echos typically occur when a sharp attack is pre-coded by silence. Quantization noise added by the coding process is normally hidden within the signal. However, the coder assumes stationarity over the window length, an assumption which breaks down in a transient situation. The noise is unmasked in the silence preceding the attack, creating an audible artifact called a pre-echo.

If the length of the noise can be shortened to about 5 ms, psycho-acoustic experiments tell us that the noise will not be audible. Using a shorter window length shortens the length of the pre-echo. However, shorter windows also have lower quality frequency selectivity and lower quality coder efficiency. One solution is to use shorter windows only when there is a quiet region followed by a sharp attack.

In order to use adaptive window length selection, a detector had to be designed. A simple detector was implemented which compares the variance within two adjacent sections of audio. In a transient situation, the variance suddenly increases from one section to the next. The coder then uses short windows to reduce the length of the pre-echo, rendering the artifact inaudible.

2.7 Video Source Coding for High-Definition Television

Sponsor
Advanced Telecommunications Research Program

Project Staff
Peter A. Monta

Efficient source coding is the enabling technology for high-definition television over the relatively narrow channels envisioned for the new service (e.g., terrestrial broadcast and cable). Coding rates are on the order of 0.3 bits/sample, and high quality is a requirement. This work focuses on new source coding techniques for video relating to representation of motion-compensated prediction errors, quantization and entropy coding, and other system issues.

Conventional coders represent video with the use of block transforms with small support (typically 8x8 pixels). Such independent blocks result in a simple


scheme for switching a predictor from a motion- 
compensated block to a purely spatial block; this is 
necessary to prevent the coder from wasting 
capacity in some situations.

Subband coders of the multiresolution or wavelet 
type, with their more desirable localization properties, 
lack of "blocking" artifacts, and better match to 
motion-compensated prediction errors, complicate 
this process of switching predictors, since the 
blocks now overlap. A novel predictive coding 
scheme is proposed in which subband coders can 
combine the benefits of good representation and 
flexible adaptive prediction.

Source-adaptive coding is a way for HDTV systems 
to support a more general imaging model than con-
ventional television. With a source coder that can 
adapt to different spatial resolutions, frame rates, 
and coding rates, the system may then make tradeoffs among the various imagery types, for 
example, 60 frames/s video, 24 frames/s film, highly 
detailed still images, etc. In general, this is an 
effort to make HDTV more of an image transport 
system rather than a least-common-denominator 
format to which all sources must either adhere or 
be hacked to fit. These techniques are also appli-
cable to NTSC to some extent; one result is an 
algorithm for improved chrominance separation for 
the case of "3-2" NTSC, that is, NTSC upsampled 
from film.

2.8 Error Concealment for an All-Digital 
HDTV System

Sponsor
Advanced Telecommunications Research 
Program

Project Staff
Aradhana Narula

Broadcasting high-definition television (HDTV) 
requires the transmission of an enormous amount 
of information within a highly restricted bandwidth 
channel. Adhering to the channel constraints 
necessitates the use of an efficient coding scheme to 
compress the data. Compressing the data dra-
matically increases the effect of channel errors. In 
the uncompressed video representation, a single 
channel error affects only one pixel in the received 
image. In the compressed format, a channel error 
affects a block of pixels in the reconstructed image, 
perhaps even the entire frame.

One way to combat the effect of channel errors is 
to add well-structured redundancy to the data 
through channel coding. Error correction schemes 
generally, however, require transmitting a significant 
number of additional bits. For a visual product like 
HDTV, it may not be necessary to correct all errors. 
Instead, removing the subjective effects of channel 
errors using error concealment techniques may be 
sufficient and require fewer additional bits for imple-
mentation. Error concealment may also be used in conjunction with error correction coding. For 
extample, it may be used to conceal errors which 
the error correction codes are able to detect but not correct.

Error concealment techniques take advantage of 
the inherent spatial and temporal redundancy within 
the transmitted data to remove the subjective 
effects of these errors once the location of the 
errors has been detected. In this research, error 
concealment techniques were developed and ana-
lyzed to help protect the system from errors occur-
ring in several parameters transmitted for HDTV 
images. Specifically, error concealment for errors 
in the motion vectors and discrete cosine transform 
(DCT) coefficients were investigated.

2.9 Transmission of HDTV Signals In a 
Terrestrial Broadcast Environment

Sponsor
Advanced Telecommunications Research 
Program

Project Staff
Julien J. Nicolas

High-definition television systems currently being 
developed for broadcast applications require 15-20 
Mbps to yield good quality images for roughly twice 
the horizontal and vertical resolutions of the current 
NTSC standard. Efficient transmission techniques 
must be found in order to deliver this signal to a 
maximum number of receivers while respecting the 
limitations stipulated by the FCC for over-the-air 
transmission. This research focuses on the princi-
ples that should guide the design of such trans-
mission systems.

The major constraints related to the transmission of 
broadcast HDTV include (1) a bandwidth limitation 
(6 MHz, identical to NTSC), (2) a requirement for 
simultaneous transmission of both NTSC and HDTV 
signals on two different channels (Simulcast 
approach), and (3) a tight control of the interference 
effects between NTSC and HDTV, particularly when 
the signals are sharing the same frequency bands. 
Other considerations include complexity and cost.
issues of the receivers, degradation of the signal as a function of range, etc.

A number of ideas are currently under study. Most systems proposed to date use some form of forward error-correction to combat channel noise and interference from other signals. The overhead data reserved for error-correction schemes represents up to 30 percent of the total data, and it is therefore well worth trying to optimize these schemes. Current work is focusing on the use of combined modulation/coding schemes capable of exploiting the specific features of the broadcast channel and the interference signals. Other areas of interest include use of combined source/channel coding schemes for HDTV applications and multi-resolution coded modulation schemes.

2.10 Position-Dependent Encoding

Sponsor
Advanced Telecommunications Research Program

Project Staff
Alexsander Pfajfer

In typical video compression algorithms, the DCT is applied to the video, and the resulting DCT coefficients are quantized and encoded for transmission and storage. Some of the DCT coefficients are set to zero. Efficient encoding of the DCT coefficients is usually achieved by encoding the location and amplitude of the non-zero coefficients. In typical MC-DCT compression algorithms up to 90 percent of the available bit rate is used to encode the location and amplitude of the non-zero quantized DCT coefficients. Therefore, efficient encoding of the location and amplitude information is extremely important for high quality compression.

Position-dependent encoding, a novel approach to encoding of the location and amplitude information, is being examined. Position-dependent runlength encoding and position-dependent encoding of the amplitudes attempts to exploit the inherent differences in statistical properties of the runlengths and amplitudes as a function of their position. This novel method is being compared to the classical, separate, single-codebook encoding of the runlength and amplitude, as well as to the joint runlength and amplitude encoding.

2.11 HDTV Transmission Format Conversion and the HDTV Migration Path

Sponsor
Advanced Telecommunications Research Program

Project Staff
Lon E. Sunshine

The current proposal for terrestrial HDTV broadcasting allows for several possible transmission formats. Because production and display formats may differ, it will be necessary to convert between formats in an effective way. A key to this process is the de-interlacing process. Since HDTV will presumably move toward progressive display systems, it will be necessary to de-interlace non-progressive source material. The research will consider topics relating to conversion among the six formats being proposed for the U.S. HDTV standard.

As HDTV evolves, it is probable that more transmission formats will be allowed. Furthermore, additional bandwidth may be allocated for some channels (terrestrial and/or cable). This research will consider the issues related to the migration of HDTV to higher resolutions. Backward compatibility and image compression and coding issues will be addressed.

2.12 Speech Enhancement

Project Staff
Chang Dong Yoo

The development of the dual excitation (DE) speech model has led to some interesting insights into the problem of speech enhancement, and based on the ideas of the DE model, a new speech model is being developed. The DE model provides more flexible representation of speech and possesses features which are particularly useful to the problem of speech enhancement. These features, along with a variable length window, are the backbone of the new speech model being developed.

Because the DE model does not place any restrictions on its characterization of speech, the enhancement system based on the DE model performs better than the one based on any of the previous speech models. While a model should be inclusive in its characterization, it should have some restrictions. Specifically, a speech model should pertain to speech. The DE model is somewhat
unrestrictive and simple in its characterization of speech. It is solely based on the separation of the voiced and unvoiced components. Whether it makes sense to represent a stop as a voiced and an unvoiced component is just one of many interesting issues which are being investigated. An extension of the DE model which deals with these issues better is currently being studied.

All model-based enhancement methods to date have been formulated on the premise that each segment of speech be stationary for a fixed window length. To improve the performance of the enhancement algorithm, this assumption of stationarity must be assured. To do this, a variable-length window should be used to capture varying durations of stationarity in the speech. There are several algorithms which adaptively detect changes in auto-regressive model parameters in quasi-stationary signals which have been successfully used in speech recognition. We propose to investigate some of these algorithms. The benefit from using a variable length window is two-fold: (1) It will allow better and “cleaner” separation of the voiced and unvoiced components; and (2) It will allow for a greater reduction in the number of characteristic parameters, such as the amplitudes of the voiced components and the LP coefficients of the unvoiced component.
Professor Emeritus William F. Schreiber
Chapter 3. Combined Source and Channel Coding for High-Definition Television

Academic and Research Staff
Professor Emeritus William F. Schreiber
Graduate Students
Michael O. Polley, Susie J. Wee
Technical and Support Staff
Deborah S. Manning

3.1 Introduction

Sponsor
Scitex America Corporation

The purpose of this project is to develop a high-definition television (HDTV) system for terrestrial broadcasting. In accordance with ground rules set up by the Federal Communications System, each TV licensee will be given a second channel within the current spectrum allocation. This new channel, to be awarded without competition, must be used for HDTV broadcasting. The existing channels will continue to be used for standard broadcasting. For much of the day, the same program must be sent on both channels, a procedure called simulcasting. In many cases, the new channels will be so-called taboo channels—those that cannot be used today for standard broadcasting because of interference. However, coverage on the new channels must be about the same as coverage on existing channels. Obviously, the HDTV format must have much better interference performance than the current format.

Two additional requirements for this new system are: (1) the picture must be the highest quality possible and (2) the system must operate satisfactorily in spite of noise, ghosts, interference, and frequency distortion. In addition, the system must have good interoperability with other media, and nondisruptive improvement over a period of time must be possible.

3.2 System Development

As discussed in RLE Progress Reports Numbers 134 and 135, in our initial plan we had intended to use a two-stage spread-spectrum (SS) system. In this scheme, the video bandwidth is divided into a large number of narrow-band signals. Then each signal is spread to full (6 MHz) channel bandwidth by a unique spreading sequence with all the spread signals added. However, we found that this method did not have adequate ghost (multipath) suppression. As discussed in last year’s Progress Report, we combined the spread-spectrum method with the orthogonal frequency-division multiplex (OFDM) method to improve ghost performance.

In both the OFDM and the SS methods, the signal is divided into many channels so that the symbol length in each channel is substantially longer than the temporal spread of the multipath. A guard interval is added to each symbol in which a portion of the symbol is repeated. By integrating over exactly one symbol length at the receiver, multipath up to the length of the guard interval is entirely eliminated. This performance has been verified in the simulation exercise.

3.2.1 Current Status

The system shown in figure 1 is under simulation. As in our earlier system, this system uses hybrid analog/digital source coding and channel coding. This gives higher spectrum efficiency than all-digital channel coding while having just as good compression performance in the source coder.

The source coder generates both analog and digital data. The analog data consists of selected Fourier coefficients. The digital data comprises DC and very low-frequency data, audio, ancillary data, and adaptive selection information for identifying the selected coefficients. As in our earlier system, analog data is subjected to a spread-spectrum operation. After forward error correction (FEC) has been added to the digital data, the analog data is added to it.

This procedure produces hybrid analog/digital symbols that are fed to the OFDM modulator. (The combination of coding with OFDM is called coded
Chapter 3. Combined Coding for High-Definition Television

3.3 Error Correction

The digital data must be received nearly error-free. We accomplish this by using a concatenated coder consisting of a Reed-Solomon coder (adding check bits to locate errors) and a convolutional coder, which increases the number of levels in the signal to regain channel capacity normally lost by quantization. This combination has proved extremely effective.

3.4 Channel Equalization

One benefit of OFDM is that, since the multipath spread is shorter than the symbol length, demodulation into the separate channels can be done before equalization, thus improving the SNR of the recovered signals. Multipath causes the gain to be nonuniform across the channel. At frequencies where the gain is low, noise is amplified. In this case, there are fewer problems if equalization is done after demodulation.

Since there are so many subcarriers, the channel gain can be described by a single gain and phase for each. These are easily measured by a training signal. Correction is easily implemented since the amplitudes of the various carriers are separated out by the demodulation process.
3.4.1 Current Status

We have fully simulated the channel coder and are now measuring its properties. We are designing the source coder which is a modification of the popular MPEG scheme but, in our scheme, the amplitude and identification of coefficients are transmitted independently. We are also developing a multiresolution version of the system. Pictures of different quality are recovered, depending on receiving conditions. Finally, we are investigating a transmission constellation that separates the digital and analog data streams so that each is recovered without interference from the other.
Section 1  Genetic Analysis

Chapter 1  Genosensor Technology Development
1.1 Introduction

The primary objective of this cooperative work with the Houston Advanced Research Center (HARC) is to develop a novel method for automated, low-cost, high-throughput DNA sequence analysis. The overall goal is to demonstrate laboratory prototypes that provide a substantial increase in speed over the conventional DNA sequencing methods now used in biomedical, pharmaceutical, and agricultural industries.

The basic approach being taken is depicted in figure 1. In a hypothetical DNA sequencing test, a solution of single-stranded "target" DNA strands of identical but unknown sequence is washed onto a specialized microelectronic chip called a genosensor. The genosensor surface contains a large array of test sites, each site containing short pieces of single-stranded DNA known as "probes." These probes are chemically attached to the site. All probes in a given site are of like sequence, and the sequence for each site is unique on the chip. The target DNA strands will bond, or hybridize, very strongly to probes containing their exact Watson-Crick complement, but much less strongly to probes on other sites. The sites containing hybridized DNA are identified via electronic sensing on the chip, and this information is used by off-chip instrumentation to reconstruct the sequence of the target strands. MIT's role in this effort is primarily the design and fabrication of the genosensor chips.

1.2 Development of Genosensor Arrays for DNA Decoding

Sponsor

Houston Advanced Research Center
Contract HRC-HG00665-01

Project Staff
Dr. Daniel J. Ehrlich, Dr. Dennis D. Rathman, Dr. Mark A. Hollis, Dr. John Melngailis

1.2.1 Genosensor Electronic-Detection Principle

The simplest electrical measurement that can be made at a test site to detect hybridization is probably a measurement of the change in local permittivity due to the addition of long target strands to the site. The complex permittivity $\varepsilon' - j\varepsilon''$ of an aqueous solution containing DNA exhibits a dispersion around a relaxation frequency which is a function of the size and conformation of the DNA molecule. A measurement of the capacitance and/or conductance between two electrodes in the solution over a range of frequency can therefore differentiate between a site that contains only short probe strands and one that contains long target strands hybridized to the probe strands. From these measurements the relative permittivity $\varepsilon'$, the dielectric loss $\varepsilon''$, and the dissipation factor $\varepsilon''/\varepsilon'$ can be obtained for the cell.

The ideal electrode structure in a test well consists of two parallel plates spaced so that the entire volume between them is filled by the hybridized DNA globules in aqueous solution. For the sizes of target DNA envisioned, this spacing ranges from approximately 200 to a few thousand angstroms. A practical, easily fabricated structure that approximates this ideal is an interdigitated design shown in figure 2. Fabricated by a combination of wet and dry etching with metal liftoff, this design can achieve the required spacings between the upper and lower electrodes at their edges.

1.2.2 Genosensor Fabrication Development

The primary emphasis of this grant is on the process development and feasibility exploration for the genosensor design of figure 2. For this chip, conventional 2-µm photolithography, wet and dry etching techniques, and electron-beam evaporation and liftoff processes are used to fabricate an interdigitated electrode structure in which the opposing electrodes are separated vertically.

---

1 Professor, Department of Electrical Engineering, University of Maryland, College Park, Maryland, as of September 1993.
Chapter 1. Genosensor Technology Development

**Figure 1. Conceptual genosensor system.**

During this past year, significant progress has been made in process development. A new 6x6 passive-array mask design has been completed and several genosensor chips have been fabricated, packaged, and delivered to HARC for testing. Figure 3 illustrates various aspects of the genosensor test chips. While the electrode-finger configuration was fixed at 2-μm width for both upper and lower electrodes, the numbers of electrode fingers and the finger lengths were systematically varied from site to site on each chip. This was done primarily for a two-fold purpose: (1) to test relative signal sensitivity as a function of cell size, and (2) to aid in the evaluation of device yield. The largest single device is 800x800 μm, which is 64 times larger than our standard design. Since the primary fatal defect affecting device yield is a metal short between the opposing electrodes, a good yield for these large devices would likely project a very high yield for the standard devices. In addition, signal sensitivity should be greater for the large-area devices. Signal sensitivity can also be enhanced by bringing the opposing electrodes closer, either by using thicker evaporated metal films or by reducing the oxide etch depth. (This has been verified by some modeling work performed under contract HRC-HG00776-01; see section 3 below.) However, this increases the probability of shorts and thus lowers the yield. For electrode vertical gaps greater than 3000 Å, the device yield at the cell site is better than 90 percent. For gaps less than 2000 Å, the yield is less than 50 percent for the largest-area...
devices. Finally, the new design also eliminated one mask level and a difficult etching step.

The other significant development for the period was the implementation of Pt instead of Au as the electrode material. Pt offers several potential advantages as an electrode material compared to Au: (1) It is more compatible with Si fabrication technology than Au, (2) Pt seems to be preferred by the chemists for DNA probe attachment, and (3) Pt is more resistant to corrosion and electrolytic etching effects than Au in the rather hostile, conductive aqueous buffer solutions important to DNA chemistry.

There are some disadvantages to using Pt for electrode fabrication, however. Because there are no good resist-compatible etches, electron-beam evaporation and liftoff must be used for patterning of the cell electrodes rather than an etching process. We have found that Pt evaporations generate a fair amount of heat owing to the high melting temperature of Pt. This adversely affects the photoresist which makes liftoffs difficult and necessitates additional cleaning procedures to ensure reliability of DNA attachment chemistry. Recent device wafers using reduced Pt-evaporation rates and additional resist postbaking have resulted in much improved device yield and surface cleanliness. Pt also has problems related to stress which limit the metal thickness that can be used and Pt is more resistive than Au (which may affect signal sensitivity—this topic is still under investigation but is probably not a problem).

An additional problem that we have encountered is breakage in the Pt interconnect lines that connect the bonding pads at the periphery of the chip to the biosites clustered in the center. Initially, we believed that this problem was in part due to the rather long and narrow (3-μm wide) nature of these lines. Further investigation and mapping of several
device wafers show that the breakages seem to be clustered in consistent regions from wafer to wafer, yet there appear to be no obvious problems on the photolithographic masks. This is still being investigated. In the meantime, we are using our direct-write laser metal-deposition capability to repair the broken lines with very good success. Pt is our preferred laser-write material, another advantage for using Pt for genosensors.

Over this time period, we have delivered 14 packaged device chips to HARC for this project. At present, several more wafers are in various stages of processing.

1.2.3 Genosensor Test Results

The results of electrical tests on a genosensor test cell are illustrated in figure 4. In this experiment, 12mer DNA probes (12 nitrogenous bases in a single strand) were attached to the Pt electrodes of a test well. Permittivity measurements were recorded using an impedance bridge connected to the leads of the test site via the genosensor package shown in figure 3. Next, high molecular weight (4000 base pairs) target DNA was puddled on the electrodes and the measurements repeated after approximately 50 minutes. Complementary target and probe DNA strands were used in this experiment to insure that hybridization occurred. From the measurements, the dissipation factor as a function of frequency was extracted for both the nonhybridized-probe case and the hybridized case (probe plus target). The nonhybridized probes exhibited maximum energy absorption at 800 kHz, while the hybridized DNA displayed maximum absorption at 80 kHz, a factor of 10 in frequency discrimination. Further, the nonhybridized probes exhibited a dissipation factor of 20 at the peak, while the corresponding value for the hybridized case was 4. These results, though preliminary, are very encouraging as they indicate that sufficient discrimination between hybridized and nonhybridized DNA can be achieved. For further discussion see section 3 below.

1.3 Microdetection Technology for Automated DNA Sequencing

Sponsor
Houston Advanced Research Center
Contract HRC-HG00776-01

Project Staff
Dr. Daniel J. Ehrlich, Dr. Dennis D. Rathman, Dr. Mark A. Hollis, Dr. John Melngailis

Figure 4. Measured dissipation factor data from a test site containing hybridized probe and target DNA plotted for comparison with data from a site containing nonhybridized-probe DNA.

The primary emphasis of this effort is to complement the program described above by optimizing the electrode geometry for maximum sensitivity. This task is being undertaken using a two-pronged experimental and theoretical approach. In the experimental approach, the electrode-finger geometry is systematically varied by varying the finger linewidth or by placing the opposing electrodes on the same horizontal plane (see figure 5 and compare to figure 2). This is in contrast to the cell-geometry variations with a fixed electrode finger width of 2 \( \mu m \) described in the previous section. Assuming that the DNA attachment chemistry is not significantly affected by these electrode variations, measurements of dissipation factor, loss, relative permittivity, and the relaxation frequency should yield direct information on relative sensitivity. During the past year we have supplied 12 packaged chips to HARC of this type for testing. Many of these chips are currently being tested. In the theoretical approach, computer modeling is used to determine the relative sensitivities of electrode geometries using known or measured properties of the attached DNA layers. The combination of the electrical measurements and computer modeling should allow us to optimize the electrode geometry and also provide the insight necessary to develop a physical model for the dispersion effects exhibited by DNA in aqueous solutions.

During this past year an electromagnetic-simulation package was purchased from ANSOFT that solves Maxwell's equations in two dimensions. Recently, this software has been upgraded to handle materials with complex permittivities such as DNA. Figure 6 shows the electrode geometry in cross section and some equipotential contours from a solution generated by ANSOFT with the metal fingers
appropriately biased. Note the hybridized DNA layer attached to the metal electrodes. The input parameters for the model are the cell geometry, applied bias, the real and imaginary parts of the permittivities for all of the materials, and the signal frequency. The simulation yields the complex admittance at the electrodes modeled as a parallel resistance, $R_p$, and capacitance, $C_p$, for the genosensor cell. These values can in turn be used to calculate $\varepsilon', \varepsilon''$, and the dissipation factor from the following relations:

$$C_p = \varepsilon' C_0$$

$$\varepsilon'' = 1/\omega R_p C_0$$

$$\varepsilon''/\varepsilon' = 1/\omega R_p C_p$$

where $C_0$ is the capacitance between the electrodes in vacuum. The parallel $R_p$ and $C_p$ obtained in this way yield a simplified equivalent circuit for the genosensor cell, which in turn can be used to simulate the behavior of the cell in a circuit environment that may include switching transistors.

The most immediate use for the modeling results, however, is to infer the dispersion characteristics for the DNA in solution by comparing the calculations with the experimental results. The experimental data shown in figure 4 were obtained for the probe and hybridized DNA using the electrode geometry of figure 5, as described in the previous section. Assuming a 4000-Å-thick DNA layer for the hybridized case, a dissipation factor of 4 (in agreement with the experimentally obtained value at a signal frequency of 80 kHz) can be obtained with a peak value of $\varepsilon''$ of 900 for this geometry. This value of $\varepsilon''$ is very consistent with literature values for similar DNA. However, for the nonhybridized case, where a dissipation factor of 20 was measured, the experimental data cannot be fit by any simple adjustment of DNA permittivity or loss in the model. Such a large loss mechanism for the nonhybridized-probe case might be explainable by a preferential and ordered alignment of the short negatively-charged DNA probes on the electrode surfaces. These probes may act as a regular array of high-Q mechanical oscillators efficiently absorbing energy from the alternating electric field at a characteristic resonant frequency. An alternative explanation might be related to the electric dipole formed by the negatively charged DNA probes and the positive ions in solution. This dipole layer, as it resonates with the applied AC field, could in turn modulate the conductivity of the buffer solution in the vicinity of the electrodes, giving rise to the characteristic peak. In any case, this phenomenon is currently under intense investigation.

### 1.4 Publications


Part V Language, Speech and Hearing

Section 1 Speech Communication
Section 2 Sensory Communication
Section 3 Auditory Physiology
Section 4 Linguistics
Section 1  Speech Communication

Chapter 1  Speech Communication
Chapter 1. Speech Communication

Academic and Research Staff
Professor Kenneth N. Stevens, Professor Jonathan Allen, Professor Morris Halle, Professor Samuel J. Keyser, Dr. Melanie Matthes, Dr. Joseph S. Perkell, Dr. Stefanie Shattuck-Hufnagel, Dr. Mario A. Svirsky, Dr. Alice Turk, Dr. Yi Xu, Peter C. Guido, Seth M. Hall, Joyce Manzella

Visiting Scientists and Research Affiliates
Dr. Vladimir M. Barsukov, Dr. Corine A. Bickley, Dr. Suzanne E. Boyce, Dr. Carol Y. Espy-Wilson, Dr. Richard S. Goldhor, Dr. Robert E. Hillman, Dr. Eva B. Holmberg, Dr. Caroline Huang, Dr. Harlan Lane, Dr. John Locke, Dr. John I. Makhoul, Dr. Sharon Y. Manuel, Dr. Takahide Matsuoka, Dr. Jorgen Pind, Dr. David Williams, Anna Esposito, Giulia Arman Nassi, Roeland P. Schaeffer, Jane Wozniak

Graduate Students
Hwa-Ping Chang, Marilyn Y. Chen, Helen M. Hanson, David M. Horowitz, Mark A. Johnson, Hong-Kwang Kuo, Sharlene A. Liu, Noel S. Massey, Kelly L. Poort, Lorin Wilde

Undergraduate Students
Howard Cheng, Leanne Clarke, Laura C. Dilley, Helena Geng, Brett A. Geoffroy, John Hardy, Marnie Harker, Peggy Li, Henry Y. Lu, Rebecca Mallin, Rachel Molinaar

Technical and Support Staff
Michelle Halverson, D. Keith North, Arlene E. Wint

1 Audiofile, Inc., Lexington, Massachusetts, and Boston University, Boston, Massachusetts.
2 Boston University, Boston, Massachusetts.
3 Massachusetts Eye and Ear Infirmary, Boston, Massachusetts.
4 MIT and Massachusetts Eye and Ear Infirmary, Boston, Massachusetts.
5 Dragon Systems, Inc., Newton, Massachusetts.
6 Department of Psychology, Northeastern University, Boston, Massachusetts.
7 Massachusetts General Hospital, Boston, Massachusetts.
8 Bolt, Beranek and Newman, Cambridge, Massachusetts.
9 Wayne State University, Detroit, Michigan.
10 Utsunomiya University, Utsunomiya City, Japan.
11 University of Iceland, Reykjavik, Iceland.
12 Sensimeetrics, Inc., Cambridge, Massachusetts.
13 Department of Physics, Salerno University, Salerno, Italy.
14 University of Utrecht, The Netherlands.
1.1 Introduction

The overall objective of our research in speech communication is to gain an understanding of the processes whereby (1) a speaker transforms a discrete linguistic representation of an utterance into an acoustic signal, and (2) a listener decodes the acoustic signal to retrieve the linguistic representation. The research includes development of models for speech production, speech perception, and lexical access, as well as studies of impaired speech communication.

Sponsors

C.J. Lebel Fellowship
Dennis Klatt Memorial Fund
National Institutes of Health
Grant R01-DC00075
Grant P01-DC00361-06A1
Grant R03-DC01721
Grant R01-DC01291
Grant R01-DC00261
Contract R01-DC00776

National Science Foundation
Grant IRC 89-05249
Grant IRC 89-10561
Grant INT 90-24713

1.2 Studies of the Acoustics, Perception, and Modeling of Speech Sounds

1.2.1 Nasal Vowels and Consonants

Some further experimental data on the acoustics of nasalization have been collected, and these data have contributed to a refinement of the acoustic theory of nasalization. An experimental approach has been developed for measuring the natural frequencies of the nasal cavity when it is closed at the posterior end. The natural frequencies are determined by measuring the response of the nasal cavity when a transient excitation is produced by rapidly closing the velopharyngeal port. In accordance with previous findings, two of the principal natural frequencies are in the range 490-640 Hz and 1880-2400 Hz for different individuals. Additional natural frequencies, presumably due to sinuses, are observed at 210-250 Hz and at frequencies between the principal resonances. Acoustic analysis of nasalized vowels in English and French shows that the lowest sinus resonance contributes to an enhancement of the spectrum amplitude at low frequencies. This increased low-frequency amplitude has been observed to be as great as 10 dB for French nasal vowels. This enhanced low-frequency amplitude is an additional attribute that contributes to the decreased first-formant prominence that occurs when a vowel is nasalized.

1.2.2 Lateral Consonants

In earlier studies of lateral consonants, acoustic events at the release and closure of these consonants were examined. These acoustic data showed the nature of the discontinuity in amplitudes and frequencies of spectral peaks, particularly the abrupt change in second-formant amplitude due to bandwidth changes associated with acoustic losses during the lateral. This study has been extended to include syllabic laterals, which contrast with the unstressed vowel [o] in word pairs like bucko-buckle. Measurements show similar values for the first two formant frequencies F1 and F2 for the two syllabic nuclei, with some spectral differences in the F3-F4 region which do not appear to be consistent. The most consistent difference between the lateral and the vowel is the F2 bandwidth. Informal listening with synthesized utterances confirms that the increased F2 bandwidth shifts the identification from the vowel to the lateral. More extensive analyses and listening tests are in progress.

1.2.3 Formant Transitions for Fricative Consonants

Measurements have been made of the transitions of the second and third formant frequencies (F2 and F3) in vowels following the release of different voiceless and voiced fricative consonants produced by four speakers. This is part of a larger study of the acoustics, perception, and modeling of fricative consonants. The data show that the range of F2 starting frequencies across different following

---

15 Under subcontract to Massachusetts Eye and Ear Infirmary.
16 Under subcontract to Massachusetts General Hospital.
17 Under subcontract to Boston University.
18 U.S.-Sweden Cooperative Science Program.
Chapter 1. Speech Communication

1.2.6 Preaspirated and Postaspirated Stops in Icelandic

We have examined the perception of two speech cues in Icelandic, Voice Onset Time (VOT) as a cue for word-initial aspiration of stops, and Voice Offset Time (VOffT) as a cue for preaspiration. These speech cues contrast in an interesting manner in perception though signaled by acoustically identical cues. The studies replicate the well-known finding that place of articulation has an effect on the placement of the VOT boundaries separating unaspirated from aspirated stops. No such effect was found for the perception of VOffT. Both cues are sensitive to the duration of the vowel carrying the relevant segments, VOffT much more so than VOT. While the effect of vowel duration on the perception of VOT has usually been attributed to the effect of rate normalization, in the case of VOffT it emerges that the ratio of VOffT at the phoneme boundary to the duration of the vowel acts as a higher-order invariant, thus obviating the need to posit a mechanism of rate normalization. The differences in the behavior of these speech cues is related to their different positions in the syllable.

1.2.7 Acoustic and Articulatory Studies of Stop Consonant Production.

The production of a stop consonant such as [p] requires that the articulator that makes the consonant closure be released rapidly, so that an abrupt onset of acoustic energy is achieved. Data on the movement of the jaw and the lips have been obtained using an electro-magnetic midsagittal articulometer (EMMA) to track movements of points on these structures during the production of stop consonants in various phonetic environments. Simultaneous recordings of the sound output were made. Data for the consonant [p] in utterances like [espV] or [epV] (where V = [i] or [a]) show that the downward movements of the jaw and lower lip appear to be coordinated to achieve maximum rate of lip opening at the time of release, in spite of the requirement that the jaw remain high during a preceding [s] in order to enhance turbulence noise generation. In general, the total time interval over which the stop consonant has an influence on the movement patterns is in the range 150-200 ms.
1.2.8 Individual Differences in Laryngeal
Waveforms for Female Speakers

The aim of this project is to develop procedures for characterizing the laryngeal behavior of individual speakers based on measurements on the speech waveform. A number of measurements were made in several vowels produced by several different speakers. The measurements are intended to provide information about the glottal area waveform and hence about the configuration and state of the glottis during modal vibration. The parameters include various measurements of spectral tilt, first-formant bandwidth, and degree of acoustic coupling through the glottis to the trachea. Substantial individual differences in these parameters have been observed, and some partial interpretations in terms of average glottal opening have been made. Further studies are needed if a compact set of descriptors for individual speakers is to be developed.

1.3 Studies of Normal Speech
Production

1.3.1 Goal-based Speech Motor Control: A
Theoretical Framework and Some
Preliminary Data

A theoretical framework for the segmental component of speech production has been outlined and some preliminary supporting data have been presented. According to the framework, articulatory movements are programmed to achieve sequences of articulatory and acoustic goals. The goals are defined partly by characteristics of speakers that enable them to produce and perceive speech sounds that have distinctive acoustic properties; these characteristics are correlates of distinctive features. Some feature correlates are determined by quantal (non-linear) relations between articulation and sound. Goal definitions also may be partly language-specific, according to other principles such as a compromise between sufficient perceptual contrast and economy of articulatory effort. When utterances are produced, goal specifications are modified by prosodic influences and reduction. Based on the sequence of modified goal specifications, articulatory movements are timed appropriately and smoothed by the control mechanism and the biomechanical properties of the articulators. To help keep acoustic variability within perceptually-acceptable limits, speech production mechanisms include a strategy of "motor equivalence", which takes advantage of the fact that for some sounds, a similar acoustic transfer function can be achieved with somewhat different area functions. This hypothesis is supported tentatively by data on motor equivalence at the level of the transformation between the vocal-tract area function and the acoustic transfer function.

1.3.2 Factors Underlying Findings of Motor
Equivalence in the Transformation Between
the Vocal-tract Area Function and the
Acoustic Transfer-function

Using measurements made with an electromagnetic mid-sagittal articulometer (EMMA), we have found evidence of trading relations between lip rounding and tongue-body raising for multiple repetitions of the vowel /u/ in seven male speakers of American English. In other words, when the tongue body is less raised, the lips are more rounded and vice versa. (An eighth speaker did not show such a trading relation.) As mentioned above, these findings support the hypothesis that there is a "motor equivalence" strategy which functions at the level of the transformation between the vocal-tract area function and the acoustic transformation. For all of the subjects, examination of subsets of data defined by a measure of acoustic distance from a normative mean has shown that there is an increasing likelihood of negative correlations for tokens that are less /u/-like. For a majority of the data sets, tokens that are less /u/-like are produced at a faster rate. Thus, this strategy could help to keep acoustic variation within perceptually-acceptable limits, and it seems to come into play more when it might be needed, to prevent production of a sound that could lead to the wrong percept.

1.3.3 Speaking Rate and Associated
Kinematic Patterns of Tongue Movements

Kinematically-based measurements of effort on tongue movements were examined for a speaker with a cochlear implant. Turning the speech processor of the implant on and off induced changes in the subject's speaking rate. Kinematic measurements were made of EMMA data on tongue-body movements toward and away from the steady-state target for the vowel /u/ in the utterance "... who hid ...". Although the movements toward the target were quite small—on the order of a millimeter—the kinematic measurements still exhibited relationships that were strikingly similar to previous results of others: there was a strong positive relation between movement distance and duration, and values of an index of velocity profile shape (the ratio of peak velocity to average velocity) were positively related to values of movement duration.
According to measurements of peak velocity and effort, the most effort was exerted in the condition when the cochlear implant had been turned off for the longest time.

### 1.4 Speech Research Relating to Special Populations

#### 1.4.1 Speech Production of Cochlear Implant Patients

This project is funded under a subcontract to the Massachusetts Eye and Ear Infirmary in collaboration with Drs. Harlan Lane, Donald Eddington and Joseph Nadol.

This research aims to characterize the speech production of postlingually deafened adults before and after they receive cochlear implants. One goal of the work is to contribute to the development of models of the role of hearing in speech. The findings can also contribute to evaluating and improving prostheses and to focusing speech therapy for implant users. Toward these goals, we have been making measurements of (1) vowel acoustics (vowel formants, F0, SPL, duration, and inferred glottal aperture); (2) consonant acoustics (voice-onset time, fricative and plosive spectra); (3) speech aerodynamics (speech initiation and termination lung volumes, volume of air per syllable, inferred subglottal pressure, oral and nasal airflow, peak glottal airflow, minimum flow, AC flow, open quotient, and maximum flow declination rate); and (4) tongue, lip and jaw trajectories. A major concern has been to separate "phonemic" changes in these parameters from changes due to "postural" adjustments in overall speaking rate, SPL, and F0.

#### Production of the Plosive Place Contrasts

To date, we have examined the spectra of plosive bursts of four implant users before and after activation of the speech processors of their cochlear implants. The spectral slopes of the bursts (dB/octave) in all three places of articulation did not change reliably with activation. Comparison with normative data suggests that these speakers had normal spectral slopes for bilabial and alveolar positions, but more compact spectra for the velar place of articulation.

### Production of Fricative Place Contrasts

Prior to activation of their cochlear implants, two deafened adult speakers produced the alveolar and palatal voiceless fricatives with spectral parameters more similar to one another than in normal hearing individuals. With the restoration of some hearing, these speakers differentiated these productions more normally. When their implant speech processors were turned off and on for short periods, they reduced and increased spectral differentiation correspondingly.

We have conducted an initial experiment to explore the articulatory basis of changes in fricative spectra with prosthetic hearing and to assess whether they are due to selective control by prosthetic hearing or a postural change in rate or both. With another implant user, an electro-magnetic midsagittal articulometer system was used to record movements of the tongue, lips and jaw. Preliminary analysis showed that tongue blade positions during the palatoalveolar fricative /sh/ changed with access to hearing; moreover, the shift in place of articulation directly correlated with the observed shift in the fricative spectrum between processor on and off conditions, a shift that sharpened the fricative place contrast with the processor on.

#### Voicing Onset Time

Implant users lengthen voice-onset time (VOT) toward normative values with processor activation. However, VOT changes may reflect postural adjustments. VOT and syllable duration have been measured for the plosives spoken by five patients pre- and post-activation of their implant speech processors over a period of several years. Pre-implant, all five speakers characteristically uttered voiced plosives with too-short VOT, compared to normative data. VOTs of voiceless plosives were also abnormally short for three of the speakers, and close to normal for the remaining two. With some hearing restored, implant users made relatively few errors with respect to voicing when identifying plosives in listening tests. Three of the five speakers lengthened VOT, and three increased the contrast. The findings are consistent with the hypothesis that deafened speakers reset articulatory routines when some hearing is restored and they can hear the relevant voicing features. However, this study controlled for the indirect contribution to VOT changes of changes in only one postural variable, speaking rate. In the case of the voiced plosives, multiple correlations between changes in the postural variables of rate, a measure of breathiness and SPL and changes in VOT were large and significant, suggesting that much, if not all, of the VOT changes could be attributed to the mediation of
postural changes. In the case of the voiceless plosives, however, multiple correlations were small and could account for only part of the observed increases in VOT, suggesting that the speaker increased VOT partly under the control of his or her renewed ability to hear VOT due to the implant.

Our findings to this point continue to indicate significant deterioration of acoustic and articulatory speech parameters in adults after becoming deaf, and substantial long-term changes, generally in the direction of normalcy, following activation of the cochlear implant.

1.4.2 Degradation of Speech and Hearing with Bilateral Acoustic Neuromas

In this project, we are studying the relation between speech and hearing in people who become deaf from bilateral acoustic neuromas (auditory nerve tumors) associated with the genetic disorder neurofibromatosis type II (NF2). The primary goal of the study is to increase our understanding of the role of hearing in the control of adult speech production. The rationale and approach for this work are similar to those of our ongoing work on the speech production of cochlear implant patients. Speech acoustic and physiological parameters are recorded and speech perception is tested in a group of (still hearing) NF2 patients. Then the same parameters are recorded at intervals for any patients who suffer further hearing loss. Thus far, we have obtained baseline recordings from 40 patients who still have some hearing.

Two of the enrolled patients have suffered further hearing loss, and we have obtained post-loss recordings on both of them. We have analyzed the pre- and post-loss recordings for one of the two. This patient had a mild-to-moderate hearing loss in one ear (aided) and was deaf in the other due to bilateral acoustic neuromas. She became deaf bilaterally when the second of these neuromas was surgically removed. Following this surgery, an auditory brainstem implant provided minimal speech reception benefit. Three speech production recordings made prior to deafening were compared to recordings made 11 and 35 weeks post-deafening. Results indicated significant changes in voice onset time (VOT), fundamental frequency (F0), F0 range, and vowel duration. VOTs, adjusted for changes in syllable duration, were significantly shorter for the voiceless plosives post-deafening. F0 was at the mean of a normative range prior to deafening, and rose significantly above this range post-deafening. An analysis of prosody indicated that the subject placed pitch accents on the same syllables post-deafening, but F0 mean and s.d. were significantly higher. Vowel formants, /s-sh/ contrast, an indirect measure of breathiness, and overall SPL, showed no significant changes by 35 weeks post-deafening.

1.4.3 Objective Assessment of Vocal Hyperfunction

This project is funded under a subcontract to the Massachusetts Eye and Ear Infirmary in collaboration with Dr. Robert E. Hillman.

The major goal of this project is to further develop and use quantitative measurements of voice production to provide objective descriptions of conditions referred to as vocal hyperfunction. This work involves the use of noninvasive acoustic, aerodynamic and physiologic measures to study both organic (nodules, polyps, contact ulcers) and non-organic (functional dysphonia/aphonia) manifestations of vocal hyperfunction. These measurements are examined in relation to encoded descriptions of the pathophysiologic status of the larynx (from videolaryngoscopy/stroboscopy), pertinent medical and psychosocial factors and judgments about perceptual characteristics of the voice. Interpretation of results will be aided through use of a simple model of vocal-fold vibration.

Expansion and Correction of the Normative Database

We have gathered and completely analyzed data for 30 additional normal speakers (15 males and 15 females) using updated analysis methods. Statistical tests comparing previous (1988) and new (1993) normative data indicated overall group differences. For most individual parameters, differences between the two groups were relatively small (less than a 10 percent change) and not statistically significant. Such variation could be attributed to minor methodological differences and sampling effects, i.e., most of the new average values fell well within one standard deviation of the previous averages. However, for the parameter maximum flow declination rate (MFDR), there were instances (mostly for loud voice) of significant differences between the 1988 and 1993 data. These differences could be attributed largely to changes (corrections) in the signal processing algorithms that are used to obtain glottal waveform measurements. The new MFDR values more closely match predictions based on speech acoustic theory and are therefore believed to be more valid.
Respiratory Function Associated with Vocal Nodules in Females

We have completed the collection of respiratory data for 10 females with vocal nodules and 10 matched (sex, age, body type) normal controls. Preliminary results indicate that, compared to matched controls, nodules subjects display reduced efficiency of respiratory function for speech production by operating at lung volume levels which require greater muscular effort. Some of the nodules subjects also tended to stop at inappropriate (non-juncture) points within sentences to inhale and demonstrated oppositional chest wall displacements after the initiation of exhalations.

Modeling of Vocal Fold Vibration

A mathematical model of vocal fold vibration has been developed during this last year. We started with the well known Ishizaka-Flanagan two-mass model of the vocal folds. The physical system is described mathematically by a set of nonlinear differential equations with singular coefficients which represent the self-oscillating voice source composed of two stiffness-coupled masses. Also included are lumped approximations for supra- and sub-glottal vocal-tract loading and lip radiation characteristics. The fourth-order-Runge-Kutta numerical method for solving differential equations is applied because it gives more precise results than the Euler method used in Ishizaka and Flanagan's original model.

Treatment Efficacy Studies of Voice Therapy Protocols

Voice therapy protocols for use in our proposed treatment efficacy studies have been developed. The protocols include: (1) detailed step-by-step procedures for each behavioral approach being used (vocal hygiene, direct facilitation, relaxation, respiration, carryover, home practice), (2) daily behavioral charts for documenting voice use outside of the therapy sessions, (3) home practice records (verified by spot checking cassette recordings of home practice sessions), (4) response charts for documenting patient performance in each therapy session, (5) scales for clinician's ratings of patients' motivation and vocal status at each therapy session, and (6) a scale for patients' self evaluation of their vocal status at each therapy session. We have completed making repeated longitudinal recordings under this protocol with one patient and are progressing with several others. Data are currently being analyzed.

Development of a Multi-user Relational Database

Significant progress has been made in the development of a multi-user relational database application ("VoiceBase") which will provide improved capabilities for the storage, retrieval, and reporting of subject background (history), perceptual, acoustic, aerodynamic, electroglottographic, and videostroboscopic project data.

1.4.4 Aiding Dysarthric Speakers

We have been collecting data on the speech patterns of several dysarthric speakers. One objective of this research is to develop procedures for using speech recognition devices to augment the communication abilities of these individuals. The speech of each speaker is evaluated through (1) listening tests, (2) tests with a speech recognizer, and (3) acoustic analysis of a series of words produced by the speaker. An outcome of this evaluation is a description of attributes of vowels and consonants which are produced adequately and consistently and those which are produced in a distorted or inconsistent manner. Each of the speakers who have been evaluated show somewhat different patterns of deficiencies, but there are some common problems, particularly with regard to fricative and lateral consonants that require shaping of the tongue blade. These kinds of evaluations provide a basis for selecting words that are likely to be recognized consistently by a speech recognizer.

1.5 Speech Production Planning and Prosody

Our work in speech production focuses on both segmental and prosodic aspects of the planning process and their interaction. To make these studies possible, we are continuing the development of a prosodic database with our colleagues Mari Ostendorf of Boston University and Patti Price of Stanford Research Institute. With the help of students in the Undergraduate Research Opportunities Program at MIT, using the ToBI transcription system recently proposed, we have labeled the pitch accents, constituent boundaries and boundary tone markers for almost one hour of FM radio news style speech produced by two professional speakers. The resulting database, which also includes part-of-speech labels and automatic phonetic alignments (10 percent hand corrected), will be made available to the speech research community through the Linguistics Data Consortium at the University of Pennsylvania. We are participating in the creation of an online training hand-
book for the ToBI transcription system, being prepared at Ohio State University, which includes audible labeled examples. We are also involved in evaluating the reliability of the system across 30 transcribers. In addition, we are training several new transcribers, who will then participate in both the ongoing creation of the FM news database and the labeling of utterances containing speech errors in the new MIT Digitized Speech Error Database (see below).

Using the FM news database, we have carried out three studies of the role of prosody in speech planning. First, we found that the apparent stress shift (in which the first syllable of late-main-stress words like "Massachusetts" is perceptually more prominent than the main stress syllable, in phrases like "the Massachusetts senator") often results from the placement of a early prenuclear pitch accent on the first accentable syllable of a new intonational phrase and the loss of pitch accent from the main-stress syllable. These findings suggest that speakers and hearers may be able to use the occurrence of early accent within the word as a cue to the onset of a new intonational phrase. Second, we found that alternating-stress words (like "Massachusetts" and "university") are pitch-accented differently from adjacent-stress words (like "campaign" and "Hongkong"): alternating-stress words are more likely to be double-accented and are more reliably early-accented in phrase-initial contexts. In addition, both intuitions and dictionary markings of the placement of lexical stress are less consistent for adjacent-stress words. These results suggest that speakers are reluctant to place pitch accents on adjacent syllables within a word, just as they are for syllables across word boundaries.

Then we analyzed the occurrence of glottal onset in more than 400 vowel-initial syllables and found that it is most likely to occur when the syllable begins an intonational phrase and/or is pitch accented (81 percent, versus 16 percent for non-phrase-initial, non-accented syllables.) The pattern of phrase-initial glottal onsets was maintained even for syllables which begin with reduced vowels, which cannot be pitch accented. This finding suggests that glottalized onset of these syllables may provide a useful cue to the beginning of an intonational phrase. Studies are currently under way to evaluate whether phrase-initial glottal onsets are associated with low F0, post-pausal position or certain specific segment types.

Finally, we are developing and prosodically labeling a second database of utterances containing segmental and word-level speech errors, in order to investigate prosodic constraints on disfluencies. Error utterances are being harvested from a number of existing digitized speech databases, and prosodic transcribers are being trained in our laboratory. We anticipate being able to label and analyse 300-500 error utterances over the next year. We asked the questions (1) what kinds of error corrections speakers mark prosodically and (2) how prosodic factors such as pitch-accent prominence and constituent boundaries constrain the pattern of segmental and morphemic errors. Results can be expected to shed light on the question of what prosodic representations speakers employ during the production planning process and how those representations guide segmental planning.

### 1.6 Models for Lexical Representation and Lexical Access

We are developing a model for lexical access in which items are stored in the lexicon in terms of arrays of features, and procedures for estimating these features from the acoustic signal are specified. The estimation of feature arrays from measurements on the signal proceeds in three steps: (1) detection of landmarks in the signal where a major articulator forms or releases a narrow constriction in the vocal tract, or forms a minimum or maximum opening for a consonant or vowel; (2) making acoustic measurements in the vicinity of the landmarks to estimate the articulator-free and articulator-bound features being implemented; and (3) organizing the feature estimates into a form that can be used to match against a feature-based lexicon. During the past year, our efforts have been directed towards detection of landmarks, developing a framework for estimating landmarks, and hand-labeling a database of sentences in terms of landmarks and features, for testing the adequacy of our analysis procedures.

#### 1.6.1 Identification of Landmarks

Two types of landmarks are created by the formation of a narrow consonantal constriction in the vocal tract. For one type, the articulator forming the constriction makes a complete closure in the midline of the vocal tract, whereas for the other type, the closure is only partial. The first of these is further subdivided into landmarks where there is no increase or decrease in intraoral pressure across the landmark (i.e., nasals and laterals), and where there is a change in intraoral pressure.

The detection of these landmarks is accomplished by examining the rate of change of energy in different frequency bands. The bands that are selected, the averaging time for determining energy in a band, and the time increment for obtaining a
first difference are all adjusted to best suit the type of landmark being sought. The general procedure is to first scan the signal with fixed filters, averaging times, and increments in order to make an initial selection of candidate landmarks, and then to examine each of these landmarks with a more finely tuned set of measurements. The algorithms that have been developed miss a few landmarks, and steps are being taken to refine the algorithms, including the measurement of time variation of first-formant prominences. Procedures for detecting regions of minimum vocal-tract opening for glides are examining changes in low-frequency amplitude and looking for minima in first formant frequency coupled with extrema in other formant trajectories.

### 1.6.2 Estimation of Features Around Landmarks

A framework is being developed for estimating the features involved in generating a given landmark. The acoustic properties that must be measured to provide cues for some features are relatively context-independent, whereas for other features the properties are dependent on the context of features that co-occur with the feature or that are associated with adjacent segments. A program is being developed to guide the set of context-dependent measurements that are required to estimate a feature. The program is given a query feature (a feature to be determined) and a set of context features, and it specifies a list of measurements to be made to determine the value of the feature. The set of measurements that is required for a given feature is determined in part from theoretical models of speech production and in part through examination of a database of utterances. They include measurements of formant trajectories near landmarks, gross spectrum shape and its variation with time, and degree of spectrum prominence. The program is being tested for a limited set of features using some databases, including one under development in our laboratory.

### 1.6.3 Feature-labeled Database

 Portions of database of 100 sentences spoken by four speakers have been hand-labeled to indicate the locations of landmarks, and the features being implemented at each landmark are listed. The feature estimates are based on careful acoustic analysis of the sound in the vicinity of each landmark and on listening to segments of the signal. This labeled database is to be used for two purposes: (1) it can serve as a standard against which various algorithms for estimating landmarks and features can be evaluated; and (2) it can provide data indicating which features in the spoken words in context differ from features in the lexicon. The second of these uses of the database can help to guide the refining of rules for describing modifications of feature patterns that occur in casual speech.

### 1.7 Publications

#### 1.7.1 Papers Published


#### 1.7.2 Papers Accepted for Publication


Perkell, J., R. Hillman, and E. Holmberg. "Group Differences in Measures of Voice Production and Revised Values of Maximum Airflow
Chapter 1. Speech Communication


1.7.3 Papers Submitted for Publication


Section 2  Sensory Communication

Chapter 1  Sensory Communication
Chapter 1. Sensory Communication

1.1 Introduction

The Sensory Communication Group is conducting research on (1) the auditory and tactual senses, (2) auditory, visual, and tactual aids for individuals who are hearing-impaired or deaf, and (3) human-machine interfaces for teleoperator and virtual-environment systems (involving the visual as well as the auditory and tactual senses). Within the domain of hearing aids, research is being conducted on systems that bypass the outer and middle ear and directly stimulate the auditory nerve electrically (cochlear prostheses), as well as on systems that stimulate the ears acoustically. The research on taction is focused not only on speech reception for the totally deaf, but also on the ability of the human hand to sense and manipulate the environment. Within the domain of human-machine interfaces, topics of special interest concern (1) development of principles for mapping the human sensorimotor system into nonanthropomorphic slave mechanisms (or the equivalent in virtual space) and (2) ability of the human sensorimotor system to adapt to alterations of normal sensorimotor loops caused by the presence of the interface.

1.2 Hearing Aid Research

Sponsor
National Institutes of Health
Grant 5 R01 DC00117

Project Staff
Professor Louis D. Braida, Lorraine A. Delhome, Dr. Kenneth W. Grant, Dr. Christine M. Rankovic, Dr. Charlotte M. Reed, Dr. Kaoru Sekiyama, Dr. Rosalie M. Uchanski, Dr. Patrick M. Zurek, Bridget L. Coffman, Paul Duchnowski, Joseph A. Frisbie, Andrew L. Howe, Michael T. Keagy, Kevin C. Knoedler, Irene Kotok, Danielle G. LeMay, David S. Lum, Jason Mueller, Philip P. Nadeau, Erik Nygren, Charlie Pan, Nicholas Ploch, Diane E. Ronan, Andrew Ugarov, Evan F. Wies, Jennifer A. Wozniak

During the past year, our research on hearing aids has focused on (1) simulation of sensorineural hearing impairment, (2) basic studies of speech intelligibility, (3) speech production factors, (4)
Chapter 1. Sensory Communication

control of acoustic feedback, and (5) aids to speechreading.

The following is a description of our research in these areas.

1.2.1 Simulation of Sensorineural Hearing Impairment

Our previous work employed noise masking to simulate the effects of hearing loss. Zurek and Delhome,1 as well as Humes et al.2 and Dubno and Schaefer,3 found good agreement between masked-normal and hearing-impaired listeners’ reception of consonants. Together with other findings,4 these results support the view that the main difficulty in speech reception comes from audibility restrictions, as opposed to supra-threshold disorders as suggested by others,5 at least for listeners with mild and moderate hearing losses.

While noise masking provides functionally accurate simulations of hearing impairment, it is phenomenologically unrealistic4 and can be used only to simulate losses of less than roughly 70 dB because of the intense noise levels needed. We have developed and evaluated an alternate technique based on multiband expansion amplification7 that simulates both threshold shift and recruitment without the use of masking noise. This simulation, a digitally-implemented elaboration of Villchur’s scheme,8 is completely specified by the audiogram of a particular hearing-impaired listener and empirical relations specifying the steepness of the recruitment function for a given hearing loss.

In initial evaluations, the performance of two normal-hearing subjects listening to the processed output of the simulator matched fairly well the consonant-reception scores of both the hearing-impaired and noise-masked normal listeners who had been previously tested by Zurek and Delhome.9 In more extensive evaluations, the performance of each of three severely hearing-impaired subjects on consonant and sentence intelligibility tests was compared to that of three normal-hearing subjects listening through the simulator. Generally, the simulation reproduced overall intelligibility scores and error patterns for individual impaired listeners very well when the frequency response was flat. When high-frequency emphasis placed more speech energy above threshold at high frequencies, the simulation gave better intelligibility than observed for two of the three impaired listeners. Overall, this study verified the equivalence of the AGC and noise simulations, explored the applicability of the AGC simulation for severe impairments, and identified frequency-specific deficits not accounted for by audibility and recruitment in some of the listeners.

---

1.2.2 Basic Studies of Speech Intelligibility

Although many current theories of speech perception assume that the perception of speech elements involves accurate cross-spectral comparisons, studies of spectral shape perception have revealed limits on our ability to make such comparisons. To gain insight into the role of cross-spectral comparisons in speech perception, we are studying the ability of normal-hearing listeners to combine cues across frequency bands when identifying speech segments. This research involves both speech identification experiments and analysis of the results of these experiments using the pre-labeling and post-labeling models\(^\text{10}\) that were initially applied to studies of audiovisual integration.

When bands arising from lowpass and highpass filtering of consonants are considered,\(^\text{11}\) the situation is similar to the audiovisual case: the pre-labeling models give a good account of the combined band (i.e., wideband) score, while the post-labeling model underestimates it. To evaluate the models’ abilities to predict cross frequency-band integration, we undertook a series of consonant-identification experiments in which the stimuli were filtered into disjoint bands, which were presented both individually and in various combinations. In the first experiment, the stimuli were CV syllables, the bands were (A) 0-0.7, (B) 0.7-1.4, and (C) 1.4-2.8 kHz, and the combined bands were A+B and A+C. Bands were combined both diotically and dichotically to evaluate peripheral interference effects. In the second experiment, the stimuli were CVC syllables, the bands were (A) 0.1-1.4, (B) 1.4-2.1, (C) 2.1-2.8, and (D) 2.8-5.5 kHz, and the combined bands (diotic presentation only) were all 15 combinations of the individual bands.

In each experiment a group of five young adult listeners with normal hearing was tested. We derived predictions for the identification score in the combined-band conditions based on performance in the single-band conditions using the following approaches: (1) the ANSI 20-band Al calculation procedure;\(^\text{12}\) (2) the ANSI one-third octave Al calculation procedure; (3) Fletcher’s rule in which the error probability for combined bands is the product of the error probabilities for the individual bands; (4) Fletcher’s rule applied to individual consonants rather than the overall score; (5) the fuzzy logical model of perception;\(^\text{13}\) (6) the pre-labeling model; and (7) the post-labeling model.

Results from the CV experiment indicate that the score in the combined band condition was generally only slightly greater than the higher score obtained for the two constituent bands. This result is inconsistent with predictions based on the articulation index. The predictions of the post-labeling model are closest to observed scores, and are statistically indistinguishable from the data in roughly half the cases. The other models generally overestimate performance in the combined-band conditions. Scores for one listener were lower in the two diotic combined-band conditions than for the better constituent band, but were greater in the dichotic combined-band conditions, suggesting that peripheral interference may have occurred in the diotic condition.

Overall scores in the CVC experiment are fairly well predicted by the Articulation Index calculation, and all of the rules considered except the post-labeling model, which tended to underestimate scores in the combined-band conditions. The other models predicted overall scores satisfactorily in five to seven of ten possible cases: (five subjects, initial and final consonants). When identification scores for individual consonants are considered, the Fletcher Rule applied to individual consonants and the pre-labeling model gave the closest predictions.

Although the results of these two experiments suggest that different integration mechanisms may apply to low- and high-frequency bands of speech, they are consistent with respect to the (lack of) need for explicit cross-spectral comparisons in consonant identification. Identification scores for combinations of bands generally can be adequately accounted for without requiring such comparison mechanisms.

We have recently developed an automatic speech recognition system for use in studies of across-


\(^{13}\) D.W. Massaro, Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry, (Hinsdale, N.J.: Lawrence Erlbaum Assoc., 1987.)
The speech sounds of English (including both consonants and vowels) can be recognized with high accuracy in running speech without relying heavily on cross-spectral comparisons.

We have recently completed our preliminary investigation of a new computational approach for determining the intelligibility of speech subjected to waveform degradations or signal-processing transformations. The approach incorporates a model of the human auditory system that derives a sequence of discrete symbols from the speech waveform. Each symbol represents a prototypical vector of parameter values characteristic of the short-term spectrum measured in a single frame (10 ms) of speech. The perceptual effect of the degradation is estimated by assessing the consistency between the symbol sequence derived from an untransformed (input) speech signal and that derived from a transformed (output) signal. This is implemented via calculation of percent transmitted information (PTI), hypothesized to be monotonically related to intelligibility for a given auditory model, listener, and speech test. Two computational models of auditory system function were studied: one (CBL) based on critical-band levels and the other (EIH) based on the ensemble interval histogram representation of auditory nerve firing patterns. Degradations studied thus far include linear filtering and additive noise, whose effects on speech intelligibility are relatively well understood. For both the CBL and EIH models, PTI decreases monotonically as the severity of filtering or level of additive noise is increased. Since intelligibility is assumed to be monotonically related to PTI, the crossover frequency (equal PTI for highpass and lowpass filtering) and the value of S/N which yields the PTI at the crossover frequency, characterize the performance of the auditory models. For the CBL model, a PTI of 55 percent is achieved at a crossover fre-

quency of 1800 Hz and also for S/N = +23 dB. For the EIH model, a PTI of 40 percent is achieved at a crossover frequency of 1000 Hz and also for a S/N = +10 dB. Thus, neither the CBL nor the EIH model provides completely satisfactory accounts of the dependence of intelligibility on the degradations considered. The CBL model predictions were much more sensitive to the effects of noise, and EIH predictions to those of highpass filtering.

1.2.3 Clear Speech

Sentences spoken clearly are more intelligible than those spoken conversationally for (1) hearing impaired listeners in a quiet background,20 (2) normal-hearing listeners in a noisy background,21 and (3) both normal-hearing and hearing-impaired listeners in reverberant and noisy-reverberant backgrounds.22 In a recent study of normal-hearing listeners we tested the interaction between three distortions multiplicative noise;23 500 Hz lowpass filtering; and reverberation with speaking rate on intelligibility. We used two sets of materials Clarke sentences24 and Harvard sentences25 which were recorded by two teachers of the deaf as part of our study of manual cued speech.26 Each type of material was spoken both while the teachers produced manual cues (speaking rate of 100 wpm) and while not producing clues (150 wpm). This difference in speaking rates is only about half that between clear and conversational speech, and little exaggeration of articulation is evident in the speech produced while cueing. Roughly equal intelligibility test scores were obtained for the two speaking rates for all conditions with the exception that the IEEE sentences were less affected by reverberation at the slower speaking rate than at the faster one. These results suggest that there may be an interaction between speaking style and type of distortion on the intelligibility advantage associated with speaking clearly. Such an interaction should aid greatly in identifying the factors responsible for the difference in intelligibility between various speaking styles.

1.2.4 Control of Acoustic Feedback

Acoustic feedback imposes a limit on hearing aid gain that can result in low-level signals not being amplified to reach audibility. Although several techniques for reducing feedback have been proposed recently, evaluations of these techniques have been limited to measurements of the increase in stable gain that can be achieved; signal distortions introduced by the processing have not been considered, and measurement conditions have differed from study to study. We have implemented the most promising signal processing techniques and compared them not only for stable gain but also for their effects on sound quality and annoyance. In the process, we have developed a novel technique for reducing the effects of feedback.

The five feedback-reduction algorithms examined were an adaptive notch filter and four variants of adaptive cancellation. The adaptive notch method has been proposed previously for use in public-address systems but has not been evaluated in the hearing aid application, though it has been used in an algorithm for detecting oscillation.28 The adaptive cancellation methods have in common the goal of estimating the feedback path in order to cancel it

---


with another out-of-phase feedback path. In the systems described by Bustamante et al. and Engebretsen et al., the adaptation takes place continuously. Kates' approach differs in that the system monitors the signal for oscillation and, if detected, interrupts the main signal path and inserts a probe noise for feedback-path estimation. The fifth system also interrupts the signal, but it does so not only when oscillation is detected, but also during selected time intervals when the input signal is estimated to be low. This approach of quiet-interval adaptation has four clear advantages: (1) it provides better cancellation because the misadjustment caused by the input signal is minimized; (2) it provides a quasi-continuous, but stable, tracking of changes in the feedback path; (3) it achieves cancellation with minimal disturbance to the signal; and (4) minimal levels of probe noise are required.

All five methods were implemented in real time with a Motorola DSP96002 DSP and evaluated for increased stable gain (measured physically) and speech quality (measured by subjective ratings). Acoustic feedback paths (two static and one time-varying) were simulated electrically. Results showed the following: First, a single adaptive notch can be effective, but only if the feedback is confined to a single frequency region that is narrower than the notch width. Second, the continuously-adapting systems of Bustamante et al. and Engebretson et al. and, as far as we can determine, the system described by Dyrlund and Bisgaard, which is now available in commercial hearing aids, are inherently unstable. This instability manifests as bursts and warbling distortions that can be minimized only at the expense of degree of cancellation. Third, the noncontinuously adapting systems and the quiet-interval system perform equivalently only under static, quiet conditions. When the feedback path changes and/or when adaptation occurs in the presence of speech, the advantages of the quiet-interval method are apparent. Overall, the quiet-interval system, by providing maximal feedback cancellation with minimal disturbance to the user, represents a substantial advance over the other techniques.

1.2.5 Aids to Speechreading

During the past grant year, our research on aids to speechreading focused on the development of low bandwidth acoustic signals that enhance speechreading for listeners with severe hearing impairments. One class of such signals consists of tones that are amplitude modulated by the envelopes of filtered bands of speech. These signals have been shown to be effective supplements to speechreading by normal-hearing listeners. However, when the center frequency of the carrier bands is lowered to match the residual hearing of listeners with severe high-frequency loss, the signals no longer successfully supplement speechreading. A

---


second class of low-bandwidth supplements consists of tones that are modulated in frequency by estimates of the time-varying fundamental frequency of the speech signal and modulated in amplitude by the broadband speech envelope.

Takeuchi and Braida\textsuperscript{38} used a match-to-sample task to study the effect of a spectrally remote amplitude-modulated signal on the ability of listeners to compare two patterns of amplitude modulation in the presence of a distractor. Discrimination performance was measured as a function of the correlation between the envelopes of the comparison and distractor signals and difference between the carrier frequencies used to convey these envelopes. Although there are some differences across listeners, the addition of a distractor, even if it was spectrally remote from the comparison signal, degraded discrimination. The degree of interference tended to increase as the degree of correlation between comparison and distractor envelopes decreased. Performance on conditions with small carrier-frequency shifts between the comparison and distractor signals was generally worse than on conditions with wide spacing. When the comparison was higher in frequency than the distractor, there was a larger effect of frequency spacing than when the comparison was lower than the distractor. These results suggest that there may be some impairment to the perception of the envelopes of individual speech bands when an additional band is presented simultaneously. Since this impairment is somewhat greater when the bands are closely spaced in frequency, hearing-impaired listeners with reduced frequency selectivity may be more affected by the simultaneous presentation of multiple bands than listeners with normal hearing. It is noteworthy, however, that even on the conditions that produced the most impairment, performance generally did not fall below 70 percent of the level with no added signal.

An alternative acoustic signal that has been proposed as a speechreading supplement is based on conveying the voice fundamental frequency (F0) and the amplitude envelope of the broadband speech waveform. For research purposes, this signal can be based on accelerometer measurements of glottal vibrations and pressure-microphone measurements of the acoustic speech signal. To compare the effectiveness of such supplements with those based on the envelopes of bands of speech, we have developed a system capable of synthesizing a sinusoidal waveform whose amplitude is controlled by the envelope of the speech waveform and whose frequency is controlled by the variation in glottal-pulse periods.\textsuperscript{39} Since the system is implemented in software on a real-time signal processor, both the mapping from speech envelope amplitude to sinusoidal amplitude and the mapping from glottal pulse amplitude to tone frequency can be specified flexibly to accommodate the hearing loss and reduced dynamic range of impaired listeners.\textsuperscript{40} Using a system-produced supplement that incorporates only frequency modulation with a linear mapping of F0 to tone frequency, five relatively untrained listeners with normal hearing showed improved scores for word recognition in sentence context (Clarke Sentences)\textsuperscript{41} from 20 to 39 percent, roughly comparable to the benefit expected from a fixed-frequency tone amplitude modulated by the envelope of an octave band of speech.

1.2.6 Publications


\textsuperscript{41} M.E. Magner, A Speech Intelligibility Test for Deaf Children (Northampton, Mass.: Clarke School for the Deaf, 1972).
Chapter 1. Sensory Communication


1.3 Cochlear Prostheses

Sponsor
National Institutes of Health
Contract 2 POI DC003611

Project Staff
Professor Louis D. Braida, Lorraine A. Delhome, Dr. Donald K. Eddington, Dr. William M. Rabinowitz

The development of cochlear prostheses is directed at aiding individuals with profound sensorineural deafness who are unable to derive useful benefits from acoustic input to the ear. The prosthesis bypasses the impaired acoustic transduction mechanism and directly stimulates surviving auditory nerve fibers with currents delivered via an electrode array implanted within the cochlea. The overall goals of our research are to understand the mechanisms responsible for the improved hearing provided by these prostheses and to exploit this understanding for the development of improved systems. Our program makes use of a dedicated group of postlingually deafened adults implanted with a multichannel cochlear prosthesis (Ineraid, Richards Medical Corporation), who participate in intensive multifactorial studies. Our research capitalizes on having direct accessibility to the implanted electrode array through a percutaneous connector.

During the past year, work has involved (1) analysis of cue integration in audiovisual speech reception, (2) acoustic simulations of cochlear-implant speech reception, and (3) alternative speech processing for improved implant performance. The work in (3) is performed with Joseph Tierney and Marc Zissman; progress is described in part 5, section 3, chapter 1, section 1.7.3.

The work on audiovisual integration assesses the ability of an implantee to combine cues that are available from separately using vision and audition. Because most implantees require audiovisual input for reliable communication, analysis of integration is of particular significance.

A series of experiments on audiovisual integration were continued. Closed-set identification for a set of ten vowels, a set of the 12 most frequently occurring consonants, and the full set of 24 consonants is being tested. The vowels are tested in a medial position, within nonsense /b-V/-d/ syllables; the consonants are tested in initial position, within nonsense C1/-a/-C2 syllables. Stimuli are from a single female talker, with multiple tokens for each stimulus, and are presented through a computer-controlled videodisc. Results are being obtained using vision alone, audio (the implant) alone, and audiovisually. Most subjects have completed the ten-vowel and 12-consonant tests. The 24-consonant tests require considerably more trials for reliable estimates of confusion patterns for each of the three test modes; these tests are in varying states of completion for the subjects. Analyses to determine the quantitative efficiency of audiovisual integration are underway. These analyses exploit an analytical framework that quantifies how performance in a combined mode (e.g., audiovisual stimulation) is influenced by integration or interference of cues available in the isolated component modes (e.g., audio and visual stimulation alone).

---

42 Subcontract from Massachusetts Eye and Ear Infirmary. Dr. Joseph B. Nadol, M.D., Principal Investigator.

The work in acoustic simulations is designed to test our understanding of the cues that implantees receive when using the prosthesis alone. Speech is processed to remove specific details and the resulting waveform serves as a new sound input for speech testing. Insofar as an implantees speech reception is unchanged, the details that were removed must have been either unimportant for speech reception or were not being received by the subject.

Several lines of evidence suggest that Inerald users perceive information about the fine structure of the speech waveform only up to a limited frequency, roughly 300 to 1000 Hz depending upon the subject. Information about higher frequency waveform components is mediated principally by the envelopes of the bandpass signals that are delivered to the electrodes. To test this hypothesis directly, experiments have begun in which the signals normally delivered to the three upper channels of the implant are replaced by their envelopes modulating tones at the respective center frequencies of the bands. The low frequency portion of the spectrum (0 to 500 Hz) was unmodified. The modified system has been tested with a single subject who exhibits exceptionally good performance with his normal system (87 percent-correct on 12-consonant identification and 49 percent-correct on whole-word identification). His performance with the modified system was 72 and 26 percent-correct on the consonant and word tests, respectively. These scores are below his normal performance, but they are remarkably good, exceeding that of the average implantee using the normal system (about 55 percent on consonants and 15 percent on words).\footnote{W.M. Rabinowitz, D.K. Eddington, L.A. Delborne, and P.A. Cuneo, "Relations Among Different Measures of Speech Reception in Subjects Using a Cochlear Implant," J. Acoust. Soc. Am. 92: 1869-1881 (1992).} Performance that was much worse (14 percent on consonants) was obtained with this same subject when the three upper-band signals were removed, and only the lowpass speech (0 to 500 Hz) was presented. These results indicate that the three upper-band envelope signals can provide spectral-shape information that is very useful and that continued study with this processing model would be fruitful.

1.4 Tactile Communication of Speech

Sponsor
National Institutes of Health
Grant 5 R01 DC00126

Project Staff
Dr. Charlotte M. Reed, Gerald L. Beauregard, Lorraine A. Delhome, Gail Denesvich, Nathaniel I. Durlach, Ashanthi Gajaweera, Irene Kotok, Charlie Pan, Geoffrey L. Plant, Dr. William M. Rabinowitz, Dr. Mandayam A. Srinivasan, Hong Z. Tan, Andrew Ugarov

The ultimate goal of this research is to develop tactual aids for the deaf and deaf-blind that can serve as substitutes for hearing in speech communication. The objectives and specific aims of our research are as follows:

1. Basic study of encoding and display schemes to develop methods of displaying acoustic signals to the tactual sense for optimal information transfer.

2. Research on tactual systems designed specifically to aid speechreading, including systems based on both acoustic and articulatory-based signal processing.

3. Evaluations of experienced deaf users of portable, wearable tactual aids to determine improvements to speech reception provided by these aids and to compare this performance to that of users of other types of auditory prostheses.

1.4.1 Basic Study of Encoding and Display Schemes

During the current year, work has been conducted on five projects concerned with increased understanding of the proprioceptive/kinesthetic as well as the cutaneous sensory system of the hand.
Amplitude and Frequency Resolution for Motional Stimulation

This study is concerned with characterizing psycho-physical performance for low-frequency (1-32 Hz), high-amplitude (up to 2 cm) signals delivered to the fingertip using a position-controlled servomotor. Measurements of amplitude discrimination have been obtained on three subjects as a function of frequency for several values of reference displacement and stimulus duration. The amplitude jnd (which is roughly independent of duration in the range of 250-1000 msec) appears to decrease slightly (from roughly 1.1 to 0.6 dB) as frequency increases from 2 to 32 Hz. The value obtained at 32 Hz is similar to results reported in the literature for vibrational frequencies. Measurements of frequency discrimination are complicated at very low frequencies by the fact that durational differences, as well as frequency differences, can be used as a cue for discriminating F and F+ ΔF. To limit performance solely to frequency cues, the overall duration of the reference and comparison stimuli are being randomized within a run of trials. Preliminary results for frequency discrimination indicate that, for d'=1, Δf/f is roughly 0.1 across the frequency range 4-16 Hz. The difference limen reported in the literature for frequencies in the range 30-200 Hz increases with frequency from roughly 0.3 to 0.5. Thus, frequency discrimination at very low frequencies appears to be somewhat better than at the higher frequencies.

Tactual Reception of Morse Code Sequences

This study, which employs the same motion-stimulation device as described above, is concerned with the tactual reception of Morse Code in subjects who are either highly skilled or naive in the sending/receiving of Morse Code. Computer-generated sequences of Morse Code are delivered to the fingertip using high-amplitude square-wave stimulation. Reception ability is being studied as a function of the rate of presentation (12, 16, and 24 words/minute) and the length of the stimulus stream (isolated letters, random three-letter strings, common words, and conversational sentences). Results for two subjects who are already highly experienced in the sending and in the auditory reception of Morse Code indicate faster learning curves and higher levels of tactual performance than those obtained for naive subjects. For example, the experienced subjects achieved criterion performance of 95 percent on the single-letter recognition task at all three presentation rates within a total of 10,000 trials. For the naive subjects, learning is proceeding at a slower pace, and differences between the skilled and naive subjects appear to increase with rate. While the skilled subjects were able to receive CUNY sentences nearly perfectly at a rate of 16 words/minute, the unskilled subjects have not advanced beyond the single-letter task. The auditory reception of sentences in Morse Code by the skilled subjects, however, is superior to their tactual reception since they can obtain near-perfect scores for auditory reception of the code at rates of 40-50 words/minute.

Discrimination of Thickness

Study of the discrimination of thickness has continued over the current year with the development of a device and software for experimentation and with collecting preliminary data on several subjects. The ability to discriminate thickness is being studied for three materials with different bending moduli (steel, brass, and plastic, in decreasing order of stiffness). Various steps have been taken to eliminate potential cues other than thickness in performing the discrimination task, including the use of different tokens (i.e., individual plates) for each material and thickness, limiting the area of the plates in which the thumb and forefinger may make contact, and placing the apparatus in a temperature-controlled box to minimize the use of thermal cues. On a given trial, two plates of different thickness (but of the same material) are presented sequentially, and the subject's task is to judge the order of presentation (thick-thin or thin-thick). Training and testing has been conducted thus far with 8-mil reference plates and with comparison values in the range of 8-20 mils. Preliminary results indicate a possible correlation between stiffness and jnd for thickness in that the jnd for steel (the stiffest of the three materials) is higher than for brass and plastic (presumably due to additional cues supplied by compression of the fingerpads as the material bends).

Design of Multifinger Tactual Stimulator

Work has been conducted on the design of a new multifinger device for use in studies of the information capabilities of the tactual (cutaneous plus kinesthetic) sense. The specifications for this device include simultaneous stimulation of three to five fingers with one degree of freedom in movement for each finger. The frequency range of stimulation is specified as 0-300 Hz, with maximum motion of one inch peak-to-peak at very low frequencies decreasing to roughly 100 microns for frequencies above 100 Hz. The geometric constraints on the display are dictated by the shape of the hand.
Study of Manual Compliance

New experiments have been completed on compliance discrimination in which work cues were eliminated and force cues were minimized. When compared to data obtained previously in which both work and terminal force cues, as well as compliance cues, were available to the subject, the new data indicate (a) that manual resolution of compliance deteriorates when force and/or work cues are reduced or eliminated and (b) that subjects appear to rely strongly on terminal force cues (when they are available) in making compliance comparisons.

1.4.2 Tactual Supplements to Speechreading

Work during the current year has been concerned with developing a more complete understanding of the differential results obtained for auditory and tactile presentation of acoustic-based low-bandwidth supplements to speechreading. Our previous work has been extended in two areas described below.

Comparison of Auditory and Tactile Presentation of a Single-band Envelope Cue as a Supplement to Speechreading

Further experiments have been conducted using a low bandwidth supplemental speechreading cue. The envelope of an octave band of speech centered at 500 Hz is used to modulate the amplitude of a 200-Hz tone. The usefulness of this cue for speechreading, when presented either auditorily through headphones or tactiley through a single-channel vibrator, was evaluated for a set of 24 initial consonants in CVC syllables and a set of 16 vowels and diphthongs in CVC syllables. Consonant results, averaged across three subjects, showed improvement over lipreading alone (53 percent-correct performance) of 10 percentage points for the tactile supplement and 22 percentage points for the auditory supplement. The improvement observed with the supplements appears to be due primarily to improved reception of the voicing feature for the tactile condition and of the voicing and nasality features for the auditory condition. Vowel results indicated near-ceiling performance on all three conditions. The segmental results will be related to those obtained previously for sentence materials.

Tactile and Auditory Measures of Amplitude-modulation Discrimination

One possible explanation for the better results obtained for auditory compared to tactile presentation of the same supplemental cue to speechreading may be differences in basic psychophysical properties of hearing and touch. In particular, the ability to discriminate differences in amplitude modulation may be highly relevant to the supplemental cue we are studying. Experiments are being conducted to assess AM discrimination as a function of modulation depth (m=0 and 1) and modulation frequency (f_m=5 or 50 Hz) for a 200-Hz carrier. Results indicate that, on average, auditory thresholds are roughly 6 dB more sensitive than tactile thresholds for a given set of stimulus parameters, with the exception of the condition m=0 and f_m=50, where auditory and tactile performance are nearly the same. Our results extend those previously available in the literature by including data for nonzero values of reference modulation, which may be particularly relevant to the information that is conveyed in the single-band speech envelope.

1.4.3 Evaluation of Practical Aids

In addition to continuing periodic laboratory evaluations of eight subjects who have been fitted with tactile devices and are wearing them in the field, two other studies have been initiated during the current year.

---


The Effects of Training on Adult Performance Using the Tactaid VII

Two additional adults have been fitted with the Tactaid VII and are participating in weekly training sessions in the use of the device as a supplement to speechreading. Pre-training measurements were obtained on the subjects' speechreading ability in two tasks (sentence reception and tracking) both with and without the tactile aid. A detailed training program using both analytic and synthetic training materials has been prepared for use in the program, emphasizing the use of the tactile aid as a supplement to lipreading but also including exercises for the tactile reception of speech alone. After a six-month training period, the speechreading tests will be administered again to determine the effects of training. Performance will then be compared with the group of adults fitted with the Tactaid VII but provided with only basic aid orientation and limited training.

The Effects of Training on the Use of Tactile Aids by Congenitally Deaf Children

This study involves long-term training of a group of congenitally deaf children (age range five to ten years) attending the Boston School for the Deaf. These children, who derive limited benefit from amplification (hearing aids and FM trainers), are provided with weekly therapy sessions using the Tactaid VII. The training sessions emphasize the use of the tactile aid for speech development, auditory/tactile speech reception, and the use of the tactile aid as a lipreading supplement. Each child's speech production and speech reception skills were assessed at the beginning of the program and will be retested at three month intervals. An additional child, who has normal hearing up to 1.5 kHz but no measurable hearing above 2 kHz, is receiving training on a modified tactile aid which provides information on high-frequency consonants. Speech-production tests consisting primarily of materials that include sounds with high-frequency energy were administered prior to the start of a training program with the aid, and will be re-administered at three-month intervals.

1.4.4 Significance

The research conducted here leads to increased understanding of tactual speech communication and contributes to the development of improved tactual aids for individuals who are deaf and deaf-blind. This research plays a significant role in helping such individuals to achieve improved speech reception, speech production, and language competence. At the more basic scientific level, this research contributes to increased knowledge about speech communication, tactual perception, manual sensing, display design, and sensory substitution.

1.4.5 Publications


Thesis


1.5 Multimicrophone Hearing Aids

Sponsor

National Institutes of Health
Grant R01-DC00270

Project Staff

Joseph G. Desloge, Nathaniel I. Durlach, Julie E. Greenberg, Michael P. O'Connell, Dr. William M. Rabinowitz, Daniel P. Walker, Dr. Patrick M. Zurek

The long-term goal of this research is the development of sensory aids that improve the ability of hearing-impaired listeners to function in complex acoustic environments through the use of microphone arrays. Since the reception of speech is the most important problem for the hearing impaired, the target signal of primary interest is speech.
We envision a microphone array system that resolves incoming signals into simultaneous directional channels followed by a coding operation that transforms these resolved signals so that resolution is preserved at the perceptual level after the signals are summed for presentation either to one or two ears. Such a system would permit even a monaural listener to monitor all directions simultaneously, detect and localize in the same operation, and focus on a single direction.

Our current work on microphone arrays is directed toward the extraction of a single target signal (assumed to arise from a target source straight ahead of the listener) and the reduction of interference from sources directionally distinct from the target source. Approaches based both on fixed and adaptive arrays, with processing for monaural or binaural outputs, are being studied.

Microphone arrays with fixed (time-invariant) weights are directed at enhancing a desired signal from one direction (straight ahead) while attenuating spatially distributed interference and reverberation. Past work applied the theory of sensitivity-constrained optimal beamforming to free-field head-sized arrays. This analysis showed that broadband intelligibility-weighted directivities of 7-8 dB can be obtained either with a broadside array of cardioid microphones or with an endfire array with weights based on analog gains and pure time delays. Current work is directed at implementing portable prototype multimicrophone arrays to be worn and evaluated by hearing-aid users.

Work on adaptive array hearing aids has recently been focused on algorithmic analysis and design. Issues of adaptation control to minimize target cancellation and misadjustment, and robustness in reverberation, issues that were outlined previously, are being thoroughly analyzed. The result of this analysis will be a specification for a relatively simple and robust five-microphone broadside adaptive array.

1.6 Superauditory Localization for Improved Human-Machine Interfaces

Sponsor
U.S. Air Force - Office of Scientific Research
Contract AFOSR-90-0200

Project Staff
Nathaniel I. Durlach, Professor Richard M. Held, Dr. William M. Rabinowitz, Barbara G. Shinn-Cunningham, Eric M. Foxlin

The motivation and background for this project were described in pages 312-313 of the RLE Progress Report Number 134 and pages 306-307 of the RLE Progress Report Number 135.

In the past year, work on facilities development included the development of specifications for an improved auditory spatial synthesis device and further development of a new inertial position tracker with enhanced resolution and speed.

In examining ways in which supernormal localization cues could be produced, the idea of synthesizing HRTFs from larger-than-normal heads by frequency-scaling normal HRTFs was considered. The theoretical effects of frequency-scaling HRTFs to approximate a larger than normal head were investigated and reported in Rabinowitz, Maxwell, Shao, and Wei (1993). In this work, it was shown that frequency-scaling normal HRTFs will produce results very similar to HRTFs from larger than normal heads, provided the sources to be simulated are relatively far from the listener.

Earlier experiments showed that some adaptation could occur when active sensorimotor tasks were employed to train subjects. These basic results were investigated in more detail in the past year by varying the (1) type of training given to subjects, (2) number of positions employed, (3) complexity of the acoustic environment, and (3) presence of visual cues.

In all experiments performed to date, introduction of the transformation produced the anticipated changes in resolution and increased resolution in

the center of the field. Furthermore, most of the experiments showed clear evidence of adaptation. Not only did the subjects in these experiments show a reduction in bias (and localization error) with exposure to the altered cues, but they also showed an increase in bias (and localization error) in the opposite direction when tested with normal cues following altered-cue exposure (a negative after-effect). Of particular interest in the context of classical adaptation work, adaptation was found to occur, with essentially comparable strength, without involvement of the sensorimotor system in the adaptation process: results were nearly identical when feedback was purely cognitive. This result contrasts with previous work (e.g., by Held, Mikaelian et al., Freedman et al., and Kalil) which focused strongly on the importance of sensorimotor involvement.

Independent of the issues of sensorimotor involvement, two results complicate the simplified picture described above. First, no adaptation occurred in the experiment in which the subjects were blindfolded. Second, in over half the experiments there appeared to be a tendency for enhanced resolution to decrease with increased exposure to the altered cues.

In the future, we will focus on analyzing the apparent decrease in resolution with exposure, examining the number of necessary and sufficient visual cues to achieve adaptation, investigating how the strength of the remapping function affects the speed and extent of adaptation, and developing a statistical model that describes experimental results.

Additional work connected with this grant has involved participation in meetings at government agencies (e.g., NASA and ONR) and the Acoustical Society of America.

An invention disclosure has been submitted to MIT’s Office of Technology Licensing for the work on an inertial tracking system. A patent may ensue for this tracker, which was supported both by this project and NASA contract NCC 2-771.

1.6.1 Publications


Talks


1.7 Mechanistic Modeling of the Primate Fingerpad

Sponsor

National Institutes of Health
Grant R29-DC00625

Project Staff

Dr. Mandayam A. Srinivasan, Kiran B. Dandekar, Walter E. Babiec

When we touch an object, the source of all tactile information is the spatio-temporal distribution of mechanical loads on the skin at the contact interface. The relationship between these loads and the resulting stresses and strains at the nerve terminals within the skin plays a fundamental role in the neural coding of tactile information. Although empirical determination of the stress or strain state of a mechanoreceptor is not possible at present, mechanistic models of the skin and subcutaneous tissues enable us to generate testable hypotheses on skin deformations and associated peripheral neural responses. Verification of the hypotheses can then be accomplished by comparing the calculated results from the models with biomechanical data on the deformation of both the skin and subcutaneous tissues and with neurophysiological data from recordings of the responses of single neural fibers. The research under this grant is directed towards applying analytical and computational mechanics to analyze the biomechanical aspects of
touch—the mechanics of contact, the transmission of the mechanical signals through the skin, and their transduction into neural impulses by the mechanoreceptors.

1.7.1 Determination of Geometric and Material Properties of the Primate Fingertip

The first step in performing mechanistic analyses of the primate fingertip (distal phalanx) is to determine its geometric and material properties. Three-dimensional (3D) external geometry of the primate fingertip was determined from accurate epoxy replicas of human and monkey fingertips. Using a videomicroscopy setup, we obtained images of orthographic projections of the epoxy replicas at various known orientations. The images were then digitized and processed to determine the boundary of the finger at each orientation. By combining the boundary data for all the different orientations, we were able to reconstruct the 3D external geometry of the fingertip. We have reconstructed several human and monkey fingertips using this method.

For mechanical modeling of the human fingerpad, the Poisson’s ratio, which is a measure of its compressibility, is required as an input to the mathematical models. The Poisson’s ratio for the human fingerpad in vivo is unknown at present. In previous noninvasive experiments on human subjects, we have measured the change in volume of the fingerpad under static indentations with different indentors. Our results show that the compressibility of the fingertip increases with increases in both the depth of indentation and the contact area with the indenter. The highest change in fingertip volume was about 5 percent. We have now developed an experimental setup involving a computer controlled linear actuator for fingertip volume change measurements under dynamic conditions.

1.7.2 Fingertip Models and Finite Element Analyses

We have performed linear and nonlinear finite element analyses of a series of mechanistic models of the fingerpad under a variety of mechanical stimuli. The models range from a semi-infinite medium to a 3D model based on the actual finger geometry and composed of either a homogeneous elastic material or a thick elastic shell containing a fluid. Simulations of the mechanistic aspects of neurophysiological experiments involving mapping of receptive fields with single point loads, determination of spatial resolution of two-point stimuli, and indentations by single bars as well as periodic and aperiodic gratings have been carried out for the 2D models. We have also solved the nonlinear contact problem of indentations by cylindrical objects. The large number of numerical calculations needed even for the linear 2D models necessitated the use of the MIT supercomputer.

The results show that the (1) model geometry has a significant influence on the spatial distribution of the mechanical signals and (2) elastic medium acts like a low-pass filter in causing blurring of the mechanical signals imposed at the surface. None of the other models predicted the experimentally observed surface deflection profiles under line loads as closely as a simple heterogeneous model that treated the fingerpad as a membrane enclosing an incompressible fluid. The results from the analyses of these models were then related to recordings of the responses of slowly and rapidly adapting mechanoreceptors (SA and RA, respectively). Under indentations by rectangular gratings, the maximum compressive strain and strain energy density at the receptor locations emerged as the two strain measures that are directly related to the response rate of SAs. The same result was true for the contact problem of indentations by cylinders. However, strain energy density is a better candidate since it is a scalar that is invariant with respect to receptor orientation and a direct measure of the distortion of the receptor caused by loads imposed on the skin. For the model calculations to match both biomechanical and neurophysiological data more


accurately, it is likely that the models will have to take into account the inhomogeneity of the primate fingerpad.

1.8 Peripheral Neural Mechanisms of Haptic Touch

Sponsor
U.S. Navy - Office of Naval Research
Grant N00014-88-K-0604

Project Staff
Dr. Mandayam A. Srinivasan, Dr. Anuradha M. Annaswamy, Dr. Robert H. LaMotte,\(^\text{56}\) Steingrimur P. Karason

We have been collaborating with Dr. LaMotte of Yale University School of Medicine in conducting psychophysical and neurophysiological studies on the tactile perception of the microtexture, shape and softness of objects. We have shown that humans can detect extremely fine textures composed of 50 nanometers high parallel bars on plane glass plates.\(^\text{57}\) Our neurophysiological recordings indicate that when such fine textures are stroked on the fingerpad skin, the fingerprint ridges vibrate and cause Pacinian Corpuscles to respond, thus enabling the detection of the microtexture.\(^\text{58}\) In studies of the tactile perception of shape, a series of two- and three-dimensional objects (e.g., cylinders, spheres, ellipsoids and wavy surfaces) were pressed or stroked across the fingerpads of anesthetized monkeys and the evoked responses in cutaneous mechanoreceptive primary afferent nerve fibers were recorded.\(^\text{59}\) Major geometrical properties of the shapes were well represented in the spatio-temporal responses of SA and RA afferent fiber populations, particularly those of the SAs. The results show that the following hypothesis explains qualitatively all the data we have obtained: the depth of indentation and the change in curvature of the skin surface are encoded by the discharge rates of SAs. In addition, the velocity and the rate of change in skin surface curvature are encoded by the discharge rates of both SAs and RAs.

Thus, intensive parameters of shapes, such as the magnitude of change in skin curvature produced by contact with the object surface, were encoded in the discharge rates of SAs and RAs, but this neural code was also influenced by changes in stroke velocity. Spatial parameters of shapes such as the curvature width and the changes in contour that characterize a shape as belonging to a particular category (such as a sphere as opposed to a cylinder) are encoded in the spatially distributed discharge rates of the SA population. This spatial response profile provides a neural code that is probably invariant with moderate changes in the parameters that govern contact conditions between the object and the skin, such as the contact force or orientation and velocity of its trajectory. Therefore, among the different possible geometric representations of the shape of objects, the intrinsic description, i.e., the surface curvature as a function of the distance along the surface, seems to be relevant for tactile sensing of shape.

Based on a theoretical analysis of the mechanics of contact, we have proposed a mechanism by which shapes of objects within contact regions are perceived through the tactile sense. The curvature of the skin surface under an object, which we know from differential geometry is approximated by the second spatial derivative of surface deflection, is coded not by differentiating (which is a noise enhancing process), but by exploiting its relation to surface pressure. Pressure peaks occur where the depths of indentation and/or changes in the skin surface curvature are high. The skin effectively acts as a low-pass filter in transmitting the mechanical signals, and the mechanoreceptors respond to the blurred versions of the surface pressure dis-

\(^{56}\) Yale University School of Medicine, New Haven, Connecticut.


We have also shown that the human discriminability of softness or compliance of objects depends on whether the object has a deformable or rigid surface. When the surface is deformable, the spatial pressure distribution within the contact region is dependent on object compliance, and hence information from cutaneous mechanoreceptors is sufficient for discrimination of subtle differences in compliance. When the surface is rigid, kinesthetic information is necessary for discrimination, and the discriminability is much poorer than that for objects with deformable surfaces.

1.8.1 Development of a Computational Theory of Haptics

Our research on computational theory of haptics is focused on developing a theoretical framework for studying the information processing and control strategies common to both humans and robots performing haptic tasks. For example, although the "hardware" of the tactile apparatus in humans and robots are different, they have the common feature of mechanosensors embedded in a deformable medium. Thus the computational problem of coding (predicting sensor response for a given mechanical stimulus at the surface) and decoding (inferring the mechanical stimulus at the surface by suitably processing the sensor response) need similar mechanistic analyses for their solution. We have developed such a "computational theory" for an idealized medium subjected to arbitrary pressure or displacement loading conditions, and give explicit formulae for the coding and decoding problems.

In our work during the past year, we were successful in answering basic identification and control questions that arise during manipulation of compliant objects using compliant fingerpads. To understand the fundamental aspects of these tasks, we have analyzed the problem of identification of compliant objects with a single finger contact, as well as under a two-finger grasp. Using lumped parameter models, we have carried out the identification of human and object parameters, using either force or displacement inputs to the rigid backing of the end-effector. Based on identified parameters, control strategies are developed to achieve a desired manipulation of the object in the workspace.

We have also modeled the dynamic interactions that occur between compliant end-effectors and deformable objects by a class of nonlinear systems. It was shown that standard geometric techniques for exact feedback linearization were inadequate. New algorithms were developed by using adaptive feedback techniques which judiciously employed the stability characteristics of the underlying nonlinear dynamics. Both theoretical and simulation results show that these adaptive control algorithms led to successful manipulation. The theoretical results can be used to generate testable hypotheses for experiments on human or robot haptics.

1.9 Biomechanics of Skin-Object Contact

Sponsor
U.S. Navy - Office of Naval Research
Grant N00014-91-J-1454

Project Staff
Dr. Mandayam A. Srinivasan, Jyh-Shing Chen, Rogeve J. Gulati

Although physical contact is ubiquitous in our interactions with objects in the environment, we do not yet understand the mechanistic phenomena occurring at the skin-object interface. As mentioned before, the spatio-temporal distribution of mechanical loads on the skin at the contact interface is the source of all tactile information. These loads, specified as pressure, displacements, etc., depend on the geometrical and material properties of both the contacting entities, as well as the overall forces of interaction. The goal of this project is to determine the growth and motion of contact regions and vari-

---


Chapter 1. Sensory Communication

ations in the associated force over time between
the human fingerpad and carefully chosen trans-
parent test objects whose microtexture, shape or
softness is varied in a controlled manner. These
results are being used to gain a deeper under-
standing of the data we have already obtained for
the same test objects, namely, electrophysiologically recorded responses of cuta-
neous mechanoreceptive afferent fibers, and
psychophysically determined human discrimin-
ations.

To measure the in vivo surface deformations of the
fingerpad under various tactile stimuli, we have
designed a videomicroscopy system together with
high precision force sensors. The videomicroscopy
system consists of a set of video zoom lenses
attached to a high-resolution CCD camera, whose
output can either be digitized directly at about 5
frames/s, or stored on a laserdisk at real-time frame
rates (30 frames/s) for off-line digitization. The
zoom lenses enable continuous variation of magni-
fication, with the field of view covering the entire
fingerpad, or just a few fingerprint ridges. High
contrast images are achieved with coaxial and other
fiberoptic lighting. In collaboration with our col-
leagues at the MIT Artificial Intelligence Laboratory,
we designed and built two 6-axis force sensors
that are customized to our application. These sensors
have much higher resolutions (10 bit) than com-
mercial sensors operating in a comparable range of
forces (5 Newtons). Transparent test objects can
be attached to these sensors for both
biomechanical and psychophysical experiments.

Using the test facility described above, we have
performed a set of experiments with human sub-
jects to investigate the relationship between the
contact force, contact area and compliance of the
object. The experiments involved active indentation
of transparent compliant rubber specimens and a
glass plate with the subjects' fingerpads. Static
video images of the contact regions were captured
at various force levels and magnifications. In order
to minimize the effects of nonuniform illumination,
we implemented homomorphic image processing
algorithms with or without image decimation. The
processed images showed that contact regions con-
sisted of discontinuous "islands" along each finger
ridge, with clear distinction between contact and
noncontact regions over the entire field of view.

Results show that for objects whose compliances
are discriminable, even when the overall contact
areas under a given contact force are the same, the
actual contact areas can differ by a factor of two or
more. The actual pressure distribution, which acts
only within the discontinuous contact islands on the
skin, will therefore be radically different for the
objects. Consequently, a spatio-temporal neural
code for object compliance emerges with far higher
resolution than an intensive code such as the
average pressure over the overall contact area.
These results are in agreement with our hypothesis
that the neural coding of objects with deformable
surfaces (such as rubber) is based on the spatio-
temporal pressure distribution on the skin. This
was one of the conclusions from our psychophys-
ical, biomechanical and neurophysiological exper-
iments in a companion project conducted in
 collaboration with Dr. LaMotte of Yale University
School of Medicine.

In another set of experiments, the images of a
human fingertip before and after indentation with
rectangular and cylindrical bars were obtained. By
superposing the images, we can measure the
surface profile as well as the displacements of the
points on the skin surface marked with ink. Such
measurements for several bars at different orien-
tations serve as biomechanical data to verify and
guide the finite element models of the fingertips
described above.

1.10 Human and Robot Hands:
Mechanics, Sensorimotor Functions and
Cognition

Sponsor
U.S. Navy - Office of Naval Research
Grant N00014-92-J-1814

Project Staff
Dr. Mandayam A. Srinivasan, Dr. J. Kenneth
Salisbury, Nathaniel I. Durlach, Dr. Robert D.
Howe,64 Dr. Robert H. LaMotte, Kiran B. Dandekar,
Louise Jandura, Kevin C. Knoedler

The premise of this University Research Initiative
project is that the integrated study of human and
robot haptics can provide complementary know-
ledge of the processes of prehension and manipu-
lation. From the human research side, we wish to
understand the basic mechanical, perceptual, and
strategic capabilities that lead to the dexterity and
deftness we observe in human task performance.
By studying the underlying competences that

64 Harvard University, Cambridge, Massachusetts.
humans bring to bear on task performance, we seek guidelines on how to better build robots. From the robotic research side, we wish to understand how mechanism and sensor design choices can best be made to maximize grasping and manipulative competences in robot task performance. By understanding better the mechanical demands of task performance, we seek to understand the performance demands which underlie skillful human manipulation. To bridge these two areas of study, we have assembled a unique team of researchers with diverse backgrounds in robotic and human studies at MIT, Harvard and Yale Universities.

The main components of the research conducted under this project during the past year are (1) development of new hardware for robotic and human studies, (2) processing of robot sensor signals and task-level control of the devices, and (3) experiments on human perception and control using some of the devices. The subsections to follow provide description of the results obtained in each of the three topics.

1.10.1 Developments of New Hardware

Hardware developments by the Dextrous Manipulation Group at the Artificial Intelligence (AI) Laboratory at MIT include two versions of glass-smooth actuator test-beds, the first version of the Precision Manipulation Instrument, and PVDF-2 contact sensors. Collaboration between the AI Lab Group and the Human Haptics Group at MIT has resulted in the design and fabrication of miniature, high-resolution six-axes force sensors, and an Instrumented ScrewDriver for experiments on human force control. The Harvard Group and the MIT Human Haptics Group have collaborated in building a reconfigured robot finger to be used as a tactile stimulator in experiments on human subjects. The Harvard Group has fabricated a prototype of a flexible tactile array pressure distribution sensor and an integrated sensor system. They have also begun development of a multi-element stress rate sensor for improved slip and contact detection along with new signal processing electronics for improved rejection of noise originating in the manipulator structure. The Yale group has designed and fabricated an instrumented probe that is capable of measuring dynamic grasp forces. The probe can be mounted on a torque motor for recording and playing back the transient force signals of impact with compliant objects in psychophysical experiments under both active and passive touch.

1.10.2 Robot Signal Processing and Task-Level Control

Interpretation of Contact Force Signals from a Robot Fingertip

Through theoretical advances in processing of temporal signals from a robot fingertip force sensor, the MIT AI Lab group has successfully applied the generalized sequential likelihood ratio test to temporal contact force signals and used it to segment them into statistically equivalent pieces indicating significant contact events. This permits detection and labeling of impacts, slip, changes in texture, and contact condition. While of immediate value in detecting and differentiating contact events, the results of the segmentation are designed to be used in a higher level procedure which will interpret the results within a manipulation context. All the algorithms are based on forms of the the generalized sequential likelihood ratio test. This test is optimal under mild conditions. Our algorithms yield significantly better performance on standard tasks, like grasping contact, than current thresholding techniques. In addition, our new algorithms allow continuous self-calibrating (adaptive) sensing and the detection/segmentation of more complex phenomena.

Task Level Control System

A task level control system has been developed for the Salisbury Hand and applied to the problem of grasping. The system is designed to grasp objects in a semi-structured environment (static objects which may be displaced randomly). The hand is outfitted with force sensing fingertips and a simple contact sensing palm. The system uses sparse contact information (spatially and in time) obtained from these sensors and coarse geometric models of objects to recursively refine its "knowledge" about object position and grasp state. Each contact event provides partial information about the task state and is used to increase the knowledge about it. Aging or inconsistent data is rejected, thus permitting a degree of robustness to error and unexpected disturbances. Local motions aimed at improving the task knowledge (and ultimately, grasp quality) are planned "on the fly" based on the current knowledge state. The system thus incorporates the responsiveness of sensor-based reactive systems and the correctness of planning systems. Proofs of stability and convergence have been obtained and the system has been demonstrated in the grasping of cylinders and blocks.
Chapter 1. Sensory Communication

Regrasping Gaits

An investigation of regrasping "gaits" has been conducted. This work is aimed at both developing methods for robot hands to continuously reorient objects by repositioning fingers and at revealing the underlying performance of this element of skillful manipulation. A systematic approach for enumerating acceptable regrasping sequences has been developed, and this permits us to identify those patterns of finger movement that allow continual reorientation of objects acknowledging friction limits, object weight, and object orientation.

1.10.3 Experiments on Human Perception and Control

Human Performance in Controlling Normal Forces of Contact

The Human Haptics Group at MIT has investigated the ability of humans in actively controlling contact forces with normal and locally anesthetized fingertips. The effect of force magnitude, time-profiles of forces, and visual feedback has been studied. The experiments consisted of tracking constant, linear-ramp and sinusoidal time-profiles of forces applied on a custom-built six-axis force sensor. In most of the experiments, visual feedback was provided by displaying on a computer monitor the time-profiles of both the desired target forces and actual forces exerted by the subjects. Some experiments on maintaining constant force were conducted without visual feedback.

The results show that for a constant target force, the mean error in tracking remained approximately constant with respect to force magnitude when visual feedback was present, but increased with target force magnitude in the absence of vision. The lack of tactile feedback due to local anesthesia caused the error to increase by 25 to 80 percent. Higher force rates also generally increased the percentage of error. Among the three target profiles, the constant force had the least error. For a given force rate, the errors for the sinusoidal and ramp targets were about the same. A design specification that emerges from these results is that haptic interfaces for virtual environments and teleoperation must have a force resolution of about 0.01 Newton at low force rates to make full use of human haptic capabilities.

Human Performance in Controlling Torque

The human ability to control torque has been investigated in preliminary experiments with the Instrumented ScrewDriver (ISD). The ISD, comprised of a single shaft which is supported by low friction bearings, is connected to a reaction torque sensor and a magnetic particle brake. Angular position of the shaft is measured by an incremental optical encoder. In this set of experiments, the subjects were asked to grasp the handle between the thumb and index finger and turn the shaft clockwise at a constant velocity. At a certain point during this maneuver, the resistive torque acting on the shaft was increased in the form of a sharp step of variable magnitude. When the step change in resistive torque occurred, the initial motion ceased while the subject increased the applied torque. After this time delay, motion of the shaft resumed and continued until the experiment was over. The results show that the time delay increased approximately linearly with resistive torque, indicating that the subjects increased the applied torque at an almost constant rate until shaft motion resumed. Increasing standard deviations about the mean applied torque after the delay indicated that increases in the resistive torque level cause decreases in the human ability to control torque.

1.10.4 Processing of Tactile Information Obtained by Touching an Object with a Tool

The Yale group has recorded and analyzed tactile signals delivered to a single fingerpad in contact with a tool (a stylus) that strikes objects of differing compliance. One end of the stylus, attached to a servocontrolled torque motor, rested against the restrained fingerpad of the subject and exerted a base force of about 0.2 Newton. The other end of the stylus was struck with a rubber object whose compliance was varied. The force exerted by the stylus against the skin upon impact with the object was recorded. In a series of experiments, this signal was edited slightly and "played back" under servocontrol of force by the motor through the stylus. The subject judged the magnitude of subjective softness by assigning a number (the method of Magnitude Estimation). On certain trials, the back end of the stylus was struck with a rubber object by the experimenter while on other trials the

---

motor played back a signal that had been previously recorded. In addition to judging the magnitude of softness, the subjects made a two-alternative forced choice as to whether the tactile signal was "live" or "played back." Results demonstrated that the subjects were unable to make this discrimination above the level of chance performance. This meant that the tactile signals played back by the motor felt natural and represented those generated when a stylus is struck by a compliant object.

A series of experiments in which the recorded traces were filtered or otherwise edited demonstrated that the critical signal associated with the compliance of an object was the slope of the rising phase of the force trace. This feature was more important than others such as peak amplitude or higher frequency oscillations in the signal. In the future, these signals will then be delivered to the cutaneous receptive fields of mechanoreceptors to learn how tool-generated tactile information is transduced and how associated neuronal events in sensory pathways are processed by the brain during haptic performance.

1.10.5 Publications


Papers Accepted for Publication


Chapter 1. Sensory Communication

**Theses**


1.11 Virtual Environment Technology for Training (VETT)

**Sponsor**

U.S. Navy - Naval Training Systems Center  
Contract N61339-93-M-1213  
Contract N61339-93-C-0055  
Contract N61339-93-C-0083

The VETT program consists of five related research projects that include a virtual environment (VE) system that serves as a research testbed and four satellite projects. The testbed and two of the satellites, sensorimotor adaptation and improved haptics interfaces, are underway at RLE. Research on the whole body motion satellite is being conducted at the Graybiel Laboratory of Brandeis University, and the training satellite research is being conducted at Bolt, Beranek and Newman, Inc., of Cambridge. This report will cover the VETT testbed, the sensorimotor satellite, and the haptics interface satellite.

1.11.1 VETT Testbed

**Project Staff**

Dr. David Zeltzer, Dr. Jeng-Feng Lee, Walter A. Aviles, Rakesh Gupta, Erik Nygren, Nicholas Ploch

The first two phases of the first year of the VETT Testbed effort have been completed as scheduled. All major computational and human/machine interface systems have been procured, installed, and tested. Computational equipment includes an SGI 2-processor ONYX system with VTX graphics subsystem, two SGI Indigo2 Extremes (one procured under another grant), a 486 PC, and a Macintosh Centris 660AV, all installed in a networked environment. Interface devices include a color video projection system capable of field-interlaced stereo display, a Virtual Research EYEGEN3 color helmet mounted display, a three-sensor Polhemus FASTRAK position/orientation sensing system, a VPL Dataglove, a PHANToM 3-axis desktop force-reflecting haptic interface device, and a BEACHTRON sound spatialization system from Crystal River Engineering.

A demonstration virtual environment—a three-dimensional model of the VETT Testbed laboratory—has been implemented and tested with all interface devices. Frame rates for the demonstration VE are approximately 40Hz for bi-ocular viewing (same image presented to each eye) and 20Hz for stereoscopic viewing.

Plans for the remainder of the first year include: (1) characterization of testbed performance (update rate and lag); (2) implementation of a prototype version of the currently envisioned first testbed-based training task (vehicle dynamics training); (3) configuration of the testbed (i.e., architecture, layout, and instrumentation) to support precisely controlled psychophysical experiments and device evaluations; (4) addition of speech-based control; and (5) incorporation of additional visual, auditory, and haptic interface devices.

1.11.2 Sensorimotor Satellite

**Project Staff**

Nathaniel I. Durlach, Professor Richard M. Held, Dr. Thomas E. Wiegand, Julien B. Beasley, Evan F. Wies

Since the start of the VETT project, we have assembled all the equipment required for the planned experiments. This equipment consists of a planar four-degree-of-freedom (DOF) manipulandum and a visual display mounted in a viewing box. The display is mounted in such a way that the visual image representing the state of the manipulandum (i.e., the subject's hand that is grasping the manipulandum) appears in the plane of the manipulandum. The relation between the state of the manipulandum and the visual image representing the manipulandum are under the control of a Macintosh computer and can be programmed to simulate various types of distortions, delays, and stochastic disturbances (noises) that are likely to arise in virtual-environment training systems.

We have built two separate jointed-arm manipulanda for use in the proposed sensorimotor experiments. Each device has four DOFs (x, y, twist, and squeeze) in a planar operating range of about 10 by 14 inches. Manipulandum number 1 also includes a brake on the squeeze dimension to simulate the presence of an object between the grip pads. Manipulandum number 2 was constructed of thin-walled aluminum tubing (without the brake) to
minimize the mass of the moving elements. Which manipulandum will be used will depend on the detailed requirements of the specific experiment under consideration.

During the last half of the first year, we expect to complete a set of preliminary experiments in which we introduce spatial, temporal, and stochastic alterations corresponding to the limitations of available interface devices. These data will enable us to make informed choices of technical tradeoffs and will also allow us to begin developing a model for predicting the effects of such alterations.

A further component of the second half of the year's research effort concerns the investigation of sensorimotor performance involving spatial auditory targets. This work will include installation of enhanced spatial tracking hardware and planning of experiments to be conducted with the new auditory display equipment being developed at Crystal River Engineering (CRE) for delivery during the second year.

1.11.3 Haptics Satellite

Project Staff

Dr. J. Kenneth Salisbury, Dr. Mandayam A. Srinivasan, Dr. David L. Brock, Gerald L. Beauregard, Nick Swarup, Craig B. Zilles, Thomas H. Massie

During the first half of the first year, we have developed device hardware and interaction software and performed psychophysical experiments pertaining to haptic interactions with virtual environments. The Haptics laboratory space in RLE has been reorganized and outfitted with new test and computer equipment as originally proposed. Two major devices for performing psychophysical experiments, the Linear and Planar Graspers, have been extensively refurbished with improved mechanics and sensors. A new force-reflecting gripper to be mounted on the planar grasper has been constructed and is currently being interfaced. Another haptic display device developed previously, the PHANToM, has been used to prototype a wide range of force-based haptic display primitives. It permits precisely controllable forces to be exerted on a user's fingertip through a thimble connection and allows simple tool interactions through a stylus-like grip. Copies of this device, along with demonstration software, were delivered to Naval Air Warfare Center Training Systems Division (NAWCTSD) at the end of December 1993 and to the VETT Testbed in January 1994.

In preparation for the design of a tool-handle haptic interface suitable for a large workspace, a full-scale mockup has been fabricated to determine workspace constraints. An analysis of force/inertia/friction requirements of the device aimed at specifying kinematic, structural and actuator requirements has been initiated. In addition, an operational prototype of a powered screwdriver tool-handle has been built and interfaced to provide an early reality check on the tool-handle concept and feedback useful for design of the full scale version. At the MIT Newman Laboratory, final commissioning and testing of the instrumentation and control package was done for MANUS, a 5-DOF haptic interface device with workspace comparable to the human arm, and a data communication link was established between MANUS and MANTA, a 2-DOF haptic interface device.

We have developed software for the demonstration of haptic displays with the PHANToM, such as interactions with virtual pushbuttons, sliding friction and surface texture displays, virtual screw turning, and virtual grasping, as well as 2-person haptic interactions with two PHANToMs. Software has also been developed for performing new psychophysical experiments with the Linear Grasper. Following analysis of the data obtained in previous experiments on the compliance just noticeable difference (JND) and its relationship to the discrimination of mechanical work and terminal force, a new series of experiments for measuring human discriminability of compliance and work done have been initiated.

During the remaining portion of the first year, we plan to further characterize human performance in the discrimination of fundamental mechanical properties of objects such as compliance, viscosity and inertia. This requires additional hardware improvements to the linear grasper, such as the mounting of tachometers and accelerometers. Other hardware developments consist of the addition of the force reflecting gripper to the Planar Grasper to permit pick-and-place tasks, and the design and construction of tool-handle prototypes 1 and 2. We will provide support for the interfacing of the PHANToM into the Testbed and the performance of psychophysical experiments under multimodal conditions. Our haptic display software efforts will focus on two areas, the packaging of display algorithms for transfer to Testbed use, and the exploration of new display techniques for evoking more subtle contact events and states, such as surface microtexture, stiction, rigidity, and edge sharpness.
Chapter 1. Sensory Communication

1.12 The Virtual Sailor: Synthetic Humans for Virtual Environments

Sponsor
U.S. Navy - Office of Naval Research
Grant N00014-92-J-4005

Project Staff
Dr. David Zeltzer, Svetlana Gaffron

1.12.1 Project Overview

Many virtual environment (VE) applications, such as a system for training team members to perform vehicle maintenance, call for the representation and display of interactive, realtime synthetic humans, which we call virtual actors. To do this, we need to develop adaptive and efficient computational models of human motor behavior.

In the past year, we have undertaken a research program designed to identify the basic issues involved in developing a system, which we call SkillBuilder, for computer-aided design of motor skills for virtual actors. The skill builder will allow us to implement a generic human model and train it to adaptively execute new skills, such as handling and operating a particular tool. As part of this effort, we have implemented a model of visually-guided reaching and grasping and controllers for human head and eye movement based on psychophysical studies. We have demonstrated that these models can support realistic, realtime interaction with human VE participants.

Real-time, interactive behavior models would be of little use without appropriate human interface software that supports task level interaction using speech and gestures. In addition to our work on SkillBuilder, therefore, we have also developed an architecture that we believe will support task level interaction using a restricted natural language interface. This work is based on our previous work on reactive planning systems, as well as the natural language research of Schank and colleagues in the early 1970s, efforts to develop a language for specifying human behavior in a clinical setting, and the notion of affordances first described by J.J. Gibson.

Finally, in addition to modeling human motor behavior for virtual actors, it is important to correctly capture human appearance, especially when close-up interactions are required. Therefore, we have also continued our efforts to develop accurate, realtime finite element models (FEM) of human facial tissue. This year we have developed a set of motor programs that coordinate the actions of our finite element model of the human face, such that we can display a range of facial expressions. Since we cannot update the FEM in realtime, we have also developed a method for pre-computing a set of key expressions and later playing them back in realtime.

1.12.2 Real-Time, Multimodal Interaction With Virtual Actors

The long-range goal of this work is to implement a virtual environment that includes autonomous and interactive synthetic humans. These virtual actors should be capable of interacting with human participants in the VE, as well as with simulated objects and events in the VE.

On July 30, 1993, we were able to demonstrate a prototypical virtual actor capable of responding, in realtime, to physical interaction with human participants using position-sensing devices. In addition, this virtual actor also performed interactive, realtime, visually-guided grasping and reaching behaviors. Finally, the actor was capable of responsive head and eye movements. Not only could it...

While there are a small number of human-modeling efforts underway elsewhere in the U.S. and Europe, our work is unique because of our focus on realtime, adaptive, multimodal interaction, rather than offline scripting of behaviors.

1.12.3 Task Level Interaction With Virtual Actors

In earlier work we developed and implemented a reactive planner, which we call the skill network, for adaptively selecting and executing behavior modules, called motor skills, in order to satisfy some motor goal, say, "Go to the door and open it." To do this, however, we need to identify the motor skills explicitly and implicitly required to accomplish the given motor act. We call this process motor goal parsing (MGP), and this year we have designed an architecture to support this process.

While most humans are capable of seemingly effortless manipulation of ordinary objects in everyday life, computational models of such behavior must account for affordances of objects and motor equivalence. That is, given a command that names one or more objects in the virtual world, how can a virtual actor compute at runtime whether and how the named objects can be manipulated? Next, once the virtual actor has determined that an object can be manipulated, the appropriate effector system must be selected. For example, can the object be grasped with one hand or must both arms be employed?

In the past year, we have described the representations we feel are necessary to transform descriptions of everyday activities into a series of motor goals that can be executed by the skill network. To account for motor equivalence and the affordances of objects in the environment, we have specified an effector-free set of task primitives to be used as an intermediate representation in the parsing process, and we have outlined the process of motor goal parsing.\footnote{D. Zeltzer and M. Johnson, "Virtual Actors and Virtual Environments: Defining, Modeling and Reasoning about Motor Skills," in Interacting with Virtual Environments, eds. L. MacDonald and J. Vince (Chichester, England: John Wiley & Sons, 1994, pp. 229-265).} It is our intention to implement this architecture in the coming year.

1.13 Research on Improved Sonar Displays: A Human/Machine Processing System

Sponsor
U.S. Navy - Office of Naval Research
Grant N00014-93-1-1198

Project Staff
Dr. David Zeltzer, Nathaniel I. Durlach, Dr. Jeng-Feng Lee, Megan C. Jasek

A decrease in the perceived threat from former Soviet-bloc countries and an increasing threat from quiet, comparatively low-cost, diesel/electric submarines and swimmer delivery vehicles have resulted in strong emphasis on littoral region surveillance and anti-submarine warfare by the U.S. Navy. The inshore, shallow-water environment, however, is extremely challenging for passive and active acoustic sensor systems. Many of the simplifications used in modeling and interpreting deepwater acoustic propagation do not hold in the highly reverberant, range dependent and noisy littoral environment. Due to these factors, not only must new sensors and sensor processing methods be developed, but also new human-machine processing systems (HMPS) which best present and exploit existing sensor systems. Emerging remote presence and virtual environment technologies and techniques now provide sufficient framework to support the type of multi-modal, interactive human-machine processing systems that are envisioned.

Ideally, a general-purpose HMPS would allow viewing and interaction with received sonar signals in light of sonar system configuration and performance information, a reasonable acoustic propagation model, current information about the undersea environment, and information about the overall tactical situation. The HMPS would receive two classes of input signals. The first consists of the outputs of underwater acoustic sensing systems (both active and passive). The second consists of the outputs of auxiliary systems designed to produce continuously updated information on the state of the surface environment (e.g., ship posi-
Chapter 1. Sensory Communication

tions) and the underwater environment (e.g., sound velocity profiles).

It is assumed that the first class of signals arises from both active sources in the environment and from reflections of sonar signals off objects in the environment and that these sounds and reflections can arise in connection with hostile targets such as submarines and swimmer-delivery vehicles, as well as various types of clutter, including biologicals, decoys, and commercial traffic. (In this document, we use the word clutter to refer not only to reflections of sonar energy off "irrelevant" objects, but also to interfering active sources.) The output of the HMPS is envisioned as consisting of a probabilistic description of the current environmental situation structured and displayed spatially utilizing multiple sensory modalities in a manner that is well-matched to the decision/action space for command and control. This display would not only facilitate analysis of incoming real-time sonar data but also help sonar personnel understand the acoustical propagation characteristics and tactical import of the current undersea environment.

Within this general context, this research is concerned with (1) determining the optimum processing for the HMPS to perform on the input signals, (2) the manner in which automatic machine processing on the one hand and processing by the human sensory systems and the human brain on the other hand can best be combined, and (3) how to display signals to the human operator to optimize the human processing. Although the focus will be on the last of these three topics, all of these are sufficiently intertwined so that consideration of the first two cannot be entirely eliminated without severely limiting the usefulness of the results.

As a first step in the development of a general purpose HMPS, we will concentrate on real-time, interactive display methodologies for ship-based active sonars. Sonar signals will be synthetically generated and their propagation through the shallow-water environment will be modeled. Inputs to the general acoustic model include bottom characteristics and water column data affecting the range and depth dependent sound velocity profile. In addition, a minimum of one acoustic source and two receivers will be modeled. Major source and receiver parameters which may be varied, in addition to the number of sources and receivers, include location, frequency response, and directional gain characteristics.

Once we have an appropriate model and testbed system up and running, we will then begin to explore display configurations that present information that we believe is particularly relevant to determining the features of the acoustic signal space that are important for ASW and present this information in a manner that takes account of both the strengths and weaknesses of the human processing system and the new technology that is now being developed in the area of human-machine interfaces for virtual environment and teleoperator systems.

Using available technologies, the testbed HMPS will allow modification of the following:

- Characteristics of the displays (in various modalities),
- Relevant undersea environment characteristics,
- Number, position and operating characteristics of sonar system elements,
- Number, position, and characteristics of clutter elements,
- Number, position and operating characteristics of targets.

A variety of spatially oriented representation techniques will be developed and evaluated against existing display techniques (such as "waterfall" techniques) for detection and classification performance. Among the issues to be considered in this proposed research are the following:

1. How best to combine human and machine processing to capitalize on the humans relatively good ability to perform complex pattern recognition tasks and to rapidly adapt to changing conditions, while simultaneously compensating for the humans relatively poor ability to store large amounts of unencoded sensory data and perform large amounts of computation rapidly;

2. How best to combine visual and auditory displays to capitalize on the respective strengths of these two modalities;

3. How best to present intersensor correlation information along with single-sensor information;

4. How best to exploit the sense of immersion and the ability to manipulate the observers point of view that can be provided by virtual environment display systems;

5. How best to partition and display information to groups of humans as well as single individuals; and

6. How best to incorporate decision support aids for localization and identification of sonar signals.
During the first half of the first year, we have evaluated a number of underwater acoustic propagation models, and we have selected a simplified propagation model for use in the program. Design of the preliminary human/machine interface is currently underway.

(Center at left) Senior Research Scientist Nathaniel I. Durlach, (center at right) Professor Louis D.B. Braida, and other staff members and students of RLE’s RLE Sensory Communication Group, some wearing the group’s research devices. Clockwise from left, Visiting Scientist Yun Shao, Research Specialist Hong Z. Tan, student Ross A. Yu, Research Specialist Lorraine A. Delhome, graduate student Barbara G. Shinn-Cunningham, Research Assistant Walter A. Aviles, Administrative Secretary Eleanora M. Luongo, Visiting Scientist Min Wei, and Administrative Secretary Ann K. Dix.
Section 3  Auditory Physiology

Chapter 1  Signal Transmission in the Auditory System
Chapter 1. Signal Transmission In the Auditory System

Academic and Research Staff
Professor Lawrence S. Frishkopf, Professor Nelson Y.S. Klang, Professor William T. Peake, Professor William M. Siebert, Professor Thomas F. Weiss, Dr. Bertrand Delgutte, Dr. Donald K. Eddington, Dr. Dennis M. Freeman, Dr. John J. Guinan, Jr., Dr. William M. Rabinowitz, Dr. John J. Rosowski, Joseph Tierney, Marc A. Zissman

Visiting Scientists and Research Affiliates
Dr. Jennifer R. Melcher, Dr. Sunil Puria, Dr. Jay T. Rubinstein, Dr. Devang M. Shah

Graduate Students
C. Cameron Abnet, Henry E. Chung, C. Quentin Davis, Scott B.C. Dynes, Michael P. McCue, Lisa F. Shatz, Konstantina Stankovic, Susan E. Voss

Technical and Support Staff
Janice L. Baizer, Mark R. Nilsen, Michael E. Ravicz, Frank J. Stefanov-Wagner, David A. Steffens, Meng Y. Zhu

1.1 Introduction
Research on the auditory system is carried out in cooperation with two laboratories at the Massachusetts Eye and Ear Infirmary (MEEI). Investigations of signal transmission in the auditory system involve the Eaton-Peabody Laboratory for Auditory Physiology, whose long-term objective is to determine the anatomical structures and physiological mechanisms that underlie vertebrate hearing and to apply this knowledge to clinical problems. Studies of cochlear implants in humans are carried out at the MEEI Cochlear Implant Research Laboratory. Cochlear implants electrically stimulate intracochlear electrodes to elicit patterns of auditory nerve fiber activity that the brain can learn to interpret. The ultimate goal for these devices is to provide speech communication for the profoundly deaf.

1.2 Signal Transmission In the External and Middle Ear
The goal of our work is to understand the relationship between the structures of the external and middle ear and their functions.

1.2.1 Structure-Function Relations In Middle Ears

Sponsor
National Institutes of Health
Grant R01-DC-00194-11

Project Staff
Dr. John J. Rosowski, Professor William T. Peake, Michael E. Ravicz, David A. Steffens

The middle-ear is a key link in the hearing process. In normal hearing, sound signals are delivered to the inner ear by the middle ear. Disease, age, and injury can degrade the performance of the middle ear. In part of our work, we investigate the relationship between the structures of the middle-ear air space and tympanic-membrane and the function of the middle ear. We are making joint anatomical and physiological measurements in a series of rodent and cat species with different tympanic membrane and middle-ear air-space morphologies. We will also use similar methods in cats to investigate whether the response of the cochlea is completely determined by the difference in pressure across the two cochlear windows.

In the last year, we have moved towards our goal in the following ways:

---

1 Research Affiliate, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts.
Chapter 1. Signal Transmission in the Auditory System

1. We have analyzed our measurements on middle-ear cavity and tympanic membrane function and incorporated the results into a paper presenting new measurements of the middle-ear input impedance in cats\(^2\) and into a model of the mammalian external and middle ear.\(^3\) These analyses support the theory that the primary effect of middle-ear cavities is consistent with a "series" model. The cavities impede the volume-velocity of sound at the tympanic membrane but have little effect on how the membrane and ossicles respond for a given volume velocity. An additional analysis of anatomical and physiological measurements, made on many mammalian species, points out that different species have greatly different middle-ear cavity impedances and that the relative magnitudes of the impedance of the cavity and the impedance of the rest of the ear varies widely among mammals.\(^4\)

2. We have taken the "series" model and our data a step further and analyzed the effect of variations in middle-ear volume on auditory thresholds in a desert rodent (the Mongolian gerbil).\(^5\) This analysis suggests that the middle-ear cavity in a gerbil plays a large role in limiting low-frequency auditory thresholds even though the cavity has a volume larger than the brain case. This conclusion is contrary to the suggestion that the large middle-ear cavities of these and other specialized desert rodents have evolved to permit maximum sensitivity to low-frequency sound.

3. A series of new measurements tested the "classic" assumption that the inner ear responds to the sound-pressure difference between the oval and round window.\(^6\) The external and middle ears of cats were prepared so that we could independently control the magnitude and phase of the sound pressure at the two cochlear windows. The cochlear potential, measured with a round-window electrode, was our indicator of cochlear response. The potential was determined with sound stimuli presented at both or either cochlear window. The results indicate that the difference in sound pressure at the windows is clearly the primary stimulus to the cochlear potential.

The results of our work have direct significance to our understanding of human disease, animal ethology, and the basic hearing process. Chronic and acute otitis media are very common causes of hearing dysfunction. One significant effect of these diseases is a reduction in the air volume of the middle ear. This reduction can be transiently caused by the accumulation of fluid or more permanently produced by the accumulation of debris within the middle ear or actual alterations in the bony walls of the cavity. Our work will enable a clearer understanding of the role of the middle-ear air spaces in both normal and diseased ears.

The ethological issues we are testing concern the role of enlarged middle-ear air spaces. Our work on a number of rodent species and a wide selection of cat species will test the popular idea that large middle-ear air spaces represent an adaptation to arid environments.

Our work on specification of the effective stimulus to the inner ear also has direct clinical consequences in that surgeons routinely manipulate the pressure difference between the cochlear windows when they reconstruct severely diseased middle ears. Our new measurements support the efficacy of such procedures. Furthermore, the results of our measurements have implications for our fundamental understanding of inner ear mechanisms. Any indication that the cochlea is sensitive to stimuli other than the difference in sound pressure at the cochlear windows will require basic changes in cochlear theories.

---


\(^6\) S.E. Voss, J.J. Rosowski, and W.T. Peake, "Is the Pressure Difference Between the Oval and Round Windows the Stimulus for Cochlear Responses?" in Abstracts of the Seventeenth Midwinter Meeting of the Association for Research in Otolaryngology, February 6-10, 1994, p. 87.
1.2.2 Basic and Clinical Studies of Middle-Ear Function

Sponsor
National Institutes of Health
Grant P01-DC00119 Sub-Project 1
Grant F32-DC00073-2

Project Staff
Dr. John J. Rosowski, Professor William T. Peake, Dr. Sunil Puria, Michael E. Ravicz

Our goal of understanding the relationship between the structure and function of the auditory periphery includes defining the effect of middle and external ear pathologies on auditory function. To help achieve this goal, we have continued measurements of function in human ears using temporal bones donated at the time of death, and we have applied the models we determined from our animal work to questions concerning middle-ear pathologies and treatments.

Dr. Puria has led a project to measure the pressure transfer function of the human middle ear. The excised temporal bones were prepared to allow reproducible placement of a calibrated sound source and microphone in the ear canal together with a miniature hydrophone in the cochlear vestibule. The preparation procedure included the placement of "flush" tubes—one near the round window, the other in the vestibular canals—to keep the cochlear fluid filled. These measurements suggest that the human middle ear acts as a bandpass pressure amplifier with a maximum gain of about 20 dB between 1 and 5 kHz. At lower frequencies, the gain is roughly proportional to frequency; at higher frequencies the gain decreases as frequency increases.

With the help of our clinical associates at the Massachusetts Eye and Ear Infirmary, we have expanded our analysis of previous models of diseased and reconstructed middle ears. The results of these analyses were presented at two conferences where we suggested that small modifications in current surgical procedures could lead to improved hearing results. These modifications included: stapes prostheses with larger areas, maintaining a volume of at least 0.5 cc for the reconstructed middle-ear cavity, and stiffer grafts for isolating the round window from sound.

1.3 Cochlear Mechanisms

Sponsor
National Institutes of Health
Contract P01-DC00119
Grant R01-DC00238
Grant R01-DC00473

Project Staff
Professor Thomas F. Weiss, Dr. Dennis M. Freeman, Dr. Devang M. Shah, C. Cameron Abnet, Henry E. Chung, C. Quentin Davis, Lisa F. Shatz

Our goal is to study the mechanisms by which the motions of macroscopic structures in the inner ear produce motions of the mechanically sensitive hair bundles of sensory receptor (hair) cells. Because of its strategic location, the tectorial membrane must play an important role in the mechanical stimulation of hair bundles. However, there have been few direct observations of the tectorial membrane, and its critical properties remain obscure.

Previously, we have measured osmotic responses to isosmotic solutions of different sodium, potassium, and calcium concentration in the isolated chick tectorial membrane. That work has been submitted for publication. However, the structure and biochemical composition of the tectorial membrane shows considerable diversity across species. Therefore, during the past year we have investigated the physicochemical properties of tectorial membranes isolated from the lizard and mouse cochlea using the same techniques. Basically, in this procedure we dissect the tectorial membrane from the cochlea, place it in a chamber through which solutions of known composition are perfused. The tectorial membrane is decorated with microscopic latex beads whose three-dimensional motions in response to solution changes are measured.

---


Chapter 1. Signal Transmission in the Auditory System

ured automatically with a system consisting of a compound microscope, CCD camera, video digitizer, and workstation. Results for all three species (chick, lizard, and mouse) show that increasing calcium concentration from the low values (20 μmol/L) typical for the solution that bathes the tectorial membrane in situ (endolymph) to the higher concentrations (2 mmol/L) found in normal extracellular fluids causes the tectorial membrane to shrink with a time course of about five minutes. The shrinkage is largely reversed when calcium concentration is lowered but with a slower time course (tens of minutes).

Tectorial membranes from all three species also swell when isosmotic, high-sodium solutions are substituted for high-potassium solutions. However, the magnitude of this response is much greater for the chick (mean of 135 percent) than for either the mouse (mean of 14 percent) or the lizard (mean of 10 percent). Swelling in high-sodium solutions is at least partially reversed on return to high-potassium solutions. However, irreversible changes result during long-duration (60-minute) exposures to high-sodium solutions. However, irreversible changes result during long-duration (60-minute) exposures to high-sodium solutions that give rise to large swelling responses.

Our results are consistent with the simple hypothesis that the tectorial membrane is a polyelectrolyte gel with fixed ionizable charge groups that can bind ions differentially. This picture of the tectorial membrane is similar to that for other connective tissues: the binding of ions modulates the fixed charge density which changes the concentration of counter ions in the tissue and hence its water content.

The shrinkage and swelling of the tectorial membrane in response to changes in ion concentration may have implications for hearing. For example, a number of factors are known to change ion composition of the lymphs including acoustic overstimulation, anoxia, systemic administration of ethacrynic acid, or perfusion of the cochlea with ouabain, etc. Our results show that changes in the ionic composition of the bath can lead to changes in the structure of the tectorial membrane. Changes in its structure are likely to cause changes in the delivery of the mechanical stimulus to the hair cells and hence to affect hearing.

1.3.1 Publication


1.4 Stimulus Coding in the Auditory Nerve and Cochlear Nucleus

Sponsor

National Institutes of Health
Grant P01-DC00119
Grant T32-DC00038P

Project Staff

Dr. Bertrand Delgutte, Dr. Peter A. Cariani

We are investigating the neural mechanisms underlying auditory perception at the level of the auditory nerve and cochlear nucleus. In the past year, we have focused on temporal patterns of neural discharge that convey pitch and speech information. These temporal patterns are analyzed by recording the responses of auditory nerve fibers and cochlear nucleus neurons in anesthetized cats. All-order interspike interval distributions (autocorrelation histograms) are computed and the properties of these distributions are compared with psychophysical data.

Our previous results had shown that the pitch of a complex tone corresponds to the most frequent interspike interval in the auditory nerve for a wide variety of stimuli. Responses of cochlear nucleus units to the same stimuli were recorded to determine whether pitch might also be represented by interspike interval patterns there. Virtually all single units in the cochlear nucleus that responded to the stimuli showed pitch-related intervals, and these were usually the most frequent in the distribution. Thus, cochlear nucleus units might convey pitch information to more central auditory stations through an interspike interval code.

We have continued our investigation of the neural mechanisms underlying the perceptual segregation of concurrently-presented vowels. Human listeners identify pairs of concurrent vowels better when they differ in fundamental frequency (F0) than when they have the same F0. Previously we had found that the constituents of a vowel pair could be identified from the aggregate interspike interval distributions of the auditory nerve with an accuracy comparable to human perceptual judgements. During the past year, we sought to explain how vowel identification might be improved for pairs with different F0s without requiring an explicit determination of F0. We found that differences in F0 create changing alignments in the onsets of vowel periods, and that these changing alignments create transient changes in interspike interval distributions of auditory-nerve fibers. One possibility is that these changes allow
Chapter 1. Signal Transmission in the Auditory System

the auditory system to perform a spectral analysis over relatively short (5-10 ms) time windows and integrate these analyses over time to identify the vowel constituents. Along these lines, we developed a new algorithm for identifying the constituents of a double vowel from running interspike interval distributions. This algorithm is consistent with human perceptual data showing effects of period asynchronies on the identification of concurrent vowels.

1.4.1 Publication

1.4.2 Binaural Interactions In Auditory Brainstem Neurons

**Sponsor**
National Institutes of Health
Grant P01-DC00119

**Project Staff**
Dr. Bertrand Delgutte, Dr. John J. Guinan, Jr., Dr. John J. Rosowski

Several acoustic cues are important for sound localization, including interaural time (ITD), interaural level (ILD) differences, and spectral shape. In collaboration with Dr. T.C.T. Yin and his colleagues at the University of Wisconsin in Madison, we are studying the relative contribution of these cues to the directional sensitivity of neurons in the inferior colliculus (IC). The activity of single units was recorded in anesthetized cats in response to "virtual space" (VS) stimuli delivered through closed acoustic systems. These stimuli were generated by passing broadband noise through digital filters constructed from head-related transfer functions measured in one cat. We focused on high-CF (> 5 kHz) units, which are likely to be sensitive to spectral features.

The sensitivity of neurons to stimuli whose virtual position varied along the horizontal and vertical planes generally resembled that found in previous free-field studies. For most units, responses to monaural VS stimuli were less directional than those for binaural stimuli, suggesting that binaural interactions were important for the directional response of these neurons. The type of binaural interactions (e.g., EE, EI, or mixed) determined with VS stimuli was consistent with that found for broadband noise.

We also generated modified VS stimuli in which one or more of the localization cues (ITD, ILD, or spectrum) were set constant for all locations. For example, "0-ITD" stimuli were constructed by delaying the waveform in the leading ear to set the ITD to zero for every azimuth. "Delta-ITD" stimuli were created by taking the waveforms for the position directly in front, then introducing the appropriate ITD for each azimuth. Results suggest that, for most of these high-CF units, ILD is the most potent cue, followed by spectral shape, then by ITD. These results are consistent with human psychophysical experiments showing the dominance of ILD cues for high-frequency stimuli. At a more general level, our results demonstrate the usefulness of virtual-space techniques for bridging the gap between free-field studies of the directional sensitivity of neurons and dichotic studies of ITD and ILD sensitivity.

1.4.3 Publication

1.4.4 Electrical Stimulation of the Auditory Nerve

**Sponsor**
National Institutes of Health
Grant P01-DC00361

**Project Staff**
Dr. Bertrand Delgutte, Scott B.C. Dynes

We are studying physiological mechanisms of electrical stimulation of the cochlea because this information will help design improved processing schemes for multiple-channel cochlear implants. During the past year, we have studied auditory-nerve fiber correlates of interactions observed psychophysically when pulsatile electric stimuli are applied in rapid succession. We measured the responses of auditory-nerve fibers in anesthetized cats for pairs of monophasic pulses separated by short intervals. Results show that neural interactions between successive current pulses occur on three different time scales:
1. The shortest interactions occur when the first pulse in a pair is subthreshold. For cathodic first pulses, the threshold for following cathodic pulses was decreased, while anodic first pulses increased the threshold of following cathodic pulses. These results are qualitatively consistent with viewing the first pulse as leaving a residual charge on the neural membrane. The time constant of these capacitive effects is 100-200 usec. These subthreshold interactions are likely to play an important role in continuous interleaved sampling (CIS) processors for cochlear implants because most pulses are probably subthreshold in this situation, so that elicitation of a spike would require interaction over multiple pulses. Particularly interesting is the result that subthreshold interactions for biphasic pulses similar to those used in CIS processors cannot be predicted from simple ideas about charge integration by the neural membrane. On the other hand, these interactions are consistent with psychophysical data showing sensitization for biphasic pulses preceded by either anodic/cathodic or cathodic/anodic pulses. Thus, the neural and psychophysical data are consistent with each other, but both are at variance with current theoretical conceptions of neural excitation.

2. When the first pulse was suprathreshold, the threshold for the second pulse was found to decrease with increasing interpulse delay, as expected from the refractory properties of nerve fibers. The decrease was such that at 2 msec interpulse delay, the threshold for the second pulse averaged about 5 dB greater than the threshold for a single pulse. Thus, these interactions due to suprathreshold pulses last longer than subthreshold interactions. These suprathreshold interactions may not severely limit the performance of CIS processors providing that there is a sufficient pool of nerve fibers that discharge at different times (volley principle).

3. In addition to interactions between two successive pulses, further results provide evidence for the existence of long-term interactions occurring cumulatively over multiple pulses. Because these multipulse interactions have not been systematically characterized, it is difficult at this point to assess their significance for CIS processors. We plan to examine these interactions in more detail during the next year.

1.4.5 Publications


1.5 Interactions of Middle-Ear Muscles and Olivocochlear Efferents

Sponsor
National Institutes of Health
Contract PO1 DC00119

Project Staff
Dr. John J. Guinan, Jr.

Our aim is to determine the actions and interactions of acoustically elicited middle-ear muscle reflexes and olivocochlear efferent reflexes.

During the past year, we have analyzed data and prepared for publication our results on the innervation of the stapedius muscle. In a previous study using intracellular labeling of physiologically identified stapedius motoneurons, we showed that there is functional spatial segregation in the stapedius motoneuron pool. In the same animals, we have now traced each sufficiently labeled axon


into the stapedius muscle to determine whether a similar functional spatial segregation is present within the muscle. Ten labeled axons of physiologically identified stapedius motoneurons were visible in the facial nerve and five could be traced to endplates within the stapedius muscle. In one case, a stapedius motoneuron innervated only a single muscle fiber; we think this is the first documented case of such remarkably fine-grained motor innervation. Overall, there were 39 observed branches from the five axons (we may have missed others). This indicates an average innervation ratio (≈7.8) which is much higher than that obtained from present estimates of the numbers of stapedius motoneurons and muscle fibers. Muscle fibers innervated by a single axon were spread over a wide area in the muscle, suggesting that spatial segregation in the stapedius muscle is unlikely. Thus, the reasons for the central functional spatial segregation in the stapedius motoneuron pool are more likely related to organizing factors which originate in the brain rather than to organizing factors which originate in the muscle.

1.5.1 Publication


1.6 Cochlear Efferent System

Sponsor

National Institutes of Health
Grant 2RO1 DC00235

Project Staff

Michael P. McCue, Konstantina Stankovic

Our aim is to understand the physiological effects produced by efferents in the mammalian inner ear including medial olivocochlear efferents which terminate on outer hair cells and lateral efferents which terminate on auditory-nerve fibers.

During the past year, we have analyzed data and prepared for publication results on our newly-discovered class of vestibular primary afferent neurons, fibers which respond to sounds at moderately high sound levels. Like their cochlear homologues, these vestibular afferent fibers receive efferent projections from brain-stem neurons. We have explored efferent influences on the background activity and tone-burst responses of the acoustically-responsive vestibular afferents. Shock-burst stimulation of efferents excited acoustically-responsive vestibular afferents; no inhibition was seen. A fast excitatory component built up within 100-200 ms of shock-burst onset and decayed with a similar time course at the end of each shock burst. During repeated 400 ms shock bursts at 1.5 s intervals, a slow excitatory component grew over 20-40 s and then decayed, even though the shock bursts continued. Efferent stimulation excited acoustically-responsive vestibular afferents without appreciably affecting their sound-evoked responses. This provides strong evidence that the excitation is due to efferent synapses on afferent fibers instead of efferent synapses on hair cells. Efferent stimulation enhanced the tone-induced within-cycle modulation of discharge rate (i.e., increased the a.c. gain) without changing the degree of phase locking to low frequency tones as measured by the synchronization index; i.e., there was little or no improvement in the bidirectionality (linearity) of nerve fiber output. The acoustically-responsive vestibular afferents provide a mammalian model for studying purely excitatory efferent effects in a hair cell system. Anatomically, these vestibular efferent synapses resemble lateral olivocochlear efferent synapses on cochlear nerve fibers, which suggests that lateral efferents may have an excitatory effect. A paper describing these results has been submitted.

During the past year, we have begun work on a project to compare efferent-evoked effects on auditory-nerve fibers with different spontaneous rates. We have implemented and done preliminary experiments with a paradigm in which auditory-nerve-fiber rate versus sound level functions with and without efferent excitation are obtained with randomized presentation of both the sound level and the presence of efferent stimulation.


Chapter 1. Signal Transmission in the Auditory System

1.6.1 Publications


1.7 Cochlear Implants

Sponsor

National Institutes of Health
Grant PO1-DC00361
Contract NO1-DC2-2402

Project Staff

Dr. Donald K. Eddington, Dr. William M. Rabinowitz, Dr. Jay T. Rubinstein, Joseph Tierney, Marc A. Zissman

1.7.1 Project A: Models of Current Spread and Nerve Excitation During Intracochlear Stimulation

Project Staff

Dr. Donald K. Eddington, Dr. Jay T. Rubinstein

The basic function of a cochlear prosthesis is to elicit patterns of activity on the array of surviving auditory nerve fibers by stimulating electrodes that are placed in and/or around the cochlea. By modulating patterns of neural activity, these devices attempt to present information that the implanted subject can learn to interpret. The spike activity patterns elicited by electrical stimulation depend on several factors: the complex, electrically heterogeneous structure of the cochlea, geometry and placement of the stimulating electrodes, stimulus waveform, and distribution of excitable auditory nerve fibers. An understanding of how these factors interact to determine the activity patterns is fundamental to designing better devices and interpreting the results of experiments involving intracochlear stimulation of animal and human subjects. As a first step toward this understanding, the goal of this project is to (1) construct a software model of the cochlea that predicts the distribution of potential produced by the stimulation of arbitrarily placed, intracochlear electrodes and (2) use these potential distributions as inputs that drive models of auditory nerve fibers.

This year, we refined our nonlinear, single-unit model to include a passive, myelinated cell body (see figures 1 and 2). Varying the membrane capacitance and resistance to simulate variations in the myelination of the cell body produces profound effects on conduction. Without myelination, spikes will not propagate across the cell body unless voltage-sensitive sodium channels are present in the soma. With increasing myelination, spikes conduct in the orthograde direction but not retrograde (see figures 3 and 4). A model cell body with myelin thickness comparable to that of the internode propagates spikes in both directions. These simulations demonstrate the biophysical feasibility of physiologically observed, somatic spike rectification and implicate the anatomic asymmetry of the nodes of Ranvier adjacent to the soma as a possible mechanism.

A multichannel stochastic model of the mammalian node of Ranvier has also been developed. This model represents N sodium channels in parallel with a membrane capacitance and leakage resistance. Each channel consists of a single-channel conductance obtained from the patch-clamp and fluctuation literature and a gate controlled by three m (activation) particles and one h (inactivation) particle. Particle open and closed times are random (exponential distribution) with mean obtained from the reciprocal of the particle's rate constant. Thus the open/closed time probability density functions are time-varying. The model has been tested by substituting Frankenhauser-Huxley kinetics and membrane parameters for the mammalian data and comparing the relative spread of threshold (RS) predicted by the model with that measured for peripheral amphibian nodes. The experimental data are very closely approximated by the simulations. This is the first direct numerical verification that the microscopic sodium channel fluctuations are sufficient to account for the macroscopic fluctuations of threshold to electrical stimulation.
Chapter 1. Signal Transmission in the Auditory System

Structure of Model for the Cat Type 1 Spiral Ganglion Cell: 17 nodes of Ranvier, 4 mm total length

![Diagram of a single-unit model with cell body, peripheral processes, and central process separated by node/internode structures.]

**Figure 1.** Overall structure of the single-unit model that includes a representation of the cell body.

McKee-type non-linear distributed circuit representation of nodes, internodes, paranodes, unmyelinated process & cell body

**Figure 2.** Node/internode structure of the single-unit model. Note that the variable conductance at the internode is a nonlinear function based on the characteristics of voltage dependent sodium channels.
Figure 3. Computed membrane potential as a function of time for the pictured model nodes. The stimulating electrode's position is indicated by the black oval and the cathodic, monophasic stimulus waveform is shown below the pictured model. Note that the spike conducts across the cell body in the orthograde direction.

Figure 4. Computed membrane potential as a function of time for the pictured model nodes. The stimulating electrode's position is indicated by the black oval and the cathodic, monophasic stimulus waveform is shown below the pictured model. Note that even though the model parameters are the same used in the model of figure 3, retrograde conduction is blocked at the cell body.
1.7.2 Project B: Psychophysics of Intracochlear Electrical Stimulation

Project Staff
Dr. Donald K. Eddington

One goal of our psychophysical studies is to provide insight into the mechanisms that limit device effectiveness. Using a masking paradigm, we demonstrated in previous years that electrical stimuli delivered simultaneously by two intracochlear electrodes produce strong interactions and that the strength of these interactions across subjects is negatively correlated \((r=-0.8)\) with their speech reception scores. Processing systems that interleave stimuli in time reduce these interactions and produce improved speech reception in many subjects. This year, we focused on measures of forward masking to investigate the strength of nonsimultaneous interactions.

The electrical stimuli used in the forward masking experiments were single, charge-balanced biphasic pulses (phase duration 50\(\mu s\)) presented to the same intracochlear electrode. The masker amplitude was constant at \(-2\) dB below the unmasked threshold with the anodic phase first and cathodic phase second \((ac)\). The following probe stimulus was presented with cathodic phase first \((ca)\), and its amplitude was varied using a two-down, one-up adaptive procedure in a three alternative, forced choice context to measure threshold. For this ac-ca masker-probe condition, the single-unit model predictions are consistent with the subject data (figure 5). The subthreshold masker reduces threshold by as much as 6 dB for a 100 us masker/probe delay. This effect diminishes with increased delay and is negligible for delays longer than 800 us.

The data plotted in figure 6 are from the same subject using a cathodic-phase-first, anodic-phase-second masker. In this condition \((ca-ca)\), the masker tends to decrease threshold less than the ac-ca case and its effect lasts longer than 3 ms. Note that the single-unit model's prediction for the ca-ca case is not consistent with the psychophysical results. This inconsistency will be explored in future single-unit studies in a cat. Because the nonsimultaneous interactions are significant, we are beginning to explore stimulation strategies that will minimize their impact.

Figure 5. Forward masking results for an anodic-phase-first, biphasic masker followed by a cathodic-phase-first, biphasic probe \((ac-ca)\). Probe threshold is plotted in dB relative to the nonmasked threshold and as a function of the delay between the masker and the probe. The subject results are represented by the symbols and the single-unit model predictions by the solid line.
1.7.3 Project C: New Sound Processors for Auditory Prostheses

Project Staff

Dr. Donald K. Eddington, Dr. William M. Rabinowitz, Joseph Tierney, Marc A. Zissman

This project is directed at the design, development, and evaluation of speech processors for use with implanted auditory prostheses in deaf humans. Work this year has focused in three areas: (1) developing a programmable, real-time, speech processing facility that can implement processor/stimulator algorithms for two independent eight-channel implants, (2) using the laboratory system to explore new sound processing algorithms in interactive test sessions with implanted subjects, and (3) designing and fabricating a programmable, wearable speech processor/stimulator that will allow subjects full-time use of processor algorithms developed and tested in our laboratory setting.

One focus of this year's testing of subjects was the exploration of input automatic-gain control. The area of signal level control/compression is important in cochlear implant systems because the wide dynamic range of acoustic signals (~100 dB) must be compressed to the relatively small dynamic range of electrical stimulation (~25 dB).

Using a real-time digital simulation of the subjects' commercial processor, the effects of systematic modifications of the AGC's dynamic characteristics were studied. Certain AGC manipulations have produced substantial increases in consonant identification. For example, on the identification of 16 intervocalic consonants (/a/-C-/a/ tokens from the Iowa videodisc), one subject moved from 32 percent correct with the standard commercial processor to 52 percent correct when the AGC release time was decreased from 250 to 50 us. A second subject improved from 35 to 68 percent correct. A third subject, who is an excellent performer with his standard processor, was tested with 24 consonants; his score increased from 82 to 92 percent correct. One practical reason these results are important is because they suggest relatively simple changes in the subjects' commercial processor that may lead to significant improvements in speech reception.

1.7.4 Publication

Section 4  Linguistics

Chapter 1  Linguistics
Chapter 1. Linguistics

Academic and Research Staff
Professor Morris Halle, Professor Noam A. Chomsky

1.1 Introduction

The work of the Linguistics group is directed towards a better understanding of the mental capacities of human beings through the study of the nature, acquisition and use of language. Language is a uniquely human faculty: only humans appear to be capable of learning and using a language, and every normal human acquires knowledge of one or more languages.

We are trying to understand how this linguistic knowledge is represented in the speaker's mind. The central issues of linguistics research are:

1. What is the nature of linguistic knowledge? What do speakers of a particular language such as Latvian, Spanish or Walpiri know, and how does knowledge of one language resemble or differ from that of another language?

2. How do speakers acquire this knowledge?

3. How do speakers put this knowledge to use in producing and understanding utterances?

4. What are the physiological mechanisms that provide the material basis for storage, acquisition and utilization of linguistic knowledge?

Our ability to answer these questions differs considerably, and our research reflects these differences. At present, we have progressed further with regard to answering the questions posed by item one and have made less progress with item four. Currently, our research is heavily concentrated on issues concerned with the nature of the knowledge that characterizes fluent speakers of various languages. However, we are making a significant effort to solve the other questions also.

We are studying these topics along a number of parallel lines. Linguists have investigated the principles by which words are concatenated to form meaningful sentences. These principles have been the primary domain of inquiry into the disciplines of syntax and semantics. Phonology studies the sound structure of words while morphology examines the manner in which different languages combine different meaning-bearing units (typically, stems, prefixes, suffixes and infixes) to form words. The latter topic has attracted increasing interest in recent years and will probably become more prominent in our research efforts in the future.

1.2 Abstracts of Doctoral Dissertations

The following are abstracts of dissertations submitted in 1993 to the Department of Linguistics and Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics.

1.2.1 Binding as an Interface Condition: An Investigation of Hindi Scrambling

Douglas Arnold Jones

Abstract

This thesis provides an analysis of scrambling, following the Minimalist program of Chomsky 1992, examining Hindi as a case study and occasionally making comparisons with Japanese.

The first major claim is that there are two kinds of operations that yield scrambling. One type, called outer scrambling, is an operator-variable construction and is the result of movement. The other type, called Inner scrambling, is the result of several conspiring factors that allow permutation without movement. These factors are triggered by (1) the strength feature of functional heads (2) economy principles that hold at phonological form and (3) full interpretation, which causes the Agr structure to be deleted after its agreement relations have been established. These factors cause TnsP to collapse into an N-ary branching structure containing the arguments of the verb. It is assumed that linear order is not fixed until spellout. Therefore, the N-ary branching structure is free to yield and hence no trace of movement, Inner scrambling is not subject to reconstruction effects. Outer scrambling, since it is an operator-variable construction, is subject to reconstruction effects.
Chapter 1. Linguistics

The second major claim is that the subject is special in that at least part of it must raise at LF to the checking domain of C degrees. If the subject cannot be split, for example, when the subject is a monomorphemic pronoun, the entire subject must raise. The result is a novel solution to a problem for the minimalist framework, namely, an account of how both strong and weak crossover may be treated as interface conditions at Logical Form in scrambling constructions.

Furthermore, these structures are integral to an account which captures the subject-orientation of reflexive anaphors. The account also captures the lack of subject orientation for reciprocal anaphors.

1.2.2 Topics in Ewe Syntax

Christopher Thad Collins

Abstract

This thesis analyzes a number of problems in Ewe syntax. Its goal is to show how several difficult problems in Ewe syntax have natural treatments in the generative grammar framework, and how these problems bear on current theoretical issues.

Chapter one gives a brief overview of the assumptions that are adopted in this thesis. The assumptions of this thesis are largely those of Chomsky (1992) and Hale and Keyser (1993).

Chapter two gives an analysis of determiner doubling in Ewe. In this construction the third person singular object pronoun doubles a DP that does not have structural case. Determiner doubling as analyzed as a kind of default case assignment. Its precise distribution has important consequences for the grammar of Ewe, including the analysis of serial verb constructions (SVCs), the verbal noun constructions, and A'-movement.

This system of default case found in Ewe will be situated in a general theory of default case, including Yoruba ni: and Russian instrumental. The implications of default case are drawn for the general theory of case given in Chomsky (1992).

Chapter three gives a theory of serial verb constructions (SVCs). The main assumptions that are made are that SVCs involve LF incorporation and that "argument sharing" is mediated by empty categories. These two assumptions lead to an account of many subtle phenomena concerning SVCs.

Chapter four gives an analysis of how the form of the third person singular subject pronoun depends on movement to Spec CP. The analysis is extended to successive cyclic movement, where it is shown that principles of economy of derivation play a role.

1.2.3 Minimalism in Syntactic Derivation

Toshifusa Oka

Abstract

This thesis proposes a theory of how syntactic derivations are constrained. Within the minimalist program of Chomsky 1992, linguistic expressions are optimal realizations of interface conditions. Optimal derivations have minimal cost. The cost of a derivation is determined by principles of economy.

A novel formulation of the system of economy principles is proposed. For example, the shortest steps principle of Chomsky (1992) is replaced by the shallowness principle, which requires that operations must be the shallowest in the phrase markers, and other principles such as the procrastinate principle are redefined. It is also proposed that economy principles are ordered in a certain way. Furthermore, a system of feature checking is proposed, generalizing abstract case to PP complements and adjuncts.

As a consequence, wh-movement phenomena that have been treated as independent phenomena, including the superiority, nesting, and ECP effects, are explained within a system based purely on economy basis, crosslinguistic and intralinguistic variations being derived from parameterized morphological properties.

1.2.4 A/A-Bar Partition Derivation

Hiroaki Tada

Abstract

This thesis investigates some relations of the problem of A/A-bar partition of movement types and the proper formulation of economy principles.

In chapter 2, I will explore two interrelated problems of scrambling. The first one, called the landing site problem, is why and how the apparently single operation scrambling behaves A-movement and A-bar-movement. The second one, called the optionality problem, is whether scrambling is an optional operation, and if so, why it can violate the last resort principle. The answer to the first question leads to the derivational view of A/A-bar partition. The answer to the second question leads to a
system of effect-based economy principles. In particular, the last resort principle is replaced by the compensation principle which requires the external effect of movement to be compensated for by its internal effect.

In chapter 3, I will discuss the problem of uniformity of chain-steps and suggest that the uniformity can be derived from step-wise application of compensation principle, and discusses complex cases such as German partial wh-movement and Mayan focus antipassives in light of the compensation principle and derivational view of A/A-bar-partition.

In chapter 4, I will discuss two phenomena related to the problem of multiple specifier in A and A-bar systems: extraction out of multiple specifier configuration and absorption of multiple wh-phrases. They are claimed to support Cheng's (1991) view that wh-movement is not driven by feature-checking as NP-movement is.

1.2.5 Agr-Based Case Theory and Its Interaction with the A-Bar System

Akira Watanabe

Abstract

This thesis proposes a modification of the Agr-based case theory of Chomsky (1992) to deal with the phenomena of case absorption. The proposed hypothesis claims that Agr must undergo further feature checking with an appropriate functional head after case checking takes place in AgrP. For this reason, CP is needed immediately above Agr-sP for nominative and null case checking; accusative case checking needs TP in simplex clauses and what will be called HP in participle constructions. When these projections are missing, case checking becomes impossible, resulting in case absorption. This explains the distribution of PRO and ECM/raising with respect to the case checking process in Agr-sP. Accusative case absorption in the participial passive and what will be called the reduced causative also fall under our account. This system is extended to the case of pre/post-positions.

Given the pivotal role of Tns in the proposed system, a dependency is expected between the feature checking in Agr-sP and the feature checking in Agr-oP. Some such cases from Irish, Japanese, and Icelandic are discussed.

The proposed modification requires the V+Infl complex to be raised to C", creating a configuration where the A-system including case and inflection interacts with the A-bar movement which makes use of Spec of CP. Wh-agreement and the comp-trace phenomena are given a uniform treatment from this perspective. At the same time, the modified case theory questions the blocking effect of negation for head movement. Instead, the influence of negation on the inflectional system is captured in terms of modality.
Appendices

Appendix A  RLE Publications and Papers Presented
Appendix B  Current RLE Personnel
Appendix C  Milestones
Appendix D  RLE Research Support Index
Appendix A. RLE Publications and Papers Presented

The first section of this bibliography includes papers and talks presented by RLE faculty, staff and students during 1993 and is in alphabetical order by conference name. Reprints of these papers may be obtained by contacting the authors directly.

Section 2 includes an alphabetical listing by author of journal articles that were published and accepted or submitted for publication. Book chapters by RLE authors are listed in Section 3. Section 4 of this bibliography lists RLE general publications and technical reports, and Section 5 is a list of RLE theses submitted during 1993. Section 6 contains miscellaneous publications.

A.1 Meeting Papers

A.1.1 Papers Presented


Liu, S.A. "Locating Landmarks in Utterances for Speech Recognition."


Stadler, R.W., and W.M. Rabinowitz. "On the Potential of Fixed Arrays for Hearing Aids (poster session)."

Takeuchi, A.H. "Perceptual Organization of Temporal Patterns in Music."


Acoustical Society of America Meeting, 126th, Denver, Colorado, October 4-8, 1993.

Chang, H.-P. "Speech Input for Dysarthric Users."


Wilde, L. "Inferring Articulatory Movements from Acoustic Properties at Fricative-vowel Boundaries."


Ceyer, S.T. "A New Mechanism for Dissociative Chemisorption on Si: Atom Abstraction."


Joffe, M. "Experiments Captured in a Dark Light Trap."


Appendix A. RLE Publications

In$_{0.55}$Ga$_{0.47}$As/In$_{0.55}$Al$_{0.45}$As Double Quantum Wells.*

American Physical Society, Division of Plasma Physics, 35th Annual Meeting, St. Louis, Missouri, November 1-5, 1993.


Migliuolo, S., B. Rogers, and B. Coppi. “Removing the External Mode Obstacle to the Second Stability Region.”


Vacca, L. “Monte-Carlo Simulation of Ion Cyclotron Heating and Transport in Tokamak Plasmas.”


Isabelle, S. “Some Issues in Chaotic Time Series Analysis.”

Wornell, G. “Wavelets and Wireless Communications (Invited talk).”

Association for Research in Otolaryngology, 18th Mid-Winter Meeting, St. Petersburg, Florida, February 6-11, 1993.

Puria, S., T. Kawase, J.J. Guinan, Jr., and M.C. Liberman. “Effects of Efferent Mediated Sound Suppression: Cap Versus Distortion Products.”


Haus, H.A. “Nonlinear Propagation in Fibers (Invited talk).”

ARPA Semiannual Principal Investigator Meeting, Washington, DC, April 14-18, 1993.

Devadas, S. “Probabilistic Manipulation of Boolean Functions.”


Smith, H.I. “Lithography: Precision, Accuracy, and Pattern Fidelity.”

Smith, H.I., and S.D. Hector. "X-ray Lithography for High Speed Devices (invited talk)."


Boston University Department of Astronomy Colloquium, Boston, Massachusetts, February 8, 1993.

Hewitt, J.N. "Gravitational Lenses, Time Delays, and Hubble's Constant."


Ram, A.K., and A. Bers. "Hamiltonian Chaos in Wave-Particle Interactions."


Dynes, S.B.C., and B. Delgutte. "Neural Response to Nonsimultaneous Electrical Stimuli: Physiological Results."

Greenberg, J. "Noise Reduction with Multimicrophone Arrays."


Ekstrom, C. "Separated Beam Experiments with an Atom Interferometer."


Ketterle, W., K. Davis, M. Joffe, A. Martin, and D. Pritchard. "High Densities of Cold Atoms in a Dark Spontaneous-force Optical Trap."


Sun, C.-K., H.K. Choi, C.A. Wang, and J.G. Fujimoto. "Femtosecond Carrier Dynamics and Stimulated Transitions Induced Carrier Temperature Change in InGaAs/AlGaAs Strained-layer SQW Diode Lasers."


Copper Mountain Conference on Multigrid Methods, Sixth, Copper Mountain, Colorado, April 4-9, 1993.

Reichelt, M.W. "Optical Convolution SOR Acceleration of Waveform Relaxation with Application to Semiconductor Device Simulation."
Appendix A. RLE Publications


Diode Laser Technology Program Meeting, Albuquerque, New Mexico, April 19, 1993.
Rediker, R.H. "History and Projection of and into the Future of Semiconductor Lasers."

Telichevesky, R., J.K. White, and W.J. Dally. "O*SA Arrays for Parallel Sparse Matrix Factorization."


Fujimoto, J.G. "Optical Coherence Tomography for Imaging and Tumor Detection."


Ram, A.K., A. Bers, and V. Fuchs. "Lower Hybrid Current Drive in the Presence of Ion Cyclotron Waves."

del Alamo, J.A., and S. Bahl. "InAlAs/InGaAs HFETs with High Breakdown Voltage."

Experimental Conference on Chaos, Second, Washington, DC, October 5-6, 1993.
Oppenheim, A.V. "Chaotic Signals in Communication (Invited talk)."


German Physical Society Meeting, Berlin, Germany, March 1993.

Natarajan, V., M.P. Bradley, F. DiFilippo, and D.E. Pritchard. "Single-Ion Mass Spectrometry at 0.1 ppb in a Penning Trap (poster session)."

Appendix A. RLE Publications


Ceyer, S.T. "A New Mechanism for Dissociative Chemisorption on Si: Atom Abstraction."

Ground Target Modeling and Validation Conference, Marquette, Michigan, August 24-26, 1993.


Harvard University Department of Physics Colloquium, Cambridge, Massachusetts, March 1, 1993.

Hewitt, J.N. "Gravitational Lenses, Time Delays, and Hubble's Constant."


Haus, H.A. "TDM Soliton Transmission and Storage."

Intelligent Vehicles Conference, Tokyo, Japan, July 1993.

Harris, J.G. "Analog Chips and Intelligent Vehicles."

International Invitational Symposium on the Unification of Analytical, Computational, and Experimental Solution Methodologies, Danvers, Massachusetts, August 18-20, 1993.

Freeman, D.M. "Micromechanics of the Inner Ear."


Monteiro, J., S. Devadas, and A. Ghosh. "Retiming Sequential Circuits for Low Power."


Bekefi, G. "Observations of Frequency, Phase and Saturation Characteristics of a Raman, Free-Electron Laser Amplifier."


Pritchard, D.E. "Atom Interferometry."


Hsu, C.C., R.T. Shin, J.A. Kong, A. Beaudoin, and T. Le Toan. "Application of Theoretical
Appendix A. RLE Publications

Model for Microwave Remote Sensing of Forest.


Ram, A.K., A. Bers, and V. Fuchs. "Lower Hybrid Current Drive in the Presence of ICRF Waves."


International Symposium on Antennas and EM Theory, Third, Nanjing, China, September 6-9, 1993.


International Symposium on Biomedical Optics (Europe '93), Budapest, Hungary, September 1-5, 1993.


Della Ratta, A.D., J. Melngailis, and C.V. Thompson. "Focused Ion Beam Induced Deposition of Copper."


Xu, X., and J. Melngailis. "Quasi-periodic Nanostructures in Focused Ion Beam Deposited Tungsten at High Angles of Incidence."

International Symposium on Radio Propagation, Beijing, China, August 18-21, 1993.


International Symposium on Recent Advancement in Microwave Technology, New Delhi, India, December 15-18, 1993.


International Television Symposium and Technical Exhibition, 18th, Montreaux, Switzerland, June 1993.

Lim, J.S. "HDTV System with Multiple Transmission Formats."


Han, H.C., and J.A. Kong. "Radiative Transfer Model for Active and Passive Remote Sensing of Vegetation Canopy."

Yuan, Y., J.A. Kong, P.S. Kao, and R.T. Shin. "Theoretical and Experimental Analysis of Multiple Antennas Attached to Coated Bodies of Revolution."


Kong, J.A. "Theoretical Model for Remote Sensing of Vegetation."


Smith, H.I. "X-ray Lithography at 100 nm and Below (invited talk)."


IEEE Design Automation Conference, 30th, Dallas, Texas, June 14-17, 1993.


Nicolas, J.J., and J.S. Lim. "Equalization and Interference Rejection for the Terrestrial Broadcast of Digital HDTV."


Richard, M.D. "Properties and Discrimination of Chaotic Maps."


Damask, J.N., V.V. Wong, J. Ferrera, H.I. Smith, and H.A. Haus. "Optical Distributed-Feedback Channel-Dropping Filters: Design and Fabrication (invited talk)."


IEEE MTT-S International Microwave Symposium, Atlanta, Georgia, June 14-18, 1993.


Appendix A. RLE Publications

Coppi, B. "High Energy Particles."


Arias, T.A. "Impurity Segregation at Semiconductor Grain Boundaries."


Kolodziejski, L.A. "Photo-Assisted Metalorganic Molecular Beam Epitaxy of ZnSe."

Meeting on Codes for Simulating Collective Fast Particle Effects, St. Louis, Missouri, October 31, 1993.

Coppi, B. "Transition in the Emission Spectrum Above the Ion Cyclotron Frequency of Fusion Reaction Products."

Microelectronic Engineering Conference, Maastricht, the Netherlands, September 27-29, 1993.

Hector, S.D., H.I. Smith, N. Gupta, M.L. Schattenburg. "Optimizing Synchrotron-based X-ray Lithography for 0.1 μm Lithography."

Midwest Symposium on Circuits and Systems, 36th, Detroit, Michigan, August 15-18, 1993.


MIT Haystack Observatory Seminar, Lexington, Massachusetts, March 5, 1993.

Hewitt, J.N. "Time Delays in Gravitational Lenses."


Troxel, D.E. "CAFE Installation at Lincoln Laboratory."

MIT Workshop on Space Life Sciences and Virtual Reality, Dedham, Massachusetts, January 6, 1993.

Shinn-Cunningham, B.G. "Auditory Virtual Environments."

National Academy of Sciences Colloquium on Human-Machine Communication by Voice, Irvine, California, February 8, 1993.

Allen, J. "Linguistic Aspects of Speech Synthesis."


Lim, J.S. "HDTV Research at ATRP."


Elfadel, I.M. "Lyapunov Functions for Rotor Neural Networks."


Ehrlich, D.J. "Prototyping to Manufacturing: Applications of Lasers in Microelectronics and Micromechanics."

NATO Workshop on Nanolithography, Rome, Italy, April 6-8, 1993.

Smith, H.I., and M.L. Schattenburg. "X-ray Nanolithography: Limits, and Application to Sub-100 NM Manufacturing."


Holley, J. "Hydrogen Maser Polarimeter (invited talk)."


Oort Lecture, Leiden University, the Netherlands, May 1993.

Burke, B.F. "Searching for Planets of Other Stars."
Appendix A. RLE Publications

Optical Society of America Annual Meeting, Toronto, Canada, October 3-8, 1993.


Brothers, L.R., and N.C. Wong. "THz Optical Frequency Comb Generation."


Klepner, D. "Rydberg Atoms in Strong Magnetic Fields."

Lee, D., N.C. Wong, and L.R. Brothers. "Frequency Tuning and Stabilization of a KTP Optical Parametric Oscillator."

Pritchard, D.E. "Atom Interferometry."

Optical Society of America Topical Meeting on Soft X-ray Lithography, Third Annual, Monterey, California, May 10-12, 1993.


Ippen, E.P. "Ultrafast Dynamics in Active Semiconductor Waveguides."

Pecek Lecture, International Union for Aeronautics and Astronautics, Graz, Austria, October 1993.

Burke, B.F. "The Search for Exoplanets."


Hewitt, J.N. "Gravitational Lenses, Time Delays, and Hubble's Constant."


Han, H.C., and J.A. Kong. "Active Radiative Transfer Theory Under the Continuous Random Medium Approach."


Appendix A. RLE Publications


Yuan, Y., and J.A. Kong. “Radiation and Scattering by Antennas Attached to Coated Bodies of Revolution.”

Yuan, Y., Y.E. Yang, and J.A. Kong. “Integral Equation Analysis of Coupled Microstrip Lines with Ferromagnetic Thin Film.”


Hewitt, J.N. “Measuring the Size of the Universe with Gravitational Lenses (Harlow Shapley Lecturer).”


Mcllrath, M.B. Meeting Chair.


Berker, A.N. “Hard-Spin Mean-Field Theory.”


Summer School on Recent Developments in Statistical Physics, Istanbul, Turkey, July 26-August 6, 1993.

Berker, A.N. “Phase Transitions in Disordered Systems.”


Coppi, B. “Common and Uncommon Problems of Laboratory and Astrophysical Plasmas.”

SEMA TECH Company, Dallas, Texas, April 22, 1993.

Mcllrath, M.B. “Process Flow Representation for TCAD Frameworks.”


Appendix A. RLE Publications


Berker, A.N. "Hard-Spin Mean-Field Theory."


U.S. Naval Research Laboratory Colloquium, Annapolis, Maryland, April 22, 1993.

Hewitt, J.N. "Gravitational Lenses, Time Delays, and Hubble's Constant."

U.S. Office of Naval Research/Houston Advanced Research Center Workshop on Atomic Coherence, Crested Butte, Colorado, July 26-29, 1993

Chapman, M.S. "Experiments with a Separated Beam Interferometer (invited talk)."

U.S.-Japan Workshop on Ion Temperature Gradient Turbulence, University of Texas, Austin, Texas, January 11-14, 1993.

Migliuolo, S. "Ion Temperature Gradient Driven Impurity Modes."

University of Maryland Department of Astronomy Colloquium, College Park, Maryland, April 21, 1993.

Hewitt, J.N. "Gravitational Lenses, Time Delays, and Hubble's Constant."

Virtual Environment Technology for Training Advisory Committee Meeting, Orlando, Florida, August 11-13, 1993.

Zeltzer, D. "VETT Testbed Progress Report."


Fujimoto, J.G. "Femtosecond Carrier Dynamics in Semiconductors and Metals."

Workshop on Coherence and Uncertainty Relations, Baltimore, Maryland, August 10-13, 1993.


Sugiyama, L., M. Nassi, and B. Coppi. "Predicted Ignition Scenarios in Ignitor and the Role of ICRH."


A.1.2 Papers To Be Presented


Wong, N.C., and D. Lee. "Frequency Tuning and Phase Locking of an Ultrastable Doubly Resonant Optical Parametric Oscillator."


Elfadel, I.M. "On the Convergence of Relaxation Labelling Processes."

Fiber Communications Conference, San Jose, California, February 20-25, 1994.


International Conference on InP and Related Compounds, Santa Barbara, California, March 1994.


Optical Society of America Meeting, Dallas, Texas, October 2-7, 1994.

Shapiro, J.H. "An Eclectic Tour of Quantum Optical Communications."


A.1.3 Published Meeting Papers


Appendix A. RLE Publications
Appendix A. RLE Publications


Appendix A. RLE Publications


A.2 Journal Articles

A.2.1 Published Journal Articles


Bahi, S.R., B.R. Bennett, and J.A. del Alamo. "Doubly-Strained In$_{x}$Al$_{0.5}$As/n − In$_{0.5}$Ga$_{0.5}$As HFET with High Breakdown Voltage." IEEE Electron Device Lett. 14(1): 22-24 (1993).


Bennett, B.R., and J.A. del Alamo. "Optimal Epilayer Thickness for In$_{x}$Ga$_{0.5}$As and In$_{0.5}$Al$_{0.5}$As Composition Measurement by High-resolution X-ray Diffraction." J. Appl. Phys. 73(12): 8304-8308 (1993).


Appendix A. RLE Publications


Greenberg, J.E., I.M. Peterson, and P.M. Zurek. “Intelligibility-weighted Measures of Speech-to-


Appendix A. RLE Publications


Appendix A. RLE Publications


A.2.2 Journal Articles Accepted for Publication


Appendix A. RLE Publications


Grot, A.C., D. Psaltis, K.V. Shenoy, and C.G. Fonstad. "Integration of LEDs and GaAs Cir-
Appendix A. RLE Publications


Smet, J.H., C.G. Fonstad, and Q. Hu. "Magnetotunneling Spectroscopy in Wide In_{0.55}Ga_{0.45}As/In_{0.55}Al_{0.45}As Double Quantum Wells." Appl. Phys. Lett.


A.3 Books/Chapters in Books/Published Proceedings


Hector, S.D., and H.I. Smith. "Soft X-ray Projection Lithography Using Two Arrays of Phase Zone..."
Appendix A. RLE Publications


Appendix A. RLE Publications


A.4 RLE Publications

RLE technical reports (RLE TR) are available from the Massachusetts Institute of Technology, Document Services, Room 14-0551, Cambridge, Massachusetts 02139-4307; e-mail: docs@mit.edu; telephone number (617) 253-5668.

Other RLE publications may be obtained by contacting the RLE Publications Office, Building 36, Room 12, Cambridge, Massachusetts 02139-4301, e-mail: maryz@rlevm.mit.edu. Fax (617) 258-7864.


A.5 RLE Theses

Theses are available from MIT Document Services (see section A.4).


Appendix A. RLE Publications


Natarajan, V. Penning Trap Mass Spectroscopy at 0.1 ppb. Ph.D. diss., Dept. of Physics, MIT, 1993.

Appendix A. RLE Publications


A.6 Miscellaneous

PTP Reports are available from the office of Professor Bruno Coppi, Building 26-215. For other publications, contact MIT Document Services.


Appendix B. Current RLE Personnel

Director: Jonathan Allen
Associate Director: Daniel Kleppner

Professors

Jonathan Allen
Boris Altshuler
Dimitri A. Antoniadis
Arthur B. Baggeoer
George Bekfi
A. Nihat Berker
Abraham Bers
Robert J. Birgeneau
Amar G. Bose
Louis D. Braida
Bernard E. Burke
Sylvia T. Ceyer
Bruno Coppi
Shaoul Ezekiel
Clifton G. Fonstad, Jr.
Lawrence S. Frishkopf
Hermann A. Haus
Albert Hill
Erich P. Ippen
John D. Joannopoulos
Marc A. Kastner
Nelson Y.-S. Klang
John G. King
Daniel Kleppner
Jin A. Kong
Patrick A. Lee
Jerome Y. Lettvin
Jae S. Lim
Alan V. Oppenheim
William T. Peake
Miklos Porkolab

David E. Pritchard
William F. Schreiber
Campbell L. Searle
Jeffrey H. Shapiro
William M. Siebert
Henry I. Smith
Charles G. Sodini
David H. Staelin
Kenneth N. Stevens
Julius A. Stratton
Donald E. Troxel
Thomas F. Weiss
Jerome B. Wiesner
John L. Wyatt, Jr.
Henry J. Zimmermann

Associate Professors

Jesús A. del Alamo
Srinivas Devadas
James G. Fujimoto
Peter L. Haglestein
Leslie A. Kolodziejski
Simon G.J. Mochrie

Jacob K. White
Jonathan S. Wurtele

Assistant Professors

Raymond C. Ashoori
Jacqueline N. Hewitt
Qing Hu
Wolfgang Ketterle

Gregory W. Wornell

Senior Research Scientists

Nathaniel I. Durlach
Joseph S. Perkell

Robert H. Rediker

Principal Research Scientists

Donald K. Eddington
John J. Guinan, Jr.
William M. Rabinowitz
Charlotte M. Reed
Philip W. Rosenkranz
Kenneth Salisbury

Mandayam A. Srinivasan
David Zeltzer
Patrick M. Zurek

1 Professor Emeritus.
### Research Scientists and Research Specialists

<table>
<thead>
<tr>
<th>Giovanni Aliberti</th>
<th>Thomas J. Lohman</th>
<th>Dennis D. Rathman</th>
</tr>
</thead>
<tbody>
<tr>
<td>John W. Barrett</td>
<td>Joyce Manzella</td>
<td>John J. Rosowski</td>
</tr>
<tr>
<td>James M. Carter</td>
<td>Ivan Mastrosky</td>
<td>Selim M. Shahriar</td>
</tr>
<tr>
<td>Bertrand Delgutte</td>
<td>Melanie L. Matthies</td>
<td>Stefanie Shattuck-Hufnagel</td>
</tr>
<tr>
<td>Lorraine A. Delhome</td>
<td>Michael B. McLrath</td>
<td>Scott E. Silverman</td>
</tr>
<tr>
<td>Gregory T. Fischer</td>
<td>Robert D. Meade</td>
<td>Linda E. Sugiyama</td>
</tr>
<tr>
<td>David W. Foss</td>
<td>Stefano Migliuolo</td>
<td>Marco A. Svirsky</td>
</tr>
<tr>
<td>Dennis M. Freeman</td>
<td>Euclid E. Moon</td>
<td>Ngai Chuen Wong</td>
</tr>
<tr>
<td>David Gosalvez</td>
<td>Abhay K. Ram</td>
<td>Y. Eric Yang</td>
</tr>
<tr>
<td>Seth M. Hall</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Administrative Staff

<table>
<thead>
<tr>
<th>Joseph F. Connolly</th>
<th>John S. Peck</th>
<th>William H. Smith</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virginia R. Lauricella</td>
<td>Gerrard F. Power</td>
<td>Vicky-Lynn Taylor</td>
</tr>
<tr>
<td>Barbara J. Passero</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Support and Technical Staff

<table>
<thead>
<tr>
<th>Mary C. Aldridge</th>
<th>Donna L. Gale</th>
<th>Mark K. Mondol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Janice L. Balzer</td>
<td>Mary S. Greene</td>
<td>Donald K. North</td>
</tr>
<tr>
<td>Felicia G. Brady</td>
<td>Peter C. Guiod</td>
<td>Denise M. Rossetti</td>
</tr>
<tr>
<td>Manuel Cabral, Jr.</td>
<td>Maureen C. Howard</td>
<td>Bruce A. Russell</td>
</tr>
<tr>
<td>Susan E. Chafe</td>
<td>Cynthia Y. Kopf</td>
<td>Maxine P. Samuels</td>
</tr>
<tr>
<td>Donald A. Clements</td>
<td>Kit-Wah F. Lai</td>
<td>Sally C. Santiago</td>
</tr>
<tr>
<td>Anne E. Conklin</td>
<td>Cindy LeBlanc</td>
<td>Robert D. Sisson</td>
</tr>
<tr>
<td>John F. Cook</td>
<td>Catherine Lorusso</td>
<td>Mabayoje Tuyo</td>
</tr>
<tr>
<td>Carol A. Costa</td>
<td>Eleanor M. Luongo</td>
<td>Laura M. von Bosau</td>
</tr>
<tr>
<td>Ann K. Dix</td>
<td>Deborah S. Manning</td>
<td>Arlene E. Wint</td>
</tr>
<tr>
<td>Francis M. Doughty</td>
<td>Rita C. McKinnon</td>
<td>Mary J. Ziegler</td>
</tr>
<tr>
<td>Dorothy A. Fleischer</td>
<td>Gina L.B. Milton</td>
<td></td>
</tr>
</tbody>
</table>

### Postdoctoral Fellows and Associates

<table>
<thead>
<tr>
<th>Robert C. Armstrong</th>
<th>Joseph A. Izatt</th>
<th>Annie H. Takeuchi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tseikouras Athanassios</td>
<td>Jeng-Feng Lee</td>
<td>Alice E. Turk</td>
</tr>
<tr>
<td>Brett E. Bouma</td>
<td>Pauline M. McMahon</td>
<td>Simon Verghese</td>
</tr>
<tr>
<td>David L. Brock</td>
<td>Fred L. Palmer</td>
<td>Thomas E. Wiegand</td>
</tr>
<tr>
<td>Kung-Hau Ding</td>
<td>Gale S. Petrich</td>
<td>Reiner Wilhelms-Tricarico</td>
</tr>
<tr>
<td>Frank DiFilippo</td>
<td>David P. Pullman</td>
<td>Xinglong Yan</td>
</tr>
<tr>
<td>Ibrahim M. Elfadel</td>
<td>Sunil Puria</td>
<td>Yi Xu</td>
</tr>
<tr>
<td>Susan L. Goldman</td>
<td>Thomas Schapers</td>
<td>Zhe Zhang</td>
</tr>
</tbody>
</table>

### Research Assistants

<table>
<thead>
<tr>
<th>Daniel P. Aalberts</th>
<th>Mark A. Armstrong</th>
<th>Deborah J. Becker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charles C. Abnet</td>
<td>Neer R. Asherie</td>
<td>Paul A. Belk</td>
</tr>
<tr>
<td>David M. Abusch</td>
<td>William W. Au</td>
<td>Vishal L. Bhagwati</td>
</tr>
<tr>
<td>Mazher M. Aldina</td>
<td>Walter A. Avles</td>
<td>Igor P. Bilinsky</td>
</tr>
<tr>
<td>Michael R. Andrews</td>
<td>Ian M. Avruch</td>
<td>Luc Boivin</td>
</tr>
<tr>
<td>John G. Apostolopoulos</td>
<td>Ziad J. Azzam</td>
<td>Michael P. Bradley</td>
</tr>
</tbody>
</table>
Appendix B. Current RLE Personnel

Louis R. Brothers
Martin Burkhardt
Carlos R. Cabrera
Xuejun Cai
David J. Carter
Claudio L. Cesar
Hwa-Ping Chang
Michael S. Chapman
Grace H. Chen
Jerry C. Chen
Jyh-Shing Chen
Marilyn Y. Chen
Tak K. Cheng
Shiufun Cheung
Dmitri Chklovskii
Kyeongjae Cho
Woo-Young Choi
Mike T. Chou
Henry E. Chung
Samuel R. Conner
Christopher A. Coronado
Michael W. Courtney
Jay H. Damask
Kiran B. Dandekar
Charles Q. Davis
Kendall Davis
Joel C. DeVries
Sean M. Donovan
David J. Dougherty
Scott B. Dynes
John D. Ellithorpe
Hong T. Ewe
Alexis Falicov
Shanhui Fan
Juan Ferrera
Siegfried B. Fleischer
Andre B. Fletcher
Eric M. Foxlin
Joseph A. Frisbie
Alex W. Fung
Irene Fung
Boris Golubovic
James G. Goodberlet
Julie E. Greenberg
Andrew E. Grumet
Nitin Gupta
Rakesh Gupta
Andrew K. Halberstadt
Troy D. Hammond
Scott D. Hector
Michael R. Hee
Lori K. Herold
Eason Ho
Jeffrey R. Holley
Jody House
Chih-Chien Hsu
Hang Hu
Wen Hu
Charles T. Hultgren
Kyle K. Iwai
Brian R. Jacobson
Louise Jandura
Hong Jiao
Mark A. Johnson
Alkan Kabakcioglu
Charles A. Katz
Sumantha Kaushik
Farzana I. Khatri
Rizwan R. Koita
Jimmy Y. Kwon
Bobby Y. Lai
Warren M. Lam
Dicky Lee
Laurence H. Lee
Zachary K. Lee
Gadi Lenz
Huiying Li
Stan Y. Liao
Kan Lu
Jeffrey T. Ludwig
Robert I. Lutwak
Paul S. Martin
Noel S. Massey
Marc O. Mewes
Jose C. Monteiro
Akbar A. Moolji
Christopher B. Moore
John D. Moores
William P. Moyné
Martin Muendel
Phillip T. Nee
Julien J. Nicolas
Michael P. O'Connell
John H. Oates
Gabrielle M. Owen
Steven G. Patterson
Gregory E. Penn
Aleksandar Pfajfer
Joel R. Phillips
Michael O. Polley
Kelly L. Poort
Khaid Rahmat
Monte J. Ramstad
Caterina Riconda
William J. Rinderknecht
Yakov Royter
Richard A. Rubenstein
Timothy A. Savas
Michael J. Schwartz
Mark N. Seidel
Satyan Shah
Lisa Shatz
Amelia H. Shen
Krishna V. Shenoy
Shih-en Shih
Gennady Shvets
Luis M. Silveira
Matthew T. Sinn
Jurgen H. Smet
Neal W. Spellmeyer
Michelle S. Spina
Richard E. Stoner
Lon E. Sunshine
Chin H. Tan
Hong Z. Tan
Mohammad A. Tassoudji
Guillermo J. Tearney
Morrison Ulman
Dimitri L. Velikov
Murat E. Veysoglu
Pavel S. Volkov
Li-Fang Wang
Susie J. Wee
Daniel P. Weiker
Lorin F. Wilde
Vincent V. Wong
William S. Wong
Rolf A. Wyss
Anto Yasaka
Kenneth W. Yee
Chang D. Yoo
Kambiz C. Zangii
Yan Zhang
Craig B. Zilles

Teaching Assistants

David M. Baylon
Dirk M. Bernold
John R. Buck
Irfan U. Chaudhary
Peter A. Garbes

Christoforos N. Hadjicostis
Joel T. Johnson
Mattan Kamon
Malini V. Narayanan

Haralabos C. Papadopoulos
Li Shu
Andrew C. Singer
Kathleen E. Wage
## Appendix B. Current RLE Personnel

### Graduate Students

<table>
<thead>
<tr>
<th>Name</th>
<th>Name</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laura E. Adams</td>
<td>Eric J. Gaidos</td>
<td>James M. Ooi</td>
</tr>
<tr>
<td>Tracy E. Adams</td>
<td>Steven H. Isabelle</td>
<td>Janet L. Pan</td>
</tr>
<tr>
<td>David B. Askey</td>
<td>Erik B. Iverson</td>
<td>Denis J. Peregryn</td>
</tr>
<tr>
<td>Karl K. Berggren</td>
<td>Mohammad J. Khan</td>
<td>Adam D. Polcyn</td>
</tr>
<tr>
<td>Keren Bergman</td>
<td>Thomas C. Killian</td>
<td>Malini Ramaswamy</td>
</tr>
<tr>
<td>James C. Blastos</td>
<td>Chwen-yuen Ku</td>
<td>Brett S. Reid</td>
</tr>
<tr>
<td>Stephen A. Boppart</td>
<td>Volkang C. Kubali</td>
<td>James W. Reiner</td>
</tr>
<tr>
<td>Stanley W. Brown</td>
<td>Hong-Kwang J. Kuo</td>
<td>Todd H. Rider</td>
</tr>
<tr>
<td>Douglas S. Brungart</td>
<td>Daniel D. Lee</td>
<td>Ante Salcedo</td>
</tr>
<tr>
<td>Palmyra E. Catravas</td>
<td>Junehee Lee</td>
<td>Steven D. Schultz</td>
</tr>
<tr>
<td>Kai P. Chan</td>
<td>Kevin Li</td>
<td>Hugh E. Secker-Walker</td>
</tr>
<tr>
<td>Diana Dabby</td>
<td>Sharlene A. Liu</td>
<td>Barbara G. Shinn-Cunning</td>
</tr>
<tr>
<td>Ali Davish</td>
<td>Nicole Love</td>
<td>Konstantina Stankovic</td>
</tr>
<tr>
<td>William S. Daughton</td>
<td>Henry Y. Lu</td>
<td>Ryan A. Tagal</td>
</tr>
<tr>
<td>Steven J. Decker</td>
<td>Kent H. Lundberg</td>
<td>Kohichi R. Tamura</td>
</tr>
<tr>
<td>Joseph S. Desloge</td>
<td>John P. Mann</td>
<td>Christopher B. Umminger</td>
</tr>
<tr>
<td>Shepard S. Doelemann</td>
<td>Robert G. McDonald</td>
<td>Evren R. Unver</td>
</tr>
<tr>
<td>Christopher R. Doerr</td>
<td>Ignacio S. McQuirk</td>
<td>Luigi Vacca</td>
</tr>
<tr>
<td>James Ernsmeyer</td>
<td>Peter A. Monta</td>
<td>Shawn M. Verbout</td>
</tr>
<tr>
<td>Paul W. Fiegut</td>
<td>Ann W. Morgenthaler</td>
<td>Wan Y. Wan-Morshidi</td>
</tr>
<tr>
<td>Jun Funatsu</td>
<td>Aradhania Narula</td>
<td>Peter S. Yesley</td>
</tr>
<tr>
<td>Lynn D. Gabbay</td>
<td>Lynn E. Nelson</td>
<td>Michael J. Young</td>
</tr>
</tbody>
</table>

### Undergraduate Students

<table>
<thead>
<tr>
<th>Name</th>
<th>Name</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walter E. Babiec</td>
<td>David H. Hijirida</td>
<td>Michael T. Pierce</td>
</tr>
<tr>
<td>Latosha Byrd</td>
<td>Jason Hintersteiner</td>
<td>Nicholas Ploch</td>
</tr>
<tr>
<td>Janis D. Castanedo</td>
<td>Philip M. Hinz</td>
<td>Edmundo J. Rivera-Santos</td>
</tr>
<tr>
<td>Beth L. Chen</td>
<td>Elliot E. Hui</td>
<td>Szymon M. Rusinkiewicz</td>
</tr>
<tr>
<td>David K. Chow</td>
<td>Zulfiquar Hyder</td>
<td>Michelle M. Scheer</td>
</tr>
<tr>
<td>Erika Chuang</td>
<td>Megan C. Jasek</td>
<td>Mehul A. Shah</td>
</tr>
<tr>
<td>John A. Crouch</td>
<td>Kevin C. Knoedler</td>
<td>David L. Slisson</td>
</tr>
<tr>
<td>Gail Deneswich</td>
<td>Danielle G. Lemay</td>
<td>Bryan D. Smith</td>
</tr>
<tr>
<td>Laura C. Dilley</td>
<td>Kuo-Yi Lim</td>
<td>Bridget E. Tannian</td>
</tr>
<tr>
<td>Ilia Entin</td>
<td>Hsiaotung Liu</td>
<td>Stanley H. Thompson</td>
</tr>
<tr>
<td>Alexander N. Ernst</td>
<td>David S. Lum</td>
<td>Ann L. Torres</td>
</tr>
<tr>
<td>Fritz N. Francis</td>
<td>Beniym Menberu</td>
<td>Evan F. Wies</td>
</tr>
<tr>
<td>Ashanthi Gajaweera</td>
<td>Prescott F. Millett</td>
<td>Joshua N. Winn</td>
</tr>
<tr>
<td>Cinawaye Gammon</td>
<td>Jason T. Mueller</td>
<td>Jennifer A. Wozniak</td>
</tr>
<tr>
<td>Brett A. Geoffroy</td>
<td>Philip M. Nadeau</td>
<td>John J. Wuu</td>
</tr>
<tr>
<td>Daniel I. Goldman</td>
<td>Erik Nygren</td>
<td>Chiann J. Yeh</td>
</tr>
<tr>
<td>Shilpa M. Hattangadi</td>
<td>Sung S. Park</td>
<td>Xiao Yu</td>
</tr>
</tbody>
</table>
Appendix B. Current RLE Personnel

Visiting Scientists

Vladimir Barsukov
Gerald L. Beauregard
Corine A. Bickley
Lars H. Bomholt
Lyle Borg-Graham
Franco Carpignano
Paolo Detragiache
Theodore W. Ducas
Anna Esposito
Marc Fleury
Regula Fluck
Rogeve Gulati

Helen M. Hanson
Yie He
Arthur K. Jordan
Hae M. Jung
Jack Kotik
Cathy M. Lerner
Hamid Nawab
Suzanne Neil
Kevin O’Neill
Karen L. Payton
Jorgen L. Pind
Geoffrey L. Plant

Stanley J. Rosenthal
Hannes J. Schmiedmayer
Francesca Scire-Scappuzzo
Kaoru Sekiyama
Chulhun Seo
Devang M. Shah
Masataka Shirasaki
Jean-Claude Souyris
Chi-Kuang Sun
George M. Svolos
Stefan Wehinger

Research Affiliates

Mitchell R. Balonon-Rosen
John S. Barlow
Herbert J. Bernstein
Giuseppe Bertin
Suzanne Boyce
Frank S. Cardarelli
Brian J. Clifton
H. Steven Colburn
Paul Duchnowski
Daniel Ehrlich
Carol Y. Espy-Wilson
Paul D. Fiore
Ignacio Garcia-Otero
Gad Geiger
Bernard Gold
Richard S. Goldhor
Kenneth W. Grant
Robert D. Hall
Katherine L. Hall
Philip R. Hemmer

Robert E. Hillman
Mark A. Hollis
Eva B. Holmberg
Caroline B. Huang
Yoshiko Ito
Joseph A. Jarrell
Lance G. Joneckis
Carlos Kennedy
John D. Kierstead
Harlan Lane
Charles P. Lin
Neil A. Macmillan
John I. Makhoul
Sharon Y. Manuel
Bruce R. Musicus
Ronel Nel
Mark R. Nilsen
Leonard L. Picard
Matthew H. Power
James C. Preisig

Mara G. Prentiss
Christine M. Rankovic
Michael E. Ravicz
Stephen A. Raymond
Joseph F. Rizzo
Diane E. Ronan
Jay T. Rubinstein
Steven R. Rueman
Robert T. Shin
Richard J. Solomon
Frank J. Stefanov-Wagner
David A. Steffens
Joseph Tierney
Michael J. Tsuk
Kenneth P. Wacks
Ehud Weinstein
David R. Williams
Jane W. Wozniak
Jiqing Xia
Marc Zissman
Appendix C. Milestones

C.1 New Faculty and Staff

Dr. Wolfgang Ketterle was appointed Assistant Professor of Physics on September 1, 1993. Professor Ketterle is a principal investigator of RLE's Atomic, Molecular, and Optical Physics Group.

Dr. M. Selim Shahriar (SB '86, SM '89, PhD '91) was appointed as a research scientist in RLE's Atomic, Molecular, and Optical Physics Group, effective September 1, 1993.

William H. Smith was appointed RLE Fiscal Officer effective July 1, 1993.

Mary E. Young was appointed RLE Assistant Fiscal Officer effective April 18, 1994.

C.2 Retirements

Donald F. Duffy, retired as RLE Fiscal Officer on June 30, 1993, after 31 years of service at MIT.

Campbell L. Searle, Professor of Electrical Engineering and Computer Science, retired after serving forty-five years at MIT.

C.3 Chair Appointments

Dr. Gregory W. Wornell (SM '87, PhD '91), Assistant Professor of Electrical Engineering and Computer Science, was appointed to the ITT Career Development Assistant Professorship, effective July 1, 1993. Professor Wornell is a principal investigator in RLE's Digital Signal Processing Group.

C.4 Awards and Honors

Dr. Boris L. Altshuler, Professor of Physics, was announced as corecipient of the 1993 Hewlett-Packard Europhysics Prize for Outstanding Achievement in Solid-State Physics by the European Physical Society. Professor Altshuler is a principal investigator in RLE's Surfaces and Interfaces Group.

Dr. Arthur B. Baggeroer (SM/EE '65, ScD '68), Ford Professor of Engineering and Professor of Electrical and Ocean Engineering, received the Ocean Engineering Society's 1992 Distinguished Technical Achievement Award. Professor Baggeroer is a principal investigator in RLE's Digital Signal Processing Group.

Dr. Bruno Coppi, Professor of Physics, received the Special Prize for Scientific Research of the Prime Minister of Italy. In addition, Professor Coppi, a principal investigator in RLE's Plasma Physics Group, was appointed to a three-year term as a senior fellow of the Italian Academy for Advanced Studies in America and also received the 1993 Italgas Prize for Energy Science.

Dr. Jesús A. del Alamo, Associate Professor of Electrical Engineering and Computer Science, was awarded MIT's Harold E. Edgerton Award. In addition to his research, Professor del Alamo, a principal investigator in RLE's Materials and Fabrication Group, was recognized for his excellent teaching skills and service to the MIT community.

Dr. James G. Fujimoto (SB '79, SM '81, PhD '84), Associate Professor of Electrical Engineering and Computer Science, was elected to fellow in the Optical Society of America. Professor Fujimoto is a principal investigator in RLE's Optics and Devices Group.

Dr. Jacqueline N. Hewitt (PhD '86), Assistant Professor of Physics and Class of 1948 Career Development Professor, was awarded the Henry G. Booker Prize from the International Union of Radio Science (URSI). Professor Hewitt is a principal investigator in RLE's Radio Astronomy Group.

Dr. Marc A. Kastner, Donner Professor of Science, was appointed director of MIT's Center for Materials Science and Engineering, effective July 1, 1993. Professor Kastner is a principal investigator in RLE's Quantum-Effect Devices Group.

Dr. Leslie A. Kolodziejski, Karl R. Van Tassel Career Development Associate Professor, was selected to serve a three-year term as Esther and Harold E. Edgerton Associate Professor, effective September 1, 1993. Professor Kolodziejski is a principal investigator in RLE's Materials and Fabrication Group.

Campbell L. Searle, Professor of Electrical Engineering and Computer Science, and Dr. William M. Siebert (SB '46, ScD '52), Ford Professor of Engineering in the Department of Electrical Engineering and Computer Science, were named corecipients of the Department Head's Special Recognition Award. Professor Siebert is a member of RLE's Auditory Physiology Group.
Appendix C. Milestones

**Dr. David H. Staelin** (BS '60, MS '61, ScD '65), Assistant Director of Lincoln Laboratory and Cecil H. Green Professor of Electrical Engineering, was awarded the IEEE Signal Processing Society's 1992 Senior Award in the image processing and multidimensional signal processing area. Professor Staelin is a principal investigator in RLE's Radio Astronomy Group.

**Dr. Jerome B. Wiesner**, Institute Professor, received the National Academy of Sciences' highest honor, its Public Welfare Medal, in 1993 for his devoted and successful efforts in science policy, education, nuclear disarmament, and world peace. Over the years, Dr. Wiesner's record of service has included tenures as RLE director, MIT president, and presidential science advisor to Presidents Kennedy and Johnson.
Appendix D. RLE Research Support Index

Advanced Research Projects Agency
(ARPA) 7-21, 26-27, 35-45, 64-70, 72-75, 79, 85-86, 149-153, 269-271, 300-301, 303, 305-308, 318-324
Advanced Telecommunications Research Program 329-334
AT&T Bell Laboratories 26-27, 319-320
Bose Corporation 318
California Institute of Technology/Jet Propulsion Laboratory 263-265
Consortium for Superconducting Electronics 269-271, 303
DEMACO 266
Digital Equipment Corporation 265-266, 303-305
Draper (Charles S.) Laboratory 29-34, 100-103, 141-143, 204-207
Electric Power Research Institute 133-137
Federal Bureau of Investigation 302
Fujitsu Laboratories 101-103
Hertz Foundation 13-14
Houston Advanced Research Center 345-349
IBM Corporation 72-73, 289-292, 300-305
Joint Services Electronics Program (JSEP) 7-11, 29-41, 51-73, 76-95, 99-101, 104-118, 161-190, 215-228, 263-266
Klatt (Dennis) Memorial Fund 355-364
LeBel (C.J.) Foundation 355-364
Lockheed Sanders, Inc. 318-320, 323
Magnetic Fusion Science Fellowship Program 244-246
Maryland Procurement Office 139-140, 323-324
MIT Leaders for Manufacturing Program 284
MIT Lincoln Laboratory 101, 124-131, 141-143, 148-149, 153-154, 282
MIT Research Laboratory of Electronics 145-147
MIT-Woods Hole Oceanographic Institution Joint Program 318
National Aeronautics and Space Administration 86-88, 153-154, 237-240, 263-265, 279-284
National Center for Integrated Photonic Technology 7-21, 26-27, 35-46, 48, 103-105
National Defense Science and Engineering 303
National Institute of Standards and Technology 141-143
National Institutes of Health 118-124, 374-382, 386-388, 397-408
Packard (David and Lucille) Foundation 279-281, 322-323
Scitex America Corporation 337-339
Semiconductor Research Corporation 61-63, 304
Sloan (Alfred P.) Fellowship 279
Texas Instruments, Inc. 29-34
U.S. Air Force - Electronic Systems Division 195-203
U.S. Army - Cold Regions Research and Engineering Laboratory 266
U.S. Army - Harry Diamond Laboratories 231-237
U.S. Army Research Office 21-23, 25-26, 62-64, 79-85, 141-143, 154-157, 221-225
U.S. Department of Energy 231-240, 244-260
U.S. Department of Transportation 266
U.S. Navy 301-304, 309-313
U.S. Navy - Naval Air Systems Command 64-70, 79, 85-86
U.S. Navy - Naval Training Systems Center 388-389
Project Staff and Subject Index
Project Staff and Subject Index

A

Aalberts, Daniel P.  161, 164
Abberathy, Douglas L.  187
Abnet, C. Cameron  397, 399
Abusch, David M.  53
Acioli, Lucio H.  99, 109
Acoustic Thermometry of Ocean Climate (ATOC) Program  325
Adams, Laura E.  99, 109
Advanced Microwave Sounding Unit  282
Advanced X-ray Astrophysics Facility  86
Aggarwal, Rajni J.  7, 21
Ahadian, Joseph F.  7, 15, 35, 48
Alcator C-MOD  249
Alldridge, Mary C.  99
Aliberti, Giovanni  317
Alidina, Mazhar M.  289, 305, 307
Allen, Jonathan  289—292, 300, 355
Altshuler, Boris  213
Altshuler, Boris L.  51—52, 455
Andreev, Anton V.  51
Andrews, Michael R.  209, 225
Annaswamy, Anuradha M.  367, 382
Antoniadis, Dimitri A.  72, 76, 79, 289, 301
Apostolopoulos, John G.  329, 330
Arecibo Observatory  277
Arias, Tomas A.  181
Ariel, Imadiel  51, 57, 161, 181
Armstrong, Robert C.  289
Array processing  325
Asherie, Neer R.  231, 249
Ashoori, Raymond C.  91—95
Atkins, Robert G.  263, 265
Atomic physics  209—228
Atoms
  Diffraction of  221
  Structure in magnetic fields  209
  Trapping and cooling  225
Au, William W.  263
Aucoin, Richard J.  61, 69, 86
Auditory physiology  397—408
Aviles, Walter A.  367, 388
Avruch, Ian M.  277
Azzam, Ziad J.  99, 134

B

Babiec, Walter E.  367, 380

Baggeroer, Arthur B.  317, 325, 326, 455
Bahl, Sandeep R.  29
Baltus, Donald G.  289, 290
Balzer, Janice L.  397
Barrett, John W.  277, 283
Barsukov, Vladimir M.  355
Basu, Santanu  99, 124, 125
Beasley, Julian B.  367, 388
Beauregard, Gerald L.  367, 375, 389
Becker, Deborah J.  277
Bekefi, George  231—237
Belk, Paul A.  53
Bergman, Keren  99, 101
Berisset Philippe  263, 266
Berker, A. Nihat  161—167, 188
Berman, David  91, 263, 270, 271
Bers, Abraham  231, 237
Bhagwati, Vishal L.  289, 305, 306
Bickley, Corine A.  355
Bilinsky, Igor P.  99, 113
Birgeneau, Robert J.  169—173
Blastos, James C.  231
Bock, Robert D.  263, 269
Boivin, Luc  99, 101
Bolt, Beranek, and Newman  388
Bomholt, Lars H.  263, 265
Boppart, Stephen A.  99, 119
Boston University  361
Bouna, Brett E.  99, 109, 119
Bounds, Jeffrey K.  99, 101
Bowring, Kristine M.  51
Boyce, Suzanne E.  355
Bradley, Michael P.  209, 218
Brady, Felicia G.  231
Braida, Louis D.  367—374
Brandeis University Graybiel Laboratory  388
Braun, Eric K.  7, 14, 19, 20
Brock, David L.  367, 389
Brodsky, Mikhail G.  91
Broner, Fernando A.  99
Brothers, L. Reginald  139, 141
Brown, Elliot  148
Brown, Stanley W.  317, 318
Buck, John R.  317, 318
Burke, Bernard F.  277—278
Burkhardt, Martin  61, 65, 76

C

Cabrera, Carlos R.  277, 283
### Project Staff and Subject Index

**CAD Framework Initiative** 292
Cai, Xuejun 289, 302
Canizaras, Claude R. 86
Capaz, Rodrigo 181
Cariani, Peter A. 400
Carpignano, Franco 231, 249
Carter, David J. 61, 76, 79
Carter, James M. 61, 62, 64, 65, 70, 72, 73, 79, 83, 86
Catipovic, Josko 318
Catravas, Palmyra E. 231
Ceyer, Sylvia T. 7, 175—179
Chafe, Susan E. 289
Chalcopryite 161
Chan, Ho-Bun 91
Chan, Kai P. 231, 237
Chang, Hwa-Ping 355
Channel dropping filter 103
Chaos 318, 319, 321, 322
Chapman, Michael S. 209, 221
Chen, Beth L. 231
Chen, Grace H. 277, 279
Chen, Jerry C. 9, 99
Chen, Jyh-Shing 367, 383
Chen, Marilyn R. 355
Chenausky, Karen 7
Cheng, Howard 355
Cheng, Tak K. 99, 106
Cheung, Shiufun 329, 331
Chinn, S. 103
Chklovskii, Dmitri 57
Cho, Kyeongjae 181
Choi, Woo-Young 7, 8, 9, 93
Chomsky, Noam A. 411—413
Chou, Michael 289, 304
Chou, Mike T. 61
Chow, Carson C. 231, 237
Chow, David K. 145
Christian, Kevin G. 277, 284
Chu, Arthur 195
Chu, William 72, 79
Chung, Henry E. 397, 399
Circuit design 309
Clarke, Leanne 355
Cochlear efferents 462, 403
Cochlear implants 359, 374, 404
Coffman, Bridget L. 367
Collins, Christopher Thad 412
Columbia University 249
Computer-integrated design 309
Conklin, Anne E. 277
Conner, Samuel R. 277
Cooper, Catherine 175
Coppi, Bruno 231, 249—260, 455
Cornell University. National Nanofabrication Facility 223
Coronado, Christopher A. 35, 37, 42
Costa, Carol A. 209
Courtney, Michael W. 209
Coutu, Pierre 263, 265
Cudjoe-Flanders, Charmaine A. 7, 29, 35, 263
Cuomo, Kevin M. 317, 318
Curd, Derek R. 263
Custom integrated circuits
Alignment and consistency 289
Computer-aided design 289—308
Device simulation 300
Embedded control systems 305
Fault tolerance 305
Microelectromechanical computer-aided design 302
Reliability 305
VLSI computer-aided design 289—293
Cyclotron resonance 218

**D**

Daimler-Benz Research Center 29
Daley, Sean P. 175
Damask, Jay N. 35, 45, 61, 79, 99, 103
Dandekar, Kiran B. 367, 380, 384
Darwish, Ali M. 99, 104
Davis, C. Quentin 397, 399
Davis, Kendall B. 209, 225
Decker, Steven J. 289, 292, 293, 296
del Alamo, Jesús A. 29, 79, 85, 145, 455
Delgutte, Bertrand 397, 400, 401
Delhomme, Lorraine A. 367, 374, 375
Delin, K.A. 270, 271
Denesvich, Gail 367, 375
Desloge, Joseph G. 367, 378
Detragiache, Paolo 231
Devadas, Srinivas 289, 305—308
DeVries, Joel C. 209, 215
DiFilippo, Frank 209, 218
Digital signal processing 103, 317—327, 329—335
Fractals 320
Dilley, Laura C. 355
Ding, Kung Hau 263
Dix, Ann K. 367
DNA decoding 345
DNA sequencing 345
Doan, Jonathan C. 231
Doeleman, Sheperd S. 277
Doerr, Christopher R. 99, 100, 101
### Project Staff and Subject Index

<table>
<thead>
<tr>
<th>Name</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donnelly, J.</td>
<td>103</td>
</tr>
<tr>
<td>Donoghue, John J.</td>
<td>195</td>
</tr>
<tr>
<td>Donovan, Sean M.</td>
<td>35, 48, 61, 88</td>
</tr>
<tr>
<td>Dougherty, David J.</td>
<td>99, 101, 104, 105</td>
</tr>
<tr>
<td>Doughty, Francis M.</td>
<td>309</td>
</tr>
<tr>
<td>Doughty, Laura B.</td>
<td>231</td>
</tr>
<tr>
<td>Dron, Lisa G.</td>
<td>269, 292, 295</td>
</tr>
<tr>
<td>Ducas, Theodore W.</td>
<td>209, 215</td>
</tr>
<tr>
<td>Duchnowski, Paul</td>
<td>367</td>
</tr>
<tr>
<td>Duffy, Donald F.</td>
<td>455</td>
</tr>
<tr>
<td>Durlach, Nathaniel I.</td>
<td>367, 375, 378, 379</td>
</tr>
<tr>
<td>Dutra, Emanuel G.</td>
<td>263, 266</td>
</tr>
<tr>
<td>Dwarf stars</td>
<td>280</td>
</tr>
<tr>
<td>Dynes, Scott B.C.</td>
<td>397, 401</td>
</tr>
<tr>
<td>Dykas, Theodore W.</td>
<td>209, 215</td>
</tr>
<tr>
<td>Duchnowski, Paul</td>
<td>367</td>
</tr>
<tr>
<td>Duffy, Donald F.</td>
<td>455</td>
</tr>
<tr>
<td>Durlach, Nathaniel I.</td>
<td>367, 375, 378, 379</td>
</tr>
<tr>
<td>Dutra, Emanuel G.</td>
<td>263, 266</td>
</tr>
<tr>
<td>Dwarf stars</td>
<td>280</td>
</tr>
<tr>
<td>Dynes, Scott B.C.</td>
<td>397, 401</td>
</tr>
<tr>
<td>Ducas, Theodore W.</td>
<td>209, 215</td>
</tr>
<tr>
<td>Duchnowski, Paul</td>
<td>367</td>
</tr>
<tr>
<td>Duffy, Donald F.</td>
<td>455</td>
</tr>
<tr>
<td>Durlach, Nathaniel I.</td>
<td>367, 375, 378, 379</td>
</tr>
<tr>
<td>Dutra, Emanuel G.</td>
<td>263, 266</td>
</tr>
<tr>
<td>Dwarf stars</td>
<td>280</td>
</tr>
<tr>
<td>Dynes, Scott B.C.</td>
<td>397, 401</td>
</tr>
</tbody>
</table>

### F

- Failcov, Alexis 161, 164
- Ferrera, Juan 61, 62, 65, 79, 83, 103
- Fleisch, Paul W. 277, 283, 284
- Field-effect transistors
  - InGaAs 29
  - InP 29
- Fischer, Gregory T. 309
- Fisher, Philip A. 35, 37, 42
- Fleischer, Dorothy A. 289
- Fleischer, Siegfried B. 99
- Fleming, Robert C. 61, 86
- Fletcher, André B. 277
- Fleury, Marc 99, 124, 130
- Fonstard, Clifton G., Jr. 7—28, 48, 83, 154
- Ford, Darlene J. 99
- Foxlin, Eric M. 367, 379
- Free-electron lasers 231
- Freeman, Dennis M. 397, 399
- Frisbie, Joseph A. 367
- Frishkopf, Lawrence S. 397
- Frustration 161
- Fuchs, Vladimir 231, 237
- Fujimoto, James G. 7, 99, 109—124, 455
- Fung, Irene 139, 140
- Fusion 249

### G

- Gabbay, Lynn D. 309
- Gaffron, Svetlana 367, 390
- Gajaweera, Ashanthi 367, 375
- Gale, Donna L. 99
- Gavrillovich, Paul 7, 8
- Gealow, Jeffrey 294
- Genetic analysis 345
- Geng, Helena 355
- Genosensor technology 345
- Geoffroy, Brett A. 355
- Geophysics 325
- Geraci, James R. 35, 37
- Gold, Bernard 317, 323
- Goldhor, Richard S. 355
- Goldman, Susan L. 367
- Golubovic, Boris 139
<table>
<thead>
<tr>
<th>Name</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodberlet, James G.</td>
<td>99, 124, 125, 129, 130</td>
</tr>
<tr>
<td>Goodhue, W.</td>
<td>13</td>
</tr>
<tr>
<td>Gosalvez, David B.</td>
<td>175, 177, 179</td>
</tr>
<tr>
<td>Gouveia-Neto, Artur D.</td>
<td>99, 109</td>
</tr>
<tr>
<td>Grant, Kenneth W.</td>
<td>367</td>
</tr>
<tr>
<td>Gravitational lenses</td>
<td>277, 279</td>
</tr>
<tr>
<td>Greenberg, Julie E.</td>
<td>367, 378</td>
</tr>
<tr>
<td>Greytak, Thomas J.</td>
<td>226</td>
</tr>
<tr>
<td>Griffith, Mark R.</td>
<td>277</td>
</tr>
<tr>
<td>Grimm, W. Eric L.</td>
<td>140</td>
</tr>
<tr>
<td>Grot, A.</td>
<td>15</td>
</tr>
<tr>
<td>Gu, Qizheng</td>
<td>263, 266</td>
</tr>
<tr>
<td>Guinan, John J., Jr.</td>
<td>397, 401, 402</td>
</tr>
<tr>
<td>Guizod, Peter C.</td>
<td>355</td>
</tr>
<tr>
<td>Gulati, R.</td>
<td>367, 383</td>
</tr>
<tr>
<td>Gupta, Nitin</td>
<td>61, 64, 73</td>
</tr>
<tr>
<td>Gupta, Rajesh K.</td>
<td>145, 149</td>
</tr>
<tr>
<td>Gupta, Rakesh</td>
<td>367, 388</td>
</tr>
<tr>
<td>Gyroscopes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Laser 195</td>
</tr>
<tr>
<td>H</td>
<td></td>
</tr>
<tr>
<td>Hadjicostis, Christoforos N.</td>
<td>317, 319</td>
</tr>
<tr>
<td>Hagelestein, Peter L.</td>
<td>99, 124—137</td>
</tr>
<tr>
<td>Hajjahmad, Ibrahim A.</td>
<td>329, 331</td>
</tr>
<tr>
<td>Hakkarainen, J. Mikko</td>
<td>289, 292, 293</td>
</tr>
<tr>
<td>Hall, Katherine L.</td>
<td>99, 104</td>
</tr>
<tr>
<td>Hall, Seth M.</td>
<td>355, 367</td>
</tr>
<tr>
<td>Halle, Morris</td>
<td>355, 411—413</td>
</tr>
<tr>
<td>Halverson, Michelle</td>
<td>355</td>
</tr>
<tr>
<td>Hammond, Troy D.</td>
<td>209, 221</td>
</tr>
<tr>
<td>Hands</td>
<td>367—389</td>
</tr>
<tr>
<td>Hansell, G.</td>
<td>14</td>
</tr>
<tr>
<td>Hanson, Helen M.</td>
<td>355</td>
</tr>
<tr>
<td>Hardy, John</td>
<td>355</td>
</tr>
<tr>
<td>Harker, Marnie</td>
<td>355</td>
</tr>
<tr>
<td>Harring, Debra L.</td>
<td>169, 329</td>
</tr>
<tr>
<td>Harvard University</td>
<td>24, 277, 384</td>
</tr>
<tr>
<td>Haus, Hermann A.</td>
<td>7, 9, 11, 45, 79, 99—104, 105</td>
</tr>
<tr>
<td>Hay, Todd A.</td>
<td>231</td>
</tr>
<tr>
<td>Heard Island Feasibility Test</td>
<td>325</td>
</tr>
<tr>
<td>Hearing</td>
<td>367—380, 397</td>
</tr>
<tr>
<td>Hearing aids</td>
<td>367—379</td>
</tr>
<tr>
<td>Cochlear prostheses</td>
<td>359, 374, 399</td>
</tr>
<tr>
<td>Tactile aids</td>
<td></td>
</tr>
<tr>
<td>Hearing-impaired individuals</td>
<td>355, 359, 360, 367—379, 397, 408</td>
</tr>
<tr>
<td>Hector, Scott D.</td>
<td>61, 64, 65, 69</td>
</tr>
<tr>
<td>Hee, Michael R.</td>
<td>99, 119</td>
</tr>
<tr>
<td>Heemeyer, Sven</td>
<td>91</td>
</tr>
<tr>
<td>Heflin, Michael B.</td>
<td>277</td>
</tr>
<tr>
<td>Held, Richard M.</td>
<td>367, 379, 388</td>
</tr>
<tr>
<td>Hemmer, Philip R.</td>
<td>195</td>
</tr>
<tr>
<td>Herold, Lori K.</td>
<td>277</td>
</tr>
<tr>
<td>Hermann, Frederick P.</td>
<td>289, 292, 294</td>
</tr>
<tr>
<td>Heteroepitaxy</td>
<td>35, 46, 48</td>
</tr>
<tr>
<td>Heterostructures</td>
<td>29, 46</td>
</tr>
<tr>
<td>Hewitt, Jacqueline N.</td>
<td>277, 279—282, 455</td>
</tr>
<tr>
<td>High-definition television</td>
<td>329—334, 337—339</td>
</tr>
<tr>
<td>Hijirida, David H.</td>
<td>231, 249</td>
</tr>
<tr>
<td>Hillman, Robert E.</td>
<td>355, 360</td>
</tr>
<tr>
<td>Hinz, Philip M.</td>
<td>209, 225</td>
</tr>
<tr>
<td>Hirayama, Yuzo</td>
<td>7, 22, 99</td>
</tr>
<tr>
<td>Ho, Easen</td>
<td>35, 37, 42</td>
</tr>
<tr>
<td>Holley, Jeffrey R.</td>
<td>209, 215</td>
</tr>
<tr>
<td>Hollis, Mark A.</td>
<td>345, 348</td>
</tr>
<tr>
<td>Holmberg, Eva B.</td>
<td>355</td>
</tr>
<tr>
<td>Horn, Bethold K.P.</td>
<td>289, 292, 293, 294, 295, 296, 297</td>
</tr>
<tr>
<td>Horowitz, David M.</td>
<td>355</td>
</tr>
<tr>
<td>Horowitz, P.</td>
<td>277</td>
</tr>
<tr>
<td>Hoshino, Isako</td>
<td>7, 26</td>
</tr>
<tr>
<td>Hoston, William C.</td>
<td>161, 165</td>
</tr>
<tr>
<td>House, Jody L.</td>
<td>35, 37, 42, 46</td>
</tr>
<tr>
<td>Howe, Andrew L.</td>
<td>367</td>
</tr>
<tr>
<td>Howe, Robert D.</td>
<td>384</td>
</tr>
<tr>
<td>Hsu, Chih-Chien</td>
<td>263, 265, 266</td>
</tr>
<tr>
<td>Hsu, Lawrence</td>
<td>294</td>
</tr>
<tr>
<td>Hu, Hang</td>
<td>61, 72</td>
</tr>
<tr>
<td>Hu, Qing</td>
<td>25, 145—158</td>
</tr>
<tr>
<td>Hu, Wen</td>
<td>231</td>
</tr>
<tr>
<td>Huang, Caroline</td>
<td>355</td>
</tr>
<tr>
<td>Huang, David</td>
<td>99, 109, 119</td>
</tr>
<tr>
<td>Huang, Gregory T.</td>
<td>263, 265</td>
</tr>
<tr>
<td>Hugunin, James J.</td>
<td>61, 79</td>
</tr>
<tr>
<td>Hui, Elliot E.</td>
<td>145</td>
</tr>
<tr>
<td>Hultgren, Charles T.</td>
<td>99, 104</td>
</tr>
<tr>
<td>Human-machine interfaces</td>
<td>367, 379, 391</td>
</tr>
<tr>
<td>I</td>
<td></td>
</tr>
<tr>
<td>IBM Corporation</td>
<td>61, 169</td>
</tr>
<tr>
<td>Ignitor-Ult experiment</td>
<td>249</td>
</tr>
<tr>
<td>Image Processing</td>
<td>317—327</td>
</tr>
<tr>
<td>Induced stochasticity and chaos</td>
<td>237</td>
</tr>
<tr>
<td>Integrated circuits</td>
<td>61</td>
</tr>
<tr>
<td>Alignment and consistency</td>
<td>289</td>
</tr>
<tr>
<td>CMOS circuits</td>
<td>291, 301</td>
</tr>
<tr>
<td>Computer-aided design</td>
<td>289—308</td>
</tr>
<tr>
<td>Computer-aided fabrication</td>
<td>309</td>
</tr>
<tr>
<td>VLSI</td>
<td>7, 13, 14, 48</td>
</tr>
<tr>
<td>VLSI computer-aided design</td>
<td>289—293</td>
</tr>
</tbody>
</table>
Intelligent vehicle 292
Interferometry
  Atom wave 221
Ions 218
Ippen, Erich P. 7, 11, 99, 100, 101, 104—106
Isabelle, Steven H. 317, 319
Ismail, Khalid 88
Iwai, Kyle K. 329, 332
Izatt, Joseph A. 99, 109, 119

J
Jacobson, Brian R. 145, 153
Jandura, Louise 367, 384
Jasek, Megan C. 367, 391
Jiao, Hong 209
Joannopoulos, John D. 165, 181—186, 188
Joffe, Michael A. 209, 225
Johnson, Joel T. 263, 265
Johnson, Mark A. 355
Joint European Tokamak 249
Johns, Douglas Arnold 411
Johns, R.V. 24
Jordan, Arthur 263
Jülich GmbH Company 145
Jung, Hae-Mook 329

K
Kabakçıoğlu, Alkan 161, 164, 165
Kamon, Mattan 289, 302, 303
Karason, Steingrimur P. 367, 382
Kärtner, Franz X. 99, 101
Kastner, Marc A. 53—56, 455
Katz, Charles A. 277, 279, 280
Katz, Daniel 195
Kaushik, Sumanth 99, 124, 125, 131, 134
Keagy, Michael T. 367
Keast, Craig L. 289, 292, 295
Ketterle, Wolfgang 209, 225, 455
Keyser, Samuel J. 355
Khan, Mohammed J. 99, 103
Khart, Farzana I. 99, 101
Kiang, Nelson Y.S. 397
Kierstead, John D. 195
Kim, Songmin 289, 303
King, Barbara A. 139, 145
Klein, Olivier 53
Kleppner, Daniel 52, 209—218

Knight, Thomas F., Jr. 296
Knoedler, Kevin C. 367, 384
Kochanek, Christopher 277
Kolt, Rizwan R. 277, 284
Kolodziejski, Leslie A. 15, 35—48, 79, 88, 103,
  455
Kong, Jin Au 263—269
Kopf, Cynthia Y. 99
Korsmeyer, F. Thomas 304
Koshida, Nubuyoshi 61
Kotok, Irene 367, 375
Kumar, Arvind 61, 76
Kuo, Hong-Kwang 355
Kwon, Jimmy Y. 309

L
Lai, Bobby Y. 139, 141
Lai, Kit-Wah F. 263
Lam, Cheung-Wei 263
Lam, Warren M. 317, 320
LaMotte, Robert H. 382, 384
Lane, Harlan 355, 359
Lasers 7—28, 48, 99—137
  Additive pulse modelocking 99, 100
  Femtosecond 109
  Fiber optic 195—207
  Free-electron 231
  Low-power 124
  Medical 118
  Ridge waveguide distributed feedback 8, 9
  Semiconductor 11
  Solid-state 109
  Solid-state far-infrared 154
  Ultrashort 109, 118
Lawrence Livermore National Laboratory 249
LeBlanc, Cindy 329
Lee, Dicky 139, 141
Lee, Hae-Seung 289, 292, 293, 294, 296, 297
Lee, Jeng-Feng 367, 388, 391
Lee, Laurence H. 263, 270
Lee, Patrick A. 57—59
Lee, William K. 263, 266
Leeb, Steven B. 289, 304
LeMay, Danielle G. 367
Lenz, Gadi 11, 99, 104
Lew, Julie C. 61, 86
Li, Huiying 61, 65
Li, Kevin 263, 265
Li, Peggy 355
Liao, Stan Y. 289, 305, 306
Lim, Jae S. 329—335
<table>
<thead>
<tr>
<th>Name</th>
<th>Page Numbers</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lim, Kuo-Yi</td>
<td>35, 37, 42, 46</td>
<td></td>
</tr>
<tr>
<td>Lin, Charles P.</td>
<td>99, 119</td>
<td></td>
</tr>
<tr>
<td>Linden, Derek S.</td>
<td>263</td>
<td></td>
</tr>
<tr>
<td>Linguistics</td>
<td>411–413</td>
<td></td>
</tr>
<tr>
<td>Liu, Shariileen A.</td>
<td>355</td>
<td></td>
</tr>
<tr>
<td>Locke, John</td>
<td>355</td>
<td></td>
</tr>
<tr>
<td>Lohman, Thomas J.</td>
<td>309</td>
<td></td>
</tr>
<tr>
<td>Lorusso, Catherine</td>
<td>231</td>
<td></td>
</tr>
<tr>
<td>Love, Nicole S.</td>
<td>309</td>
<td></td>
</tr>
<tr>
<td>Lu, Henry Y.</td>
<td>355</td>
<td></td>
</tr>
<tr>
<td>Lu, Kan</td>
<td>35, 42, 48</td>
<td></td>
</tr>
<tr>
<td>Ludwig, Jeffrey T.</td>
<td>317, 320</td>
<td></td>
</tr>
<tr>
<td>Lum, David S.</td>
<td>367</td>
<td></td>
</tr>
<tr>
<td>Lumsdaine, Andrew</td>
<td>289, 296, 300</td>
<td></td>
</tr>
<tr>
<td>Luongo, Eleanora M.</td>
<td>367, 378</td>
<td></td>
</tr>
<tr>
<td>Lutwak, Robert I.</td>
<td>209, 215</td>
<td></td>
</tr>
<tr>
<td>Lyons, W. Gregory</td>
<td>270</td>
<td></td>
</tr>
<tr>
<td>Lyubomirsky, Ilya</td>
<td>99, 101</td>
<td></td>
</tr>
<tr>
<td>McQuirk, Ignacio S.</td>
<td>289, 292, 293</td>
<td></td>
</tr>
<tr>
<td>Meade, Robert D.</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td>Medical lasers</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>Meek, Kathy</td>
<td>7, 8</td>
<td></td>
</tr>
<tr>
<td>Melcher, Jennifer R.</td>
<td>397</td>
<td></td>
</tr>
<tr>
<td>Melloch, Michael R.</td>
<td>56, 76</td>
<td></td>
</tr>
<tr>
<td>Melngailis, John</td>
<td>345, 348</td>
<td></td>
</tr>
<tr>
<td>Mercil, Juliet</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td>Mesoscopic systems</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>Messier, Mark D.</td>
<td>277, 279</td>
<td></td>
</tr>
<tr>
<td>Metal surface studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemical reactions</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>Phase transitions</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>Semiconductors</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td>Structural analysis</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>Mewes, Marc O.</td>
<td>209, 225</td>
<td></td>
</tr>
<tr>
<td>Migliuolo, Stefano</td>
<td>231, 249</td>
<td></td>
</tr>
<tr>
<td>Middleson, J.</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>MIT Advanced Telecommunica Research Program</td>
<td>329—334</td>
<td></td>
</tr>
<tr>
<td>MIT Artificial Intelligence Laboratory</td>
<td>140, 384</td>
<td></td>
</tr>
<tr>
<td>MIT Audio Coder</td>
<td>331</td>
<td></td>
</tr>
<tr>
<td>MIT Center for Electromagnetic Theory and Applications</td>
<td>263</td>
<td></td>
</tr>
<tr>
<td>MIT Center for Space Research</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>MIT Digitized Speech Error Database</td>
<td>361</td>
<td></td>
</tr>
<tr>
<td>MIT Integrated Circuits Laboratory</td>
<td>65, 103, 312</td>
<td></td>
</tr>
<tr>
<td>MIT Laboratory for Information and Decision Systems</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>MIT Lincoln Laboratory</td>
<td>124, 148, 153, 270, 271, 309, 317</td>
<td></td>
</tr>
<tr>
<td>MIT Microsystems Technology Laboratories</td>
<td>154</td>
<td></td>
</tr>
<tr>
<td>MIT Microwave Temperature Sounder</td>
<td>283</td>
<td></td>
</tr>
<tr>
<td>MIT NanoStructures Laboratory</td>
<td>61—90</td>
<td></td>
</tr>
<tr>
<td>MIT-Woods Hole Oceanographic Institution Joint Program</td>
<td>317</td>
<td></td>
</tr>
<tr>
<td>Mnushkina, Irina</td>
<td>35, 37</td>
<td></td>
</tr>
<tr>
<td>Mochrie, Simon G.J.</td>
<td>169, 187</td>
<td></td>
</tr>
<tr>
<td>Moel, Alberto M.</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Molecular beam epitaxy</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Molinaar, Rachel</td>
<td>355</td>
<td></td>
</tr>
<tr>
<td>Mondol, Mark K.</td>
<td>61, 65</td>
<td></td>
</tr>
<tr>
<td>Monta, Peter A.</td>
<td>329, 331, 332</td>
<td></td>
</tr>
<tr>
<td>Monteiro, José C.</td>
<td>289, 305, 307</td>
<td></td>
</tr>
<tr>
<td>Moolji, Akbar A.</td>
<td>29, 61, 85</td>
<td></td>
</tr>
<tr>
<td>Moon, Euclid E.</td>
<td>61, 66, 79</td>
<td></td>
</tr>
<tr>
<td>Moore, Christopher B.</td>
<td>277, 279, 280</td>
<td></td>
</tr>
<tr>
<td>Moors, John D.</td>
<td>99, 101</td>
<td></td>
</tr>
<tr>
<td>Morshidi, Wan</td>
<td>209, 225</td>
<td></td>
</tr>
<tr>
<td>Mou, Alex X.</td>
<td>263</td>
<td></td>
</tr>
<tr>
<td>Moyne, William P.</td>
<td>309</td>
<td></td>
</tr>
<tr>
<td>Muccio, Eduardo R.</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Mueller, Jason</td>
<td>367</td>
<td></td>
</tr>
<tr>
<td>Muendel, Martin H.</td>
<td>99, 124, 129, 130</td>
<td></td>
</tr>
</tbody>
</table>
Music
  Mathematical models  323
  Musicus, Bruce R.  317

N
  Nabors, Keith S.  289, 303
  Nadeau, Philip P.  367
  Nadol, Joseph  359
  Naqvi, Asad A.  161, 165
  Narula, Aradhana  329, 333
  Nassi, Giulia Arman  355
  Nassi, Marco  231, 249
  Natarajan, Vasant  209, 218
  National Aeronautics and Space Administration  277
  National Center for Integrated Photonic Technology  48
  National Nanofabrication Facility  223
  National Radio Astronomy Observatory  277, 279
  National Synchrotron Light Source  169, 187
  Nawab, S. Hamid  320
  Nelson, Lynn E.  99, 100
  New England Eye Center  119
  Nicolas, Julien J.  329, 333
  Nilsen, Mark R.  397
  Noise
    Active Noise Cancellation  322, 324
    Nonlinear waves in plasmas  237
  North, D. Keith  355
  Nuytken, P. R.  18, 21
  Nygren, Erik  367, 388

O
  O'Connell, Michael P.  367, 378
  O'Neill, Kevin  263, 266
  Oates, John H.  263, 265
  Oceanography
    Signal processing  325
  Odoardi, Angela R.  7, 29, 35, 263
  Ohkawa, Hana  231, 249
  Oka, Toshifusa  412
  Ooi, James M.  317, 321
  Oppenheim, Alan V.  317—327
  Optical coherence tomography  118
  Optical communication  139—143
    Devices  7—28
    Fiber optic rings  19
    Optical fibers  19
    Squeezed states  139
  Optical fibers  100, 101, 105
  Optical frequency  141
    Division  141
    Synthesis  141
  Optics  99—137
  Orlando, Antonio  263, 265
  Orlando, Terry P.  76, 79, 263, 269—273, 269, 303
  Osgood, Gene T.  289
  Ostendorf, Mari  361
  Owen, Gabrielle M.  61, 64, 65
  Palmer, Fred L.  209, 218
  Pan, Charlie  367, 375
  Pant, Amrit R.  209, 221
  Papadopoulos, Haralabos C.  317, 321
  Parkes-MIT-NRAO Survey  277
  Passive waveguides  45
  Pava, Jerôme M.  99
  Peake, William T.  397, 399
  Pelley, J. David  209, 225
  Peng, Lung-Han  7, 11, 22, 24
  Penn, Gregory E.  231, 249
  Peregryn, Denis J.  317, 325
  Perilli, Richard R.  7
  Perkell, Joseph S.  355
  Perreault, Brian M.  317, 322
  Petrich, Gale S.  35, 37, 42, 45, 46, 48
  Pfeifer, Alexsander  329, 334
  Phase transitions  161
  Phillips, Joel R.  263, 269, 289, 303
  Phillips Laboratory  249
  Photonic devices
    Channel-dropping filters  45
  Pind, Jorgen  355
  Ploch, Nicholas  367, 388
  Plant, Geoffrey L.  367, 375
  Plasma physics  231—260
  Plasma wave interactions  237
  Plotnik, Irving  61
  Polley, Michael O.  337
  Poort, Kelly L.  355
  Porter, Jeanne M.  61, 86
  Portnoy, Vladislav  231
  Power, Matthew H.  367
  Prasad, Sheila  7, 26
  Preisig, James C.  317
  Prentiss, Mara G.  195
  Price, Patti  361
  Princeton University  249
Project Staff and Subject Index

Pritchard, David E.  209, 218—228
Process flow representation  309
Psaltis, D.  15
Pullman, David P.  175, 177, 178, 179
Puria, Sunil  397, 399

Q
Quantum dots  51, 57, 91, 161
Quantum heterostructures  7—28
Quantum mechanics  91
Quantum studies
Optics  139
Quantum-effect devices  61—90

R
Rabinowitz, William M.  367, 374—375, 378, 379, 397, 404, 408
Radar
Laser  140
Radio astronomy  277—284
Radio telescopes
Very Large Array  277, 279
Very Long Baseline Array  277, 279, 280
Radio transients  280
Rahmat, Khalid  289, 300, 301
Ram, Abhay K.  231, 237
Ramaswamy, Malini  99, 109, 113
Ramstad, Monte J.  169
Rankovic, Christine M.  367
Rathman, Dennis D.  345, 348
Ravicz, Michael E.  397, 399
Rediker, Robert H.  139, 143
Reed, Charlotte M.  367, 375—378
Reichelt, Mark W.  289, 292, 296, 300
Relativistic electron beams  231
Remote sensing  263
Renormalization-Group Theory  161
RF heating and current drive  237
Rhee, Kee  72
Ricci, Kenneth N.  231
Richard, Michael D.  317, 322
Riconda, Caterina  231, 249
Rider, Todd H.  231, 249
Rinderknecht, W. John  289, 305
Robots  367, 380, 384
Rogers, Barrett  231, 249
Ronan, Diane E.  367
Rosenkranz, Philip W.  277, 282, 283
Rosowski, John J.  397, 399, 401
Rossetti, Denise M.  329
Royter, Yakov  7, 19, 20
Rubenstein, Richard A.  209, 221
Rubinstein, Jay T.  397, 404

S
Sakamoto, Makoto  317
Salcedo, Ano  263, 266
Sallsbury, J. Kenneth  367, 384, 389
Santiago, Sally C.  317
Savas, Timothy A.  99, 124, 125, 129
Scheffler, Roeland P.  355
Schäpers, Thomas  145
Schattenburg, Mark L.  61, 64, 65, 69, 86
Schrock, Stephen F.  317, 323
Schmidt, Martin A.  289, 302
Schmiedmayer, H. Joerg  209, 221
Schreiber, William F.  337—339
Schultz, Steven D.  231, 237
Schwartz, Michael J.  277, 282, 283
Scire-Scappuzzo, Francesca  263
Searle, Campbell L.  455
Seidel, Mark N.  289, 292, 296, 298
Sekiya, Kaoru  367
Semi-conductor alloys  161
Semi-conductor surface studies  161, 181
Semi-conductors  91, 104
Compound  29, 35—48
II-VI  37
Single-electron transistors  53
Surface studies  187
Sensory aids  397—408
Sensory communication  367—393
Sendur, Stephen D.  289, 302
Seo, Chulhun  7
Shah, Devang M.  397, 399
Shah, Satyen  61, 70, 86
Shahrar, Selim M.  195, 455
Shao, Michael  277
Shapiro, Jeffrey H.  139—143
Shattuck-Hufnagel, Stephanie R.  355
Shatz, Lisa F.  397, 399
Shen, Amelia H.  289, 305, 306
Shenoy, Krishna V.  7, 13, 14, 15, 18, 20
Sheard, Scott R.  139
Shih, Shih-en  263
Shin, Robert T.  263, 265
Shinn-Cunningham, Barbara G.  367, 379
Shirasaki, Masataka  99, 101
Shu, Li  317
Siebert, William M.  397, 455
<table>
<thead>
<tr>
<th>Topic</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal processing</td>
<td>7, 284, 317—327, 355, 404</td>
</tr>
<tr>
<td>Chaos</td>
<td>321</td>
</tr>
<tr>
<td>Chaotic systems</td>
<td>318, 319, 322</td>
</tr>
<tr>
<td>Mathematical models</td>
<td>323</td>
</tr>
<tr>
<td>Noise</td>
<td>322</td>
</tr>
<tr>
<td>Speech</td>
<td>323</td>
</tr>
<tr>
<td>Silitch, Alexis P.</td>
<td>225</td>
</tr>
<tr>
<td>Silveira, Luis M.</td>
<td>289, 292, 298, 304</td>
</tr>
<tr>
<td>Silverman, Scott E.</td>
<td>61, 62, 70</td>
</tr>
<tr>
<td>Simons, Benjamin D.</td>
<td>51</td>
</tr>
<tr>
<td>Singer, Andrew C.</td>
<td>317, 323</td>
</tr>
<tr>
<td>Sisson, Robert D.</td>
<td>61, 65</td>
</tr>
<tr>
<td>Smet, Jurgen H.</td>
<td>7, 11, 22, 25, 145, 154</td>
</tr>
<tr>
<td>Smith, Brian M.</td>
<td>61</td>
</tr>
<tr>
<td>Smith, Clare F.</td>
<td>277</td>
</tr>
<tr>
<td>Smith, Henry I.</td>
<td>7, 8, 45, 48, 61—90, 103</td>
</tr>
<tr>
<td>Smith, Stephen P.</td>
<td>195</td>
</tr>
<tr>
<td>Smith, William H.</td>
<td>455</td>
</tr>
<tr>
<td>Smullin, Louis D.</td>
<td>99</td>
</tr>
<tr>
<td>Sodini, Charles G.</td>
<td>289, 292, 293, 294, 296, 297</td>
</tr>
<tr>
<td>Solid-state far-infrared lasers</td>
<td>154</td>
</tr>
<tr>
<td>Solid-state materials</td>
<td>106</td>
</tr>
<tr>
<td>Sonar</td>
<td>391</td>
</tr>
<tr>
<td>Song, Seungheon</td>
<td>187</td>
</tr>
<tr>
<td>Souyris, Jean-Claude</td>
<td>263</td>
</tr>
<tr>
<td>Speech communication</td>
<td>355—364</td>
</tr>
<tr>
<td>Speech intelligibility</td>
<td>367</td>
</tr>
<tr>
<td>Speech models</td>
<td>334</td>
</tr>
<tr>
<td>Speech physiology</td>
<td>355—364</td>
</tr>
<tr>
<td>Speech processing</td>
<td>334</td>
</tr>
<tr>
<td>Speech reception</td>
<td>367—379</td>
</tr>
<tr>
<td>Speech recognition</td>
<td>323</td>
</tr>
<tr>
<td>Speechreading</td>
<td>367, 375</td>
</tr>
<tr>
<td>Spellmeyer, Neal W.</td>
<td>209</td>
</tr>
<tr>
<td>Spina, Michelle S.</td>
<td>277, 283, 284</td>
</tr>
<tr>
<td>Spinnrad, H.</td>
<td>277</td>
</tr>
<tr>
<td>Srinivasan, Mandayam A.</td>
<td>367, 375, 380—388, 389</td>
</tr>
<tr>
<td>Stadler, Robert W.</td>
<td>367</td>
</tr>
<tr>
<td>Staelin, David H.</td>
<td>277, 282—284, 456</td>
</tr>
<tr>
<td>Stankovic, Konstantina</td>
<td>397, 403</td>
</tr>
<tr>
<td>Stannite</td>
<td>161</td>
</tr>
<tr>
<td>Statistical mechanics</td>
<td>161</td>
</tr>
<tr>
<td>Stefanov-Wagner, Frank J.</td>
<td>397</td>
</tr>
<tr>
<td>Steffens, David A.</td>
<td>397</td>
</tr>
<tr>
<td>Stevens, Kenneth N.</td>
<td>355—364</td>
</tr>
<tr>
<td>Stoner, Richard E.</td>
<td>231</td>
</tr>
<tr>
<td>Stroock, Abraham D.</td>
<td>209, 218</td>
</tr>
<tr>
<td>Su, Lisa T.-F.</td>
<td>72</td>
</tr>
<tr>
<td>Sugiyama, Linda E.</td>
<td>231, 249</td>
</tr>
<tr>
<td>Sun, Chi-Kuang</td>
<td>99, 113</td>
</tr>
<tr>
<td>Sun, Ke-Xun</td>
<td>139</td>
</tr>
<tr>
<td>Sunshine, Lon E.</td>
<td>329, 334</td>
</tr>
<tr>
<td>Superconducting Josephson devices</td>
<td>149, 269</td>
</tr>
<tr>
<td>Superconductors</td>
<td>145</td>
</tr>
<tr>
<td>Surface studies</td>
<td>161—167, 169—173</td>
</tr>
<tr>
<td>Svirsky, Mario A.</td>
<td>355</td>
</tr>
<tr>
<td>Svolos, George M.</td>
<td>231, 249</td>
</tr>
<tr>
<td>Swarup, Nick</td>
<td>367, 389</td>
</tr>
<tr>
<td>Synthetic humans</td>
<td>390</td>
</tr>
<tr>
<td>T</td>
<td></td>
</tr>
<tr>
<td>Tada, Hiroaki</td>
<td>412</td>
</tr>
<tr>
<td>Takeuchi, Annie H.</td>
<td>367</td>
</tr>
<tr>
<td>Tamura, Kohichi R.</td>
<td>99, 100</td>
</tr>
<tr>
<td>Tan, Chin Hwee</td>
<td>289, 290</td>
</tr>
<tr>
<td>Tan, Hong Z.</td>
<td>367, 375</td>
</tr>
<tr>
<td>Taniguchi, Nobuhiko</td>
<td>51</td>
</tr>
<tr>
<td>Tannian, Bridget E.</td>
<td>209, 221</td>
</tr>
<tr>
<td>Tassoudji, M. Ali</td>
<td>263, 265, 266</td>
</tr>
<tr>
<td>Tearney, Guillermo J.</td>
<td>99, 119</td>
</tr>
<tr>
<td>Telecommunications research</td>
<td>329—334</td>
</tr>
<tr>
<td>Teleoperator systems</td>
<td>388</td>
</tr>
<tr>
<td>Television research</td>
<td></td>
</tr>
<tr>
<td>High-definition</td>
<td>329—334, 337—339</td>
</tr>
<tr>
<td>Thermionuclear plasmas</td>
<td>249</td>
</tr>
<tr>
<td>Thin films</td>
<td>169</td>
</tr>
<tr>
<td>Thompson, Carl V.</td>
<td>48, 88</td>
</tr>
<tr>
<td>Thompson, Stanley H.</td>
<td>209, 225</td>
</tr>
<tr>
<td>Tierney, Joseph</td>
<td>397, 404, 408</td>
</tr>
<tr>
<td>Tokamaks</td>
<td>249</td>
</tr>
<tr>
<td>Touch</td>
<td>367, 375, 380—391</td>
</tr>
<tr>
<td>Trautman, Theodore R.</td>
<td>175</td>
</tr>
<tr>
<td>Troxel, Donald E.</td>
<td>309—313</td>
</tr>
<tr>
<td>Tsang, Leung</td>
<td>263</td>
</tr>
<tr>
<td>Tsekouras, Athanasios A.</td>
<td>175, 177, 179</td>
</tr>
<tr>
<td>Tsuk, Michael</td>
<td>263, 265</td>
</tr>
<tr>
<td>Tufts New England Medical Center</td>
<td>119</td>
</tr>
<tr>
<td>Turk, Alice</td>
<td>355</td>
</tr>
<tr>
<td>Tuyo, Michael T.</td>
<td>367</td>
</tr>
<tr>
<td>Twente University (The Netherlands)</td>
<td>149</td>
</tr>
<tr>
<td>U</td>
<td></td>
</tr>
<tr>
<td>U.S. Navy. Naval Research Laboratory</td>
<td>61</td>
</tr>
<tr>
<td>Uchanski, Rosalie M.</td>
<td>367</td>
</tr>
<tr>
<td>Ugarov, Andrew</td>
<td>367, 375</td>
</tr>
<tr>
<td>Ugarte, Alberto</td>
<td>263, 266</td>
</tr>
<tr>
<td>Ultschmied, Jiri</td>
<td>231</td>
</tr>
<tr>
<td>Ulman, Morrison</td>
<td>99, 109, 113</td>
</tr>
<tr>
<td>Umminger, Christopher B.</td>
<td>289, 292, 294</td>
</tr>
<tr>
<td>Underwater acoustics</td>
<td>325, 391</td>
</tr>
<tr>
<td>University of California, Berkeley</td>
<td>277</td>
</tr>
<tr>
<td>University of Pennsylvania, Linguistics Data Con-</td>
<td>361</td>
</tr>
<tr>
<td>sortium</td>
<td></td>
</tr>
</tbody>
</table>
Project Staff and Subject Index

University of Texas  249
University of Wisconsin  401
Unver, Evren R.  309
Utz, Arthur L.  175

V
Vacca, Luigi  231, 237
van der Zant, Herre S.J.  263, 269, 270, 271, 289, 303
Verbout, Shawn M.  317, 323
Noise  324
Signal processing
Verghese, George C.  319
Verghese, Simon  145, 148
Vestibular system  403
Veysooglu, Murat E.  263, 265
Virtual environments  367, 379, 388—393
Visible emitters  37
Vision chip project  292
Vollbehn, Pavel S.  231
von Bosau, Laura M.  231
Voss, Susan E.  397

W
Wage, Kathleen E.  317, 326
Wang, Jing  181
Wang, Li-Fang  263, 266
Watanabe, Akira  413
Wave propagation  325
Weber, E.  185
Wee, Susie J.  337
Wehinger, Stefan  209, 221
Weiner, Miriam  231
Weinstein, Ehud  317
Weisner, Jerome B.  456
Weiss, Thomas F.  397—408
Welker, Daniel P.  367, 378
Wells, Barry  169
White, Jacob K.  269, 270, 289, 292, 296, 298, 300—305
Wiegand, Thomas E.  367, 388
Wies, Evan F.  367, 388
Wilde, Lorin F.  355
Williams, David  355
Willsky, Alan S.  140
Wind, Shalom  53
Wint, Arlene E.  355
Wong, Ngai C.  139, 141
Wong, Vincent V.  61, 62, 64, 65, 79, 83, 103
Wong, William  99, 101
Woods Hole Oceanographic Institution  317
Wornell, Gregory W.  317, 320, 321, 455
Wozniak, Jane  355
Wozniak, Jennifer A.  367
Wright, Alan  277
Wurtele, Jonathan S.  231
Wuu, John J.  209, 225
Wyatt, C.  14
Wyatt, John L., Jr.  289, 292—300
Wyss, Rolf A.  145

X
X-ray lithography  8, 61—90
Xia, Jiqing  263
Xu, Yi  355

Y
Yale University School of Medicine  384
Yamaguchi, Masanori  263
Yang, Isabel Y.  61, 64, 65, 72
Yang, Julius J.  175, 177, 179
Yang, Y. Eric  263, 265, 266
Yasaka, Anto  61, 70
Yee, Kenneth W.  61
Yesley, Peter S.  209, 225
Yie, He  289, 302
Yin, T.C.T.  401
Yoo, Chang Dong  329, 334
Young, Mary E.  455
Young, Michael J.  169
Yu, Charles  99
Yu, Paul C.  289, 292, 294
Yuan, Yi  263, 265

Z
Zangi, Kambiz C.  317, 324
Zeltzer, David  367, 388—393
Zhang, Zhe  175, 177, 179
Zhu, Meng Y.  397
Zilles, Craig B.  367, 389
Zincblende  161
Zissman, Marc A.  397, 404, 408
Zue, Victor W.  367
Zurek, Patrick M.  367, 378—379