OPTICAL SCIENCES

A. Krier (Ed.)

Mid-infrared Semiconductor Optoelectronics

Springer Series in

OPTICAL SCIENCES **118**

---

founded by H.K.V. Lotsch

Editor-in-Chief: W.T. Rhodes, Atlanta

Editorial Board: T. Asakura, Sapporo
K.-H. Brenner, Mannheim
T.W. Hänsch, Garching
T. Kamiya, Tokyo
F. Krausz, Wien and Garching
B. Monemar, Linköping
H. Venghaus, Berlin
H. Weber, Berlin
H. Weinfurter, München

Springer Series in

OPTICAL SCIENCES

The Springer Series in Optical Sciences, under the leadership of Editor-in-Chief William T. Rhodes, Georgia Institute of Technology, USA, provides an expanding selection of research monographs in all major areas of optics: lasers and quantum optics, ultrafast phenomena, optical spectroscopy techniques, optoelectronics, quantum information, information optics, applied laser technology, industrial applications, and other topics of contemporary interest.
With this borad coverage of topics, the series is of use to all research scientists and engineers who need up-to-date reference books.
The editors encourage prospective authors to correspond with them in advance of submitting a manuscript. Submission of manuscripts should be made to the Editor-in-Chief or one of the Editors.

Editor-in-Chief

William T. Rhodes
Georgia Institute of Technology
School of Electrical and Computer Engineering
Atlanta, GA 30332-0250, USA
E-mail: bill.rhodes@ece.gatech.edu

Editorial Board

Toshimitsu Asakura
Hokkai-Gakuen University
Faculty of Engineering
1-1, Minami-26, Nishi 11, Chuo-ku
Sapporo, Hokkaido 064-0926, Japan
E-mail: asakura@eli-hokkai-s-u.ac.jp

Karl-Heinz Brenner
Chair of Optoelectronics
University of Mannheim
Institute of Computer Engineering
B6, 26
68131 Mannheim, Germany
E-mail: brenner@uni-mannheim.de

Theodore W. Hänsch
Max-Planck-Institut für Quantenoptik
Hans-Kopfermann-Straße 1
85748 Garching, Germany
E-mail: t.w.haensch@physik.uni-muenchen.de

Takeshi Kamiya
Ministry of Education, Culture, Sports
Science and Technology
National Institution for Academic Degrees
3-29-1 Otsuka, Bunkyo-ku
Tokyo 112-0012, Japan
E-mail: kamiyatk@miad.ac.jp

Ferenc Krausz
Vienna Institute of Technology
Photonics Institute
Gusshausstraße 27/387
1040 Wien, Austria
E-mail: ferenc.krausz@tuwien.ac.at
and
Max-Planck-Institut für Quantenoptik
Hans-Kopfermann-Straße 1
85748 Garching, Germany

Bo Monemar
Department of Physics
and Measurement Technology
Materials Science Division
Linköping University
58183 Linköping, Sweden
E-mail: bom@ifm.liu.se

Herbert Venghaus
Heinrich-Hertz-Institut
für Nachrichtentechnik Berlin GmbH
Einsteinufer 37
10587 Berlin, Germany
E-mail: venghaus@hhi.de

Horst Weber
Technische Universität Berlin
Optisches Institut
Straße des 17. Juni 135
10623 Berlin, Germany
E-mail: weber@physik.tu-berlin.de

Harald Weinfurter
Ludwig-Maximilians-Universität München
Sektion Physik
Schellingstraße 4/III
80799 München, Germany
E-mail: harald.weinfurter@physik.uni-muenchen.de

A. Krier (Ed.)

Mid-infrared Semiconductor Optoelectronics

With 443 Figures

Anthony Krier, PhD
Physics Department
Lancaster University
Lancaster
LA1 4YB
UK
E-mail: a.krier@lancaster.ac.uk

British Library Cataloguing in Publication Data
Mid-infrared semiconductor optoelectronics. - (Springer
series in optical sciences)
1.Optoelectronics 2.Infrared technology 3.Semiconductors
4.Lasers
I.Krier, A.
621.3'81045
ISBN-10: 184628208X

Library of Congress Control Number: 2006924636

Springer Series in Optical Sciences ISSN 0342-4111
ISBN-10: 1-84628-208-X e-ISBN 1-84628-209-8 Printed on acid-free paper
ISBN-13: 978-1-84628-208-9

© Springer-Verlag London Limited 2006

Whilst we have made considerable efforts to contact all holders of copyright material contained in this book, we may have failed to locate some of them. Should holders wish to contact the Publisher, we will be happy to come to some arrangement with them.

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

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

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Printed in Germany

9 8 7 6 5 4 3 2 1

Springer Science+Business Media
springer.com

Preface

The mid–infrared spectral region is of enormous interest as the practical realisation of optoelectronic devices operating in the 2–10 μm wavelength range offers potential applications in a wide variety of areas including; optical gas sensing and environmental monitoring, free-space optical communications, infrared counter-measures, clean energy generation, biomedical and thermal imaging. For example, the mid–infrared contains the fundamental fingerprint absorption bands of a number of pollutant and toxic gases and liquids; methane (3.3μm), CO_2 (4.6μm), CO (4.2μm), NOx (6.5μm) and SOx (7.3μm) require accurate, in situ multi-component monitoring in a variety of different situations (*e.g.* oil-rigs, coal mines, landfill sites, car exhausts) and in concentrations, ranging from ppb to almost 100%. The mid–infrared region also provides unique fingerprints with strong absorption bands for drug intermediates, pharmaceuticals, narcotics and bio-chemcials allowing highly selective and sensitive detection or imaging. Consequently, the mid–infrared is very attractive for the development of sensitive optical sensor instrumentation. In addition there is an atmospheric transmission window between 3 and 5 μm which enables free space optical communications and thermal imaging applications in both civil and military situations as well as the development of infrared countermeasures for homeland security. However, the advantages of this wavelength range have not been fully exploited due to the limitations in current technology. Development of high bandwidth sources and detectors for optical fibre based telecommunications at wavelengths up to 1.55μm has achieved a high art. However, suitable sources and detectors for longer wavelength mid–infrared applications do not yet exist. Most high sensitivity optical sensors and optical spectroscopy applications require that the lasers operate single mode and at room temperature. Biochemical/chemical measurements need to be made *in situ* or at point of care to be useful and this requires miniaturised technology which is sensitive and inexpensive. Further development is needed in order to realise the optoelectronic devices for these requirements. Many of the necessary improvements are intimately tied to our immature ability to fabricate high-quality epitaxial materials and device structures, as well as our incomplete understanding of how to manipulate their optoelectronic properties. Mid-infrared semiconductor quantum heterostructures that require precisely-controlled nanometer-sized layers significantly challenge the available growth and processing techniques, methodologies for band-structure modelling, characterization tools, and even the knowledge of fundamental materials parameters.

This book is an overview of the current status of science and technology associated with the development of semiconductor opto-electronic devices for the mid–infrared spectral range. The book is divided into four sections and although each of the chapters are individually self-contained we begin with a consideration of some fundamental physics which is more generally relevant to the narrow-gap III-V materials. In particular, the reader is given an overview of the main problems facing the design engineer in these materials and some of the possible solutions based on strained layer engineering (primarily in lasers) are considered and compared. The book is focused exclusively on *semiconductor* optoelectronic devices and so lasers,

light-emitting diodes, photodetectors and thermo-photovoltaic cells are discussed but optical parametric oscillators, fibre and crystal lasers are not considered here.

The section on lasers includes up-to-date contributions from experts working in a number of different areas including; strained Type I and Type II lasers, quantum cascade lasers, interface lasers and vertical cavity surface emitting lasers (VCSELs), principally in III-V compounds but also in those based on IV-VI compounds. Methods and structures to achieve high-power operation using both electrical and optical pumping are included to give a comprehensive picture of the current state of the art in each of these technologies. Following the development of the quantum cascade laser it has recently emerged that there is a technology gap in the 3-4 μm region and at present there are no room temperature CW lasers which operate in this window. Consequently, there is much interest in the development of an efficient room temperature semiconductor laser for this region. A number of different approaches are being investigated and at the moment it is still an open question as to which one will be successful.

The section on light-emitting diodes (LEDs) and detectors reviews recent progress and considers the key developments in these devices and also includes a unique chapter on negative luminescence. In addition to the more conventional photodetector structures, quantum well (QWIP), quantum dot (QDIP) and strained layer superlattice detectors are considered and particularly with respect to their use in focal plane arrays. Meanwhile, avalanche photodiodes for the mid–infrared are also reviewed and considered in detail.

The final section is concerned with applications and highlights once again the diversity and technological importance of the mid–infrared spectral region. Contributions include applications in the areas of optical gas sensing, thermo-photovoltaic (TPV) cells for clean energy conversion, biomedical applications for diagnosis and treatment, as well as infrared countermeasures. It has not been possible to include everything of interest in this rapidly expanding field however we have endeavoured to provide the reader with an up to date overview of the current status in mid–infrared semiconductor optoelectronics which it is hoped will at least provide a starting point for further research and investigation as well as a useful reference for the future.

Acknowledgement

I should like to thank all the contributors for their enthusiastic participation in this project, without whom it would not have been possible to provide such a comprehensive and in depth coverage of semiconductor mid–infrared opto-electronics. I am also grateful to Oliver Jackson and colleagues at Springer for their help and guidance in the technical production of this work. Finally, I should like to thank my wife, Susan for her patient and careful help in many ways, including proof–reading and preparation.

Tony Krier
Lancaster University
November, 2005

Contents

Part I Materials and Device Design Considerations

Theory of Mid-wavelength Infrared Laser Active Regions: Intrinsic Properties and Design Strategies
J.T. Olesberg and M.E. Flatté 3
1 Challenges and Opportunities in MWIR Laser Active Region Design 3
1.1 Intrinsic Material Properties Dominate Device Performance 4
1.2 Opportunities Provided by Heterostructure Design 5
2 Metrics for Material Comparison and Device Design 6
2.1 Single-stage Active Regions 8
2.1.1 Threshold Figure of Merit for Bulk GaAs 12
2.2 Choice of Active Region Thickness 14
2.2.1 Optimal Thickness of Bulk GaAs 14
2.3 Maximizing the Slope Efficiency 16
2.3.1 Maximising Efficiency for a specific Optical Power 16
2.4 Cascaded Active Regions 17
3 Band-edge Optimization of Active Region Materials 21
3.1 Density of States Imbalance 21
3.2 The Effects of Density of States Imbalance 25
3.3 Reducing Density of States Imbalance Using Strain 30
3.4 Reducing Density of States Imbalance Using Quantum Confinement 30
3.5 Development of 2–3 μm GaInAsSb Diodes 32
3.6 Optimizing Band Edge Properties Using InAs/GaSb/AlSb Materials 40
4 Final State Optimization of Active Region Materials 47
4.1 Suppression of Intersubband Absorption 48
4.2 Suppression of Auger Recombination 50
5 Cavity-integrated Active Region Design for Optimal Laser Performance 53
5.1 Optimization of the Active Region 53
5.2 Transport and Optical Cladding 56
6 Calculating the Electronic and Optical Properties of Semiconductors 59
6.1 Electronic Structure and Momentum Matrix Elements 59
6.1.1 *k.p* methods for Bulk Materials 61
6.1.2 *k.p* Methods for Heterostructures 63
6.1.3 Influence of Interface Bonds 68

6.2 Comparison with Experiment ... 70
6.2.1 Non-equilibrium Absorption ... 70
6.2.2 Auger Recombination ... 73
7 Concluding Remarks ... 75
Appendix. Derivation of Expressions for the Electronic and Optical Properties of Semiconductors ... 76
A1 Density of States ... 77
A2 Optical Properties ... 77
A3 Carrier Recombination Rates ... 80
A4 Carrier Mobilities ... 82
References ... 84

Bandstructure and High-pressure Measurements
B.N. Murdin, A.R. Adams and S.J. Sweeney ... 93
1 Bandstructure Calculation by the k.p Method ... 93
1.1 Bulk Bandstructure ... 94
1.1.1 Zinc-blende Crystals ... 96
1.1.2 Lead-salts ... 99
1.2 Quantum Well Structures ... 101
1.3 Temperature Dependence of the Bandstructure ... 103
1.4 Hydrostatic Pressure and Strained Layer Quantum Wells ... 103
2 Transition Rates ... 105
2.1 Interband Optical Transition Rates ... 106
2.2 Auger Recombination ... 106
2.3 Auger Rates in Bulk Zinc Blende Materials ... 109
2.4 Auger Suppression in Bulk Materials: Pb-salts ... 110
2.5 Auger Suppression in Bulk Materials: Dilute III-V Nitrides ... 111
2.6 Auger Suppression in Quantum Wells ... 112
3 High Pressure ... 114
3.1 Near-infrared Applications ... 114
3.1.1 Radiative Recombination ... 114
3.1.2 Non-radiative Recombination ... 118
3.2 Determination of Mid-IR Material Pressure Coefficients ... 120
3.3 Interband Mid-IR Devices ... 121
3.3.1 InAs LEDs ... 121
3.3.2 Comparison of Near-IR and 2.37μm InGaAsSb Lasers ... 122
4 Conclusion ... 125
References ... 126

Part II Lasers

III-Sb-based Type-I QW Diode Lasers
M. Rattunde, J. Schmitz, C. Mermelstein, R. Kiefer and J. Wagner ... 131
1 Introduction ... 131
2 III-Sb-based Material System ... 132
2.1 AlGaAsSb ... 132

2.2 Strained GaInAsSb Layers 132
2.3 Laser Structure 134
3 Fabrication of (AlGaIn)(AsSb)-based Diode Lasers 136
3.1 Growth 136
3.2 Epitaxial Layer Characterization 137
3.3 Device Processing 139
4 Gain and Loss Mechanisms in III-Sb-based Lasers Emitting at Around 2 μm 140
4.1 Optical Gain and Electric Loss in the QW Active Region 140
4.2 Optical Losses in the Laser Structure 143
5 High-power Performance of 2 μm III-Sb-based Diode Lasers 146
6 Long-wavelength III-Sb-based Type-I QW Lasers 150
7 Outlook 151
Acknowledgments 153
References 153

VCSELs Emitting in the 2 – 3 μm Wavelength Range
F. Genty, A. Garnache and L. Cerutti 159
1 Introduction 159
2 A Short Description of Vertical-cavity Lasers 160
2.1 Conventional Microcavity VCSELs 160
2.2 External-cavity VCSELs (VECSELs) 163
2.2.1 Structure Description 163
2.2.2 Geometrical Stability of a Concave/plane Laser Cavity 165
3 VCSELs Emitting in the 2–3 μm Range 168
3.1 CdHgTe-based Structures – CEA, Grenoble (France) 169
3.2 InP-based Structures – Walter Schottky Institute, University of Munich, Garching (Germany) 171
3.3 GaSb-based Structures I – Naval Research Laboratory/ Hughes Research Laboratory (USA) 173
3.4 GaSb-based Structures II – University of Montpellier 2 (France) 176
3.4.1 Electrically-pumped Microcavity VCSEL Emitting Near 2.2 μm 176
3.4.2 Optically-pumped GaSb-based VCSEL Structures 177
4 Conclusion 184
Acknowledgments 185
References 185

Antimonide Type-II “W” Lasers
I. Vurgaftman, W.W. Bewley, C.L. Canedy, C.S. Kim, J. R. Lindle, M. Kim, and J.R. Meyer 189
1 Introduction 189
2 Advances in the MBE Growth of “W” Laser Structures 191
3 Optically Pumped “W” Lasers 196
4 Single-stage “W” Diode Lasers 199
5 Interband Cascade “W” Lasers 201
6 “W” VCSELs and PCDFB Lasers 203
7 Critical Issues in Improving the Performance of “W” Lasers 208

8 Conclusions ... 212
References ... 213

Interface Lasers with Asymmetric Band Offset Confinement
K.D. Moiseev and Y.P. Yakovlev ... 219
1 Introduction ... 219
2 The 2D-Electron Channel in a Type II Broken-gap p-GaInAsSb/p-InAs Heterointerface ... 220
3 Interface Luminescence Properties of the Type II Broken-gap Single p-GaInAsSb/p-InAs Heterojunction ... 220
3.1 Interface EL in a Single p-GaInAsSb/p-InAs Heterojunction ... 220
3.2 Tunnelling-injection Laser Based on the Type II p-$GaIn_{0.17}AsSb$/p-$In_{0.83}GaAsSb$ Heterojunction ... 224
3.3 Suppression of Auger-recombination Processes at the Type II Heterointerface ... 226
4 Interface Laser with Improved Temperature Dependence ... 228
4.1 Electroluminescence in a Single p-$GaIn_{0.17}AsSb$/n-$In_{0.83}GaAsSb$ Heterostructure ... 228
4.2 Advanced Tunnel-injection Heterostructure Laser ... 230
4.3 Prospects for High-temperature Luminescence in Type II Broken-gap Heterojunctions ... 233
5 Conclusions ... 235
References ... 235

IV-VI Semiconductors for Mid-infrared Optoelectronic Devices
P. J. McCann ... 237
1 Introduction ... 237
2 Spectroscopy with IV-VI Semiconductor Lasers ... 237
3 IV-VI Semiconductor Growth and Characterization ... 241
3.1 IV-VI Layers on BaF_2 ... 242
3.2 IV-VI Layers on Silicon ... 245
4 Self-heating Effects in IV-VI Mid-IR Lasers ... 248
5 Electrophonon Resonance in PbSe QWs ... 249
6 Progress in Laser Fabrication Using Substrate Removal ... 254
7 Summary ... 259
Acknowledgements ... 260
References ... 260

Mid-infrared Vertical Cavity Surface Emitting Lasers Based on the Lead Salt Compounds
G. Springholz, T. Schwarzl and W. Heiss ... 265
1 Introduction ... 265
2 Vertical Surface Emitting Lasers ... 266
3 Lead Salt-based Bragg Interference Mirrors ... 268
3.1 Mirror Materials ... 268
3.2 Examples and Results ... 270
4 Lead Salt Vertical Cavity Surface Emitting Lasers ... 272
4.1 Design Issues and Resonator Structure ... 272

4.2 PbTe/EuTe Quantum Well VCSELs ... 274
4.2.1 Structure and Optical Characterization ... 274
4.2.2 Threshold and Laser Emission ... 275
4.2.3 Temperature Dependence of Emission ... 276
4.3 Room Temperature PbTe QW VCSELs ... 277
4.4 PbSe/PbSrSe VCSELs ... 280
5 Self-assembled Infrared Quantum Dot Lasers ... 281
5.1 Self-assembled PbSe Quantum Dots ... 281
5.2 Quantum Dot VCSELs: Growth and Characterization ... 284
5.3 Quantum Dot Laser Emission ... 286
6 Lead Salt VCSELs with Different Active Regions ... 288
6.1 VCSEL Structures and Optical Properties ... 288
6.2 Laser Emission ... 289
6.2.1 Bulk-like PbTe VCSEL ... 289
6.2.2 PbTe/PbEuTe Quantum Well VCSEL ... 291
6.2.3 Lasing Properties of the Quantum Dot VCSEL ... 292
6.2.4 Tuning Properties ... 292
7 CW-VCSELs Emitting at 6-8 μm ... 293
7.1 Structure and Optical Properties ... 293
7.2 Laser Emission ... 294
8 Conclusions ... 298
Acknowledgements ... 298
References ... 299

Optically Pumped MIR Lasers
R. Kaspi, G.C. Dente and A.P. Ongstad ... 303
1 Introduction ... 303
2 Optically Pumped Laser Design ... 303
2.1 Guidelines for Optical Pumping and the MIR-OPSL ... 303
2.2 Type-II Quantum Wells ... 306
2.3 Dielectric Waveguide Design for the MIR-OPSL ... 309
2.3.1 Connection Between Confinement Factor and Beam Quality ... 309
2.3.2 Ghost Modes in the MIR-OPSL ... 313
3 Laser Characteristics ... 315
3.1 Spectral Power and Loss Measurements ... 315
3.2 Beam Quality ... 319
4 Conclusion ... 321
References ... 321

Mid-infrared Quantum Cascade Lasers
J. Cockburn ... 323
1 Introduction ... 323
1.1 General QCL Concepts and Current Status ... 323
2 QCL Active Region Design ... 326
2.1 Intersubband Population Inversion ... 327
2.2 Two-well and Three-well Active Regions ... 328
2.3 Double Phonon Resonance Active Region ... 330

2.4 Bound to Continuum Active Region 330
2.5 Comparison of InP-based and GaAs-based QCLs 333
3 Quantum Cascade Lasers for 3-5µm Operation 336
3.1 Strain Compensated InGaAs-AlInAs-InP 338
3.2 InGaAs-AlAsSb-InP 340
3.3 InAs-AlSb 344
4 Quantum Cascade Lasers Grown by Metal-Organic Vapour Phase Epitaxy 347
5 Conclusions 351
Acknowledgements 351
References 351

Part III LEDs and Detectors

Mid-infrared Electroluminescence in LEDs Based on InAs and Related Alloys
A. Krier, X.L. Huang and V.V. Sherstnev 359
1 Introduction 359
1.1 Background 360
2 Limitations to LED Performance 362
2.1 Internal Quantum Efficiency 362
2.2 Purification Using Rare Earth Gettering and Pb Neutral Solvent Epitaxy 363
2.3 Purification of Epitaxial InAs Using Gd Gettering 364
2.4 Fabrication of InAsSb LEDs at 4.6µm 366
2.5 InAs LEDs for Methane Detection at 3.3µm 368
2.6 Neutral Solvent Epitaxy 370
3 High-pressure Measurements 374
4 Optical Extraction 377
5 Comparison of Devices 379
6 InAsSb Quantum Dot Light Emitting Diodes Grown by Liquid Phase Epitaxy 381
7 Superluminescence and Ring Lasers 386
8 Conclusion 389
Acknowledgements 390
References 390

LED-Photodiode Opto-pairs
B.A. Matveev 395
1 Introduction 395
2 Device Configuration and Fabrication 396
2.1 Substrate and Buffer Layers 397
2.2 Active Layer Properties 400
2.2.1 Layer Doping 400
2.2.2 Layer Thickness 401
2.2.3 Active Layer Mesa Diameter 402
2.3 Contacts 403
3 Choice of the Operating Mode 408

4 Out-coupling of Radiation in Mid-IR Devices 415
5 The Use of Diode Opto-pairs for Chemical Sensing 419
6 Summary 424
Acknowledgements 424
References 425

QWIP Detectors for the MWIR
S. Haywood, K.T. Lai and M. Missous 429
1 Introduction 429
1.1 QWIPs *vs* Interband Detectors 429
2 MWIR Transitions in QWs 431
2.1 Square QW Structures 431
2.1.1 (In)GaAs/AlGaAs on GaAs 432
2.1.2 InGaAs/AlInAs on InP 433
2.2 Double-barrier QWs 433
2.2.1 DBQWs on GaAs Substrates 434
2.2.2 DBQWs on InP Substrates 435
3 Strain-balanced QWIPs for High-temperature Operation 436
3.1 Materials Growth and Characterisation 436
3.2 Stoichiometric Growth Conditions 437
3.3 Structural and Electrical Properties 438
3.4 Modelling the Intersubband Transitions 441
3.5 Measured Transitions 442
3.6 Predicted Temperature Performance 443
4 Enhanced Detector Performance and Functionality 444
4.1 Asymmetric Wells 445
4.1.1 Normal Incidence Absorption 445
4.1.2 Voltage Tuning via the Stark Effect 445
4.2 Photovoltaic Operation 446
4.3 High-speed Operation 447
5 Conclusions 447
Acknowledgements 448
References 448

Negative Luminescence
T. Ashley and G.R. Nash 453
1 Introduction 453
2 Negative Luminescent Parameters 455
2.1 Key Parameters 455
2.2 Optical Concentrators 456
3 The Magnetoconcentration Effect 458
3.1 Principle of Operation 458
3.2 Indium Antimonide and Cadmium Mercury Telluride Devices 459
3.3 Other Material Systems 461
3.4 Conclusions 462
4 Photodiodes 462
4.1 Principle of Operation 462
4.2 Indium Antimonide Devices 464

4.3 Cadmium Mercury Telluride Devices 467
4.4 Other Device Structures and Material Systems 471
4.5 Conclusions 472
5 Applications of Negative Luminescence 473
5.1 IR Sources for Gas Sensing 473
5.1.1 Temperature Stabilisation 474
5.1.2 Efficient Long Wavelength IR Sources 474
5.2 Dynamic Infrared Scene Projection 476
5.3 Radiative Cooling 477
5.4 Uncooled IR Radiation Shields for BLIP IR Detectors 478
5.5 Radiometric Reference Planes for Thermal Imagers 479
5.5.1 DC Restoration in Scanned Thermal Imagers 479
5.5.2 Non-uniformity Correction in Staring Arrays 481
5.6 Camouflage 482
6 Summary 482
References 483

Mid-infrared Quantum Dot Photodetectors
P. Bhattacharya, A. D. Stiff-Roberts and S. Chakrabarti 487
1 Introduction 487
1.1 Quantum Dot Photodetectors in the Mid-Infrared 487
1.2 Comparison of HgCdTe Photodiodes and InAs/GaAs Quantum Dot Infrared Photodetectors 490
2 Infrared Detection with Quantum Dots 491
2.1 Bound-state Energy Levels in QDs 492
2.2 Density of States and Carrier Distribution in QDs 494
2.3 Intraband Absorption in QDs 495
2.4 Phonon Bottleneck and Effective Carrier Lifetime in QDs 497
3 Self-assembly of Quantum Dots by the Stranski-Krastanow Growth Mode 498
4 Characterization of Mid-infrared Quantum Dot Photodetectors 500
4.1 QDIP Heterostructure Designs 500
4.2 Dark Current 501
4.3 Mid-infrared Spectral Response in QDIPs 502
4.4 State-of-the-art Performance in MIR QDIPs at High Operating Temperatures 503
5 Conclusions 506
Appendix. Infrared Photodetector Figures of Merit 507
A1 Photocurrent 507
A2 Dark Current 507
A3 Noise Current 507
A4 Normalized Spectral Response and Peak Responsivity 508
A5 Peak Specific Detectivity 508
A6 Photoconductive Gain and Quantum Efficiency 509
Acknowledgements 510
References 510

Quantum Photovoltaic Devices Based on Antimony Compound Semiconductors
Y. Wei, A. Gin and M. Razeghi 515
1 Introduction 515
2 Theoretical Modeling 516
2.1 The Type-II InAs/GaSb Superlattice 516
2.2 Nanopillar Structures 521
3 Material Growth and Characterization 526
4 Photodiodes with Cut-off Wavelength ~8 μm 531
5 Type II Focal Plane Arrays 533
6 Nanopillar Fabrication 535
6.1 Electron Beam Lithography 535
6.2 Nanopillar Device Fabrication Process 535
6.3 Nanopillars in GaSb Material 537
6.4 Nanopillars in InAs/GaSb Superlattice Material 538
6.5 Reactive Ion Etching Using BCl_3:Ar and CH_4:H_2:Ar 538
6.6 Reactive Ion Etching Using Cyclic CH_4:H_2:Ar / O_2 539
6.7 Device Fabrication 540
6.7.1 Polyimide Planarization 540
6.7.2 Polyimide Etchback 540
6.7.3 Top Contact Deposition 541
6.7.4 Dark Current Measurements 542
7 Conclusion 543
Acknowledgement 543
Appendix. Superlattice Hamiltonian Matrix 543
References 545

High-speed Avalanche Photodiodes for the 2-5 μm Spectral Range
M.P. Mikhailova and I.A.Andreev 547
1 Introduction 547
2 Impact Ionization in III-V Semiconductors 548
2.1 Interband Ionization in Semiconductors 548
2.2 Threshold Energy of Impact Ionization 549
2.3 The Dependence of the Ionization Coefficients on the Electric Field 551
2.4 Two-valley Model 552
2.5 Inter-relationship Between Multiplication Coefficients and Ionization Coefficients of Electrons and Holes 553
2.6 Noise and Response Speed of APDs 554
3 Ionization Coefficients in III-V Semiconductors and their Alloys 557
3.1 Electron Impact Ionization 557
3.1.1 Anisotropy of Ionization Coefficients in Multi-valley Semiconductors of GaAs, InP Type 557
3.1.2 Anisotropy of Ionization Coefficients in Multi-valley Semiconductors of GaSb and their Alloys 558
3.2 Electron Impact Ionization in Semiconductors of InAs, InSb Type 560
3.3 Hole Impact Ionization 561
3.4 Hole Impact Ionization in Semiconductors with Band Gap "Resonance", E_g=Δ(InAs, GaSb) 562
4 Avalanche Photodiodes for the 2-2.5 μm Spectral Range 564

4.1 Experimental Ionization Coefficients in Solid Solutions Based on GaSb 565
4.1.1 Ionization Coefficients in GaSb, GaAlSb and GaAlAsSb 565
4.1.2 Experimental Ionization Coefficients in GaInAsSb 567
4.2 Dark Current in APDs with a 'Resonant' GaAl(As)Sb Composition 570
4.3 GaInAsSb/GaAlAsSb Avalance Photodiode with Separate Absorption and Multiplication Region (SAM APD) 572
4.4 Noise and Response Speed of GaInAsSb/GaAlAsSb APDs 576
5 Avalanche Photodiodes for the 3-5 μm Spectral Range 581
5.1 Experimental Investigation of Ionization Coefficients in InAs, InGaAs and InAsSb 582
5.2 Noise and Response Speed of Long Wavelength APDs 585
6 Methods of Separating Ionization Coefficients Using Quantum Structures 586
7 Conclusion 588
References 589

Part IV Applications

Infrared Methods for Gas Detection
J.G. Crowder, S.D. Smith, A. Vass and J. Keddie 595
1 Introduction 595
2 Gas Absorption Spectra 595
3 Methods of Gas Detection 597
4 Infrared Sources for Gas Detection 598
4.1 Thermal Sources 599
4.2 Semiconductor Sources 600
5 Infrared Detectors for Gas Detection 601
5.1 Optical Immersion 602
6 Design of Optical and Gas Sampling Systems 604
6.1 A Non-imaging Gas Sensor with Thermal Source and Detector 606
6.2 A Long-path, Imaging Gas Sensor with Semiconductor Source and Detector 608
7 Laser Techniques 609
8 Conclusions 610
References 610

Mid-infrared Biomedical Applications
I.K. Ilev and R.W. Waynant 615
1 Introduction 615
2 Mid-IR Biophotonics 616
2.1 Biophotonics 616
2.2 Fundamentals of Mid-IR Biophotonics Applications 617
2.3 Mid-IR Biophotonics Delivery Systems 619
2.3.1 Basic Mid-IR Biophotonics Delivery System 619
2.3.2 Mid-IR Biomedical Lasers 620
2.3.3 Mid-IR Incoherent Light Sources 621

2.3.4 Mid-IR Bio-medical Delivery Fibers ... 622
2.3.5 All-Hollow-Waveguide Mid-IR Laser Delivery System ... 624
3 Mid-IR Bio-photonics Applications ... 625
3.1 Mid-IR Laser Surgery and Tissue Ablation ... 625
3.2 Mid-IR Bioimaging ... 627
3.3 Mid-IR Spectroscopy and Biosensors in Biomedicine ... 628
3.3.1 Non-invasive Blood Glucose Monitoring ... 628
3.3.2 Breath Analysis for Medical Applications ... 629
3.3.3 Summary of Mid-IR Spectroscopy and Biosensors ... 630
4 Conclusion ... 631
References ... 631

Development of Infrared Countermeasure Technology and Systems
D.H. Titterton ... 635
1 Introduction ... 635
2 Historical Development ... 637
3 Propagation and Atmospheric Windows ... 639
3.1 Humidity ... 640
3.2 Haze, Fog, Cloud and Rain ... 641
3.3 Turbulence ... 641
3.4 Extinction ... 643
3.5 Thermal Blooming ... 643
3.6 Ionisation ... 643
3.7 Wakes and Plumes ... 644
3.8 Aero-optical Effects ... 644
4 Defeat Mechanisms ... 645
4.1 Denial ... 645
4.1.1 Smokes and Obscurants ... 646
4.1.2 Pyrotechnic Smoke ... 647
4.1.3 Rapid-bloom Obscurant ... 647
4.1.4 Large-area Smoke Screening Systems ... 647
4.2 Deception ... 647
4.2.1 Expendables (Flares) ... 649
4.2.2 On-board Techniques (Jammers) ... 651
4.3 Dazzle ... 653
4.4 Damage ... 654
4.4.1 In-band Damage Route ... 655
4.4.2 Out-of-band Damage ... 657
4.5 Destruction ... 659
5 The Evolution of IR-jammer Systems ... 661
5.1 Incoherent IR Sources ... 662
5.2 Arc Lamps ... 664
5.3 Coherent Sources ... 666
5.4 DIRCM Systems ... 667
6 Characteristics of Laser-based Jammers ... 669
7 Future Developments of Laser-based Systems ... 670
References ... 671

Survey of Thermophotovoltaic (TPV) Devices
M.G. Mauk 673
1 Introduction and Overview 673
2 TPV Literature and Other Sources of Information 679
3 Basic Operation of TPV Cells 680
4 TPV System Thermodynamic Limits and Modeling 684
5 TPV Cell Modeling 689
5.1 TPV Device Structure and Delineation 690
5.2 Minority Carrier Recombination and Limits to Open-circuit Voltage 692
6 Survey of TPV Materials and Devices 697
6.1 Silicon, Germanium, and Si-Ge Alloy TPV Cells 697
6.2 GaSb TPV Cells 699
6.3 TPV Cells Based on Bulk III-V Alloy Ternary Crystals 701
6.4 InGaAs and InAsP/InP TPV Cells and MIMs 703
6.5 Low-bandgap (< 0.5 eV) InAs and InAsSbP TPV Cells 706
6.6 TPV Cells Based on Other Semiconductors 710
6.7 Quantum Well TPV Cells 710
6.8 Isolation for Series Interconnection, Integrated Reflectors, and Wafer-bonding for TPV Cells 713
6.9 Tandem TPV Cells 714
6.10 Micron-gap TPV Cells 717
6.11 Spectral Control, Microstructured Emitters, Filters and Photonic Crystals 720
6.12 Thermophotonics 720
7 Concluding Remarks and Outlook for TPV Technology 721
Appendix. Modeling of InGaAsSb TPV Cells 723
A1 Materials Properties 723
A2 TPV Device Model 728
References 731

Index 739

Part I

Materials and Device Design Considerations

Theory of Mid-wavelength Infrared Laser Active Regions: Intrinsic Properties and Design Strategies

J. T. Olesberg[1,2] and M. E. Flatté[1,3]

[1] Optical Science and Technology Center, The University of Iowa, Iowa City, IA 52242, USA

[2] Department of Chemistry, The University of Iowa, Iowa City, IA 52242, USA
jonathon-olesberg@uiowa.edu

[3] Department of Physics, The University of Iowa, Iowa City, IA 52242, USA
michael-flatte@uiowa.edu

1 Challenges and Opportunities in MWIR Laser Active Region Design

Ongoing research efforts to expand the spectral range of semiconductor laser diodes follow three principal tracks at present: (1) improving operation in the blue-green portion of the visible spectrum and extending the short-wavelength limit into the ultraviolet, largely utilizing the group III-nitride material system [1]; (2) developing high-power sources for photon energies below the typical optical phonon energy of the semiconductor material (usually corresponding to optical wavelengths $\lambda > 40\mu$m) [2]; and (3) achieving room-temperature, high-power, continuous-wave operation of semiconductor lasers in the 3–30 μm range. These frontiers reflect the reality that semiconductor laser diodes operate well in the wavelength range 200nm $\lesssim \lambda \lesssim 3\mu$m, which we will refer to as the optimized wavelength range.

What distinguishes semiconductor active regions in the optimized range from those at longer wavelengths? The long-wavelength semiconductor materials that are theoretically expected to perform best are based on complex layered structures in which some of the layers are thinner than 2 nm. Dramatic advances in the growth of atomically-sharp interfaces between these layers provide confidence that grown structures are very similar to designed structures. High-quality native substrates for growth of these structures are readily available, although they may be more expensive than native substrates in the optimized wavelength range (*e.g.* GaAs and InP). The most significant materials challenges now concern the attempt to exert even greater control of segregation and cross-incorporation at heterointerfaces, particularly in the short-period superlattices most desirable for long-wavelength devices.

Progress in the control of material growth has in some areas exceeded progress in understanding the fundamental optical and carrier dynamics in these materials. Thus advances in the fundamental physical understanding of long-wavelength materials are likely to produce significant improvements in device predictability, performance, and reliability. Ideally, identifiable differences in these fundamental device physics

properties will lead to new categories of "designer materials" with superior properties compared to more traditional semiconductor laser active region designs.

The 3–30 μm spectral region is an excellent range within which to put device design and theory concepts to the test, for there are a wide variety of laser diode active region options. Two qualitatively different approaches to active region design have been pursued with success: the gain can be obtained either from transitions between the conduction and valence band of the active region (interband transitions), or from transitions between conduction subbands or between valence subbands (intersubband transitions). In addition to this interband/intersubband distinction, the interband active regions themselves come in a multitude of options. The highest-performing interband active regions are based on quantum well or superlattice heterostructures, and can be differentiated based on whether the conduction electrons and valence holes preferentially reside before recombination in the same material (type-I heterostructures) or in different materials (type-II heterostructures). The breadth of competing structures, and the importance of fundamental device physics in comparing the ideal and actual performance of these structures, makes this area fertile ground for theoretical device physicists.

1.1 Intrinsic Material Properties Dominate Device Performance

Several intrinsic material factors conspire to make the operation of laser diodes in the 3–30 μm wavelength range significantly more difficult than at near-infrared wavelengths. The best known of these factors is Auger recombination, where nonequilibrium electron-hole pairs recombine and transfer the energy and momentum to another carrier rather than emitting a photon. The primary consequence of Auger recombination is that, left alone, the nonequilibrium carrier density would be rapidly depleted, and large drive currents are required to replenish the carrier density in an operating device. A secondary effect is that the excess energy transferred to the excited carrier heats the carrier distribution.

The intrinsic Auger recombination process is very sensitive to momentum and energy conservation conditions [3]. Approximate analytic forms appropriate for bulk semiconductors with $m_v \gg m_c$ indicate that the recombination probability depends exponentially on the fundamental energy gap, according to expressions of the form $\exp(-AE_g/k_BT)$, where A varies from 0 to 0.5 in a manner that depends on the particular Auger process. Thus Auger recombination is a problem of greater importance for narrow-gap semiconductors than it would be for semiconductors emitting in the optimal wavelength range.

An additional complication is that, because narrow-gap semiconductor materials usually come from elements of higher atomic number, they have large spin-orbit interactions. In the 3–5 μm range, many of the materials have spin-orbit splittings comparable to the band gap. The near-resonance of the spin-orbit splitting with the energy gap provides an additional channel for Auger recombination. In a material such as InAs, where the two energies are roughly the same, the Auger recombination coefficient is $C = 1.1 \times 10^{-26}$ cm^6/s, which is more than two orders of magnitude larger than for 1.55 μm InGaAsP, where Auger recombination was a significant chal-

lenge for many years. For InGaAsP devices, the problem of Auger recombination has been largely overcome through the use of strained quantum wells. In this chapter, we will explore similar, but more extreme, approaches in materials design to deal with the more serious problem of Auger recombination in the 3–30 μm wavelength range.

Another consequence of a decreasing band gap and increasing spin-orbit splitting is the enhanced impact of intervalence band absorption in the mid-infrared. Intervalence band transitions between the top heavy hole band and the light hole or spin-orbit bands occur at energies comparable to the fundamental gap, which can lead to a significant internal optical loss in a laser diode, both from doped clads and in the active region itself. As with Auger recombination, intervalence band absorption has the unfortunate side-effect of heating the carrier distribution. When using heterostructure designs to optimize device performance, considerable care needs to be taken to minimize this undesirable optical loss mechanism.

As electron-hole pairs are injected into the material, the nonequilibrium Fermi energies in the conduction and valence band adjust according to the conduction and valence band densities of states, which for a bulk material depend on the effective mass according to $m^{(3/2)}$. Although the valence hole mass is relatively insensitive to the fundamental gap, E_g, the conduction electron mass depends sensitively on the energy gap in a form that approaches $m_c \propto E_g$ for $E_g < 0.5$ eV [4]. Therefore, the conduction band density of states decreases as the band gap decreases, which leads to a large imbalance in the magnitudes of the valence and conduction band densities of states. A significant imbalance in the densities of states has a negative impact on nearly every facet of active region performance [5,6]. The joint density of states and the differential gain are diminished, the linewidth enhancement factor increases, and problems with Auger recombination and intervalence band absorption are amplified.

1.2 Opportunities Provided by Heterostructure Design

Many of the problems impeding mid-infrared device development are *intrinsic* material issues. Solutions to these problems will come from consideration of the fundamental physics of the materials and optimization of the material properties that influence semiconductor laser diode performance. Layered arrangements of different materials (heterostructures) can have widely different material properties depending on layer thickness, layer composition, and configuration of the layers — properties that are predictable from theoretical calculations. For example, the valence band density of states can be reduced through the introduction of strain and quantum confinement. In the InAs/GaSb/AlSb material system there are three binary semiconductors all nearly lattice matched to each other, and both InAs and GaSb are available as native substrates for growth. The introduction of In into GaSb, Sb into InAs or As into AlSb provides additional opportunities for strain engineering.

The use of heterostructures generates new challenges in addition to opportunities. For example, the thermal conductivity of semiconductor superlattices may not be as high as for bulk materials [7,8]. Materials employing large valence band offsets might significantly reduce the density of states imbalance, but also impede vertical hole mobility. Furthermore, there is often a difference between what is designed and

the precise structure grown; thus, device designs must be robust enough to accommodate slight variations in growth.

The treatment in this chapter will focus on the key material properties that influence semiconductor laser diode performance and strategies for optimizing their characteristics. Section 2 shows that it is possible to identify the ways that the active region's material properties impact device performance in a device- and geometry-independent manner. This process identifies two figures of merit that can be used as a guide for device optimization and to compare different material systems. Sections 3–5 consider strategies for improving the performance of active region materials using these figures of merit as a guide. In order to provide a concise account of the subject, the presentation in these sections is limited to issues related to interband diodes, and the principal focus will be on emission wavelengths shorter than 5 μm. Section 3 examines performance improvements from modifications to the band edge density of states, using simple models and examples from type-I, 2–3 μm diode development, followed by examples from type-II, 3–5 μm diode development. Section 4 explores the additional opportunities provided by type-II materials through modifying the electronic structure deep in the valence band in order to eliminate final states for intervalence band absorption and Auger recombination. Section 5 presents the best designed structure we have been able to create.

As all of these design strategies rely on accurate methods for calculating the electronic and optical properties of heterostructures, these methods will be summarized in Sect. 6. We conclude with open questions and additional remarks in Sect. 7.

2 Metrics for Material Comparison and Device Design

The operating characteristics of present day mid-infrared laser diodes at wavelengths greater than 3 μm lag far behind near-infrared laser diodes. Even when mid-infrared diodes operate at room temperature, their threshold current density is excessively high compared to near-infrared devices. This discrepancy has persisted for many years despite intensive effort.

Threshold current densities of early GaAs-based near-infrared diodes were also exceedingly high: $\sim$10 kA/cm^2 at 77 K [9–11]. Only shortly before there had been considerable disagreement over the feasibility of laser operation using a semiconductor gain medium [12]. After the first demonstration of lasing at 77 K, it was seven years before the first room temperature, "low-threshold" (*i.e.*, 10 kA/cm^2) diodes were available [13, 14]. The years since 1969, though, have witnessed dramatic reductions in threshold current densities and corresponding improvements in device efficiency. Some of these improvements are a result of advances in material quality, but much of the progress has come about through optimizations in the design of the active region. A major design contribution to progress has come through shrinking the volume of the active region; initial homojunction diodes gave way to single and then double heterostructures, which allowed smaller, more defined inversion regions. Double heterostructures then were made thinner and thinner, necessitating the introduction of separate confinement regions that allowed decoupling of the waveguiding

function of the active region from the gain function. Active regions were then further thinned until they became quantum wells. The advent of quantum wells also allowed the incorporation of strain, which effectively reduces the active region density of states still more.

How does one determine the optimal parameters for an active region (*e.g.*, its thickness) and compare alternative material strategies? Experimental trial and error is an indispensable part of the process, but it is important to have a theoretical metric for understanding the implications of choices in the optimization process. This is especially important in the mid-infrared wavelength range where laser diodes have been grown using a variety of active region material systems. For example, there are type-I interband lasers, such as InAsSb/AlAsSb, InAsSb/InAsP, or InSb/AlInSb, type-II broken gap lasers, either in traditional or cascade configuration, and intersubband cascade lasers.

There are a variety of quantities used in the literature to compare the expected performance of active region materials. Most notable among them are the Auger coefficient, the differential gain, the active region recombination current density, and the magnitude of intersubband absorption at the lasing energy. Each of these individually, however, tells only part of the story.

In this section we will describe two figures of merit that follow directly from minimizing the applied electrical power necessary to operate a diode. These figures of merit properly integrate all of the above parameters in a device independent manner. We find that each active region material has a preferred density where it operates most efficiently. This optimal operating point can be used to choose the optimal active region thickness once the laser cavity parameters are known.

There are important aspects of active region material design that are not included in the optimizations discussed in this section. These include the linewidth enhancement factor, the vertical mobility of carriers in the active region, and the thermal conductivity of the active region. Many of the optimizations that reduce threshold current densities also improve the linewidth enhancement factor [15]. Vertical transport, however, is an example of a relevant property that tends to become less optimal when the two figures of merit are optimized. Fortunately, optimizing the figures of merit for an active region in a well-designed cavity can minimize the importance of vertical transport for the laser diode, as only a very thin active region will be required. Little is known, unfortunately, about the thermal conductivity of these optimized active region designs. Perhaps future work will be able to incorporate these additional properties into suitably generalized figures of merit.

The derivation of the figures of merit is a generalization of insights described by Thompson in an early article [16] and his text on laser devices [17]. Similar ideas were also developed in a 1985 article by McIlroy [18] and by Coldren and Corzine in Appendix 17 of their text on laser diodes [19]. Thompson and McIlroy both use the ratio of gain to current density as a tool for optimizing device designs (*e.g.*, choice of active region thickness or cavity length). In the language of Thompson, optimal performance is obtained when the gain to current characteristics of the active region are properly matched to the "optical loading" of the cavity. Coldren, in addition to using the ratio that he calls the "current to gain factor" to optimize the laser cavity

design, also uses the ratio as a tool for comparing the relative performance of three near-infrared active region materials.

There are thus two primary uses of figures of merit for active region materials. The first is to design properly the cavity of a laser diode around the particular optimal operating conditions of the active region in order to minimize threshold current densities or maximize the device slope efficiency. This will be addressed in this section after a derivation of the figures of merit. The second use of the figures of merit is to compare the efficiency of alternative material systems and gauge improvements in material performance that flow from various optimization strategies. This will be the focus of later sections where we address material optimizations. Both of these uses of the figures of merit are a consequence of the fact that the optimal performance of a material can be quantified in terms that are largely independent of the geometry of the device in which the material is used.

Along the way, we will make three extensions to the approach originally presented by Thompson and McIlroy. First, we will generalize the quantity being minimized to be the applied ohmic power density rather than simply the applied current density. This allows us to compare conventional single-stage designs with multistage, cascaded active regions [20]. Second, operation above threshold is included in our analysis, which exposes a second figure of merit relevant for maximizing the slope efficiency of a diode [21,22]. Third, the role of intersubband absorption in the active region is included as a separate term in the analysis, which is essential for the comparison of min-infrared materials where intervalence band absorption can be large [21,22].

We will begin by considering the relatively simple case of traditional, noncascaded active region designs. Afterwards, the results will be generalized to identify the benefits that can be obtained using a cascaded active region.

2.1 Single-stage Active Regions

The primary limitation in present high-temperature mid-infrared laser diode operation is device heating due to the applied current and voltage. The appropriate figures of merit are obtained by minimizing the required ohmic power for a given output optical power. The ohmic power density delivered to a device is

$$P_{2D}^{in} = J_{2D} V_{op} + J_{2D}^2 \rho_s \, , \tag{1}$$

where J_{2D} is the areal current density (in A/cm^2), V_{op} is the operating voltage, and ρ_s is the series resistivity (in Ωcm^2). The operating voltage, V_{op}, is the sum of the ideal diode voltage (*i.e.*, the band gap voltage of the active region) and a series voltage term coming from elsewhere in the device:

$$V_{op} = V_{AR} + V_s \, . \tag{2}$$

In Equation (1), the only variable that depends strongly on the active region, other than the band gap voltage, is the areal current density, J_{2D}. The series voltage and series resistance are normally dominated by non-active-region elements of the diode.

It is important, especially for high-power operation, to optimize these parameters, but they do not involve the active region.

For comparing non-cascade laser diodes with similar band gaps, it can be assumed that operating voltages are comparable, and we seek to minimize the required current density for a given output power. This assumption will be relaxed later when cascaded active regions are considered. The external current, I, required to obtain a given output power, P_0, from a laser diode is

$$I = \frac{J_r A D}{\eta_i} + \frac{1}{\eta_d}\frac{e P_0}{h\nu} \, . \tag{3}$$

The first term describes the external current required to maintain the active region at its threshold density. The second term describes the additional external current that is necessary to produce the desired optical output power. In the threshold term, J_r is the volumetric recombination current density in the active region, A is the area of the stripe (length times width), D is the active region thickness (in the growth direction), and η_i is the internal quantum efficiency (the fraction of carriers injected at the leads that recombine either radiatively or non-radiatively in the active region). The product $A\,D$ is the effective volume of the active region. J_r can be written in terms of the per-carrier recombination rate, r, as

$$J_r = ner \, . \tag{4}$$

In the above-threshold term of Equation (3), η_d is the differential quantum efficiency, e is the fundamental charge, $h\nu$ is the lasing photon energy, n is the conduction electron density, and we assume an undoped active region ($n = p$, where p is the valence hole density).

Dividing Equation (3) by the stripe area gives the expression for the external current density, J_{2D},

$$J_{2D} \equiv \frac{I}{A} = \frac{J_r D}{\eta_i} + \frac{1}{A}\frac{1}{\eta_d}\frac{e P_0}{h\nu} \, . \tag{5}$$

The threshold condition for lasing is that the modal gain be equal to the modal losses in one round trip through the cavity. We write this as

$$\Gamma\,(\gamma - \alpha_a) = \alpha_m + \langle \alpha_w \rangle \ , \tag{6}$$

where Γ is the optical confinement factor (the fraction of the optical mode that overlaps the active region), γ is the fundamental gain coefficient in the active region, α_a is the optical loss coefficient in the active region (due, for example, to intervalence band absorption), α_m is the mirror reflectivity loss distributed over the cavity length, and $\langle \alpha_w \rangle$ is the modal waveguide loss coefficient (losses not associated with the active region or the mirrors, such as those due to free carrier or inter-valenceband absorption in the clads). The reason that α_a is kept separate from the other loss terms is that its impact scales with the confinement factor (*i.e.*, the thickness of the active region). The distinction between active-region and non-active-region losses is normally blurred at shorter wavelengths where non-active-region losses are dominant.

That is not always the case with mid-infrared materials, however, and so it is useful to keep them separate.

Equations (5) and (6) can be joined through the definition of an effective mode width, D_{mode}, which is given by

$$D_{mode} \equiv D/\Gamma\,. \tag{7}$$

D_{mode} is *not* the full-width at half-maximum of the mode, but merely a tool for estimating the confinement factor for small perturbations of active region thickness [16, 17].

Combining Equation (7) with Equations (5) and (6) gives

$$J_{2D} = \frac{D_{mode}\,(\alpha_m + \langle\alpha_w\rangle)}{\eta_i}\left[\frac{J_r}{\gamma - \alpha_a}\right] + \frac{1}{A}\frac{1}{\eta_d}\frac{eP_0}{h\nu}\,. \tag{8}$$

The final step is to write out the differential quantum efficiency, η_d, in terms of the optical losses, since it implicitly contains the active region loss. The differential quantum efficiency is the fraction of carriers injected at the terminals that produce photons that are coupled out of the cavity. The fraction of photons that are coupled out of the cavity to the total number that are produced is the ratio of the mirror "loss" to the total cavity loss (the photons "lost" out the facets are the ones that we are interested in). Thus we can write

$$\eta_d = \eta_i \frac{\alpha_m}{\alpha_m + \langle\alpha_w\rangle + \Gamma\alpha_a}\,. \tag{9}$$

Using Equation (6), we can replace the confinement factor in Equation (9) to obtain

$$\eta_d = \eta_i \left(\frac{\alpha_m}{\alpha_m + \langle\alpha_w\rangle}\right)\left(\frac{\gamma - \alpha_a}{\gamma}\right). \tag{10}$$

Substituting this into Equation (5) gives the final result,

$$\begin{aligned} J_{2D} = {} & \frac{\{D_{mode}\,(\alpha_m + \langle\alpha_w\rangle)\}}{\eta_i}\left[\frac{J_r}{\gamma - \alpha_a}\right] \\ & + \frac{1}{A}\frac{1}{\eta_i}\frac{eP_0}{h\nu}\left\{\frac{\alpha_m + \langle\alpha_w\rangle}{\alpha_m}\right\}\left[\frac{\gamma}{\gamma - \alpha_a}\right]. \end{aligned} \tag{11}$$

By writing the expression for the operating areal current density in the form of Equation (11), we have factored both the threshold and the slope efficiency terms into parts that depend only on the active region (shown in the square braces) or the waveguide (shown in the curly braces). Because the factors are largely independent, they can be optimized separately. The inverse of the active region terms can be used as figures of merit for optimizing the active region or for comparing different active region materials. The figure of merit for minimizing the threshold current density is the maximum of the net material gain per volumetric recombination current density in the active region,

$$\text{Threshold figure of merit} = \left[\frac{\gamma - \alpha_a}{J_r}\right]_{max} \tag{12}$$

whereas the figure of merit for maximizing the slope efficiency is the maximum of the net material gain per fundamental gain,

$$\text{Slope efficiency figure of merit} = \left[\frac{\gamma - \alpha_a}{\gamma}\right]_{max} \tag{13}$$

Both of the active region figures of merit are *material* quantities that do not depend on the geometry of the active region or cavity. Because of this, these figures of merit can be used to compare bulk, quantum well, or superlattice active regions in a device-independent, head-to-head fashion. The threshold figure of merit includes and properly weighs the impact of lasing transition strengths, Auger recombination, intersubband absorption, differential gain, gain saturation, and the balancing and reduction of the valence and conduction band densities of states. The slope efficiency figure of merit measures the impact of intersubband absorption and gain saturation on high-power operation.

These figures of merit are appropriate tools for comparing different active regions, and can be determined theoretically or experimentally. A theoretical approach to obtain the ratio of net material gain to volumetric current density would be to calculate the net gain as a function of carrier density and then either calculate directly or use approximate expressions for carrier recombination due to surface, Shockley-Read-Hall, radiative, and Auger recombination. A direct experimental approach using functional laser diodes is to extract the transparency current density and differential gain by performing Hakki-Paoli measurements on Fabry-Perot devices operating just below threshold. A direct experimental approach using only the active region materials, without a cavity, is to measure the net gain spectrum and density dependent recombination rate using ultrafast pump-probe measurements.

Commonly used simplified expressions for the material gain as a function of carrier density or current density can also be used to obtain simple expressions for these figures of merit. The simplest analytic model expressions for the material gain have the form $\gamma = \gamma_0 \ln(J/J_{tr})$ or $\gamma = \gamma_0 \ln(n/n_{tr})$. In these cases the analytic expressions for the optimal operating point are:

For $\gamma = \gamma_0 \ln(J/J_{tr})$,

$$\gamma_{opt} = \gamma_0 \tag{14}$$

$$J_{opt} = \exp(1) J_{tr} = 2.7 J_{tr} \tag{15}$$

$$\left(\frac{\gamma}{J_r}\right)_{opt} = \exp(-1)\frac{\gamma_0}{J_{tr}}\,. \tag{16}$$

For $\gamma = \gamma_0 \ln(n/n_{tr})$ and $J \approx e\beta n^k$,

$$\gamma_{opt} = \gamma_0/k \tag{17}$$

$$n_{opt} = \exp(1/k) n_{tr} \tag{18}$$

$$\left(\frac{\gamma}{J_r}\right)_{opt} = \exp(-1)\frac{\gamma_0}{e\beta n_{tr}^k}\,, \tag{19}$$

where k is the power of the dominant recombination mechanism (*i.e.*, 1 for surface or Shockley-Read-Hall recombination, 2 for radiative recombination, or 3 for Auger recombination). In Equations (17) and (18), the impact of the dominant non-linear recombination process is to reduce the optimal carrier density. We note the optimal carrier density does not itself depend on the magnitude of the radiative or Auger coefficient. The figure of merit, however, is inversely proportional to the dominant coefficient β.

2.1.1 Threshold Figure of Merit for Bulk GaAs

As an example, the calculated gain spectra of bulk GaAs for a series of nonequilibrium carrier densities are shown in Figure 1. The gain spectra were calculated using a 14-band $\boldsymbol{k} \cdot \boldsymbol{p}$ formalism that has been developed for the study of mid-infrared materials. This formalism, which will be described in greater detail in Section 6, has been used extensively to model the optical and electrical characteristics of near- and mid-infrared quantum wells and superlattices [23–29].

The peak values from each of the spectra in Figure 1 are plotted as a function of carrier density in Figure 2 (dashed line). The value of the peak gain rises from zero at the transparency carrier density ($n_{tr} \sim 1.9\times10^{18}$ cm^{-3}) at first linearly and then sub-linearly. The shape of the gain curve with increasing density depends on the type of material. For example, the roll-off from linearity is often more pronounced in quantum well materials than is evident in a bulk material, as is shown here.

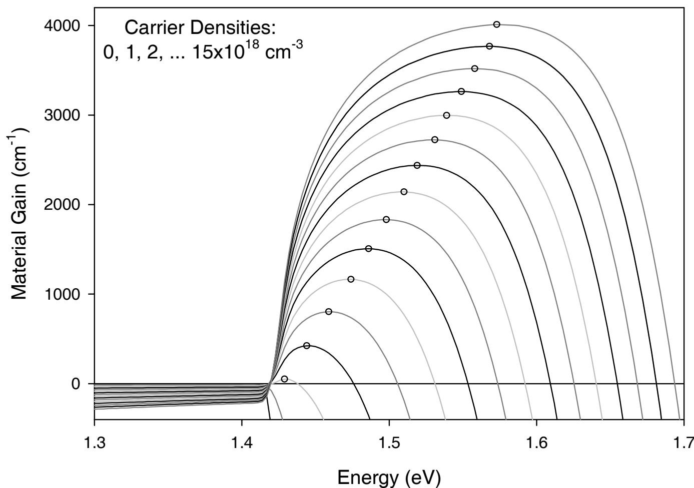


Figure 1. Calculated room temperature gain spectra for bulk GaAs at a series of carrier densities. The broad negative features below 1.4 eV are due to intersubband absorption. The *circle symbols* indicate the peak amplitude and photon energy for each gain spectrum

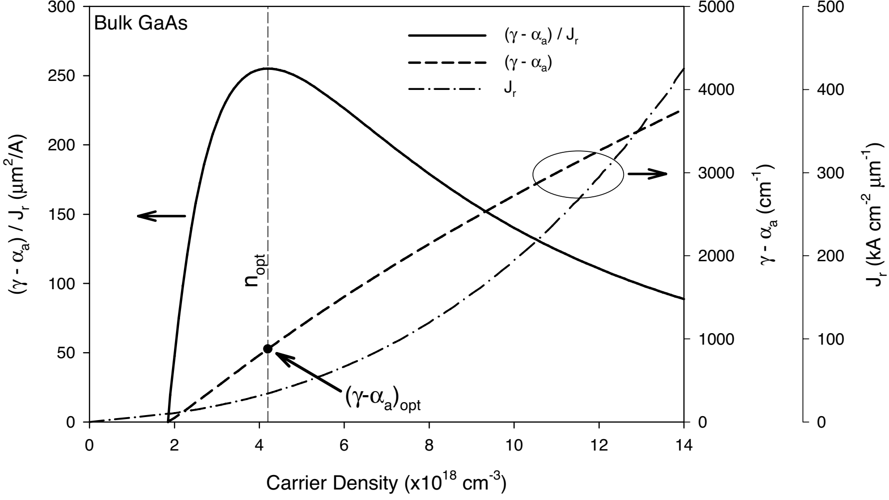


Figure 2. Calculated figure of merit curve (*solid line, left axis*) for bulk GaAs using recombination parameters given in the text. Also shown are the net material gain (*dashed line, right axis*) and volumetric recombination current density (*dash-dotted line, far right axis*). The peak of the figure of merit curve represents the optimal operating point for bulk GaAs, which occurs near a carrier density of $n_{opt} \sim 4{\times}10^{18}$ cm^{-3}. The net material gain at this density is indicated with a *dot*

For high quality GaAs, the recombination rate is determined by surface, radiative, and Auger recombination, according to

$$J_r = q\Big(\frac{2(L+W)}{LW}v_s n + Bn^2 + Cn^3\Big)\,, \tag{20}$$

where L and W are the ridge length and width, v_s is the surface recombination velocity, B is the radiative recombination coefficient, and C is the Auger recombination coefficient. The volumetric recombination current density for a 50 μm x 1 mm ridge waveguide with a surface recombination velocity of $5{\times}10^5$ cm/s, a radiative coefficient of $5{\times}10^{-11}$ cm^3/s, and an Auger coefficient of $5{\times}10^{-30}$ cm^6/s is shown as the dash-dotted line in Figure 2. The curve starts out linear at low densities where it is dominated by surface recombination but becomes superlinear with increasing density. In the case of bulk GaAs, the radiative recombination term dominates the volumetric recombination current density over much of the density range shown, and the curve has a parabolic shape. For most mid-infrared materials, however, the Auger term dominates and the recombination current density increases with the cube of the carrier density.

The ratio of net material gain to volumetric recombination current is shown as the solid line in Figure 2. The curve reaches a maximum value of 255 μm^2/A at a carrier density of $4.2{\times}10^{18}$ cm^{-3} and a net material gain of 880 cm^{-1}. This optimal point represents the conditions under which the material operates with maximum efficiency.

That there is an optimal point on this curve is a consequence of three factors. The first is the presence of a transparency density below which there is no positive gain. The second is the sublinear increase of gain with carrier density once gain saturation begins. The third is the superlinear increase in the volumetric recombination current density due to radiative and Auger recombination. At carrier densities just above the transparency density, a small increase in density produces a larger *fractional* change in gain than in recombination rate. But as the gain curve begins to saturate and nonlinear recombination terms become important, the benefits from increasing the carrier density are overcome by the increasing costs in terms of current density. The carrier density at which the threshold figure of merit curve takes on its maximal value represents the optimal conditions for extracting gain from the material.

2.2 Choice of Active Region Thickness

The key point in this analysis is the fact that the optimal operating point is a *material* quantity that can be determined independently of a particular cavity design. Once the optimal operating point of a material has been determined, the next step is to match the material to a particular cavity, which can be described using a characteristic modal loss and effective mode width. The cavity should be optimized independently based on the desired characteristics of the emitter (*e.g.*, low-threshold or high power operation). Once this is done, the active region material and cavity should be matched by choosing the thickness of the active region.

The goal is to have the device reach threshold at the optimal carrier density. If we define $(\gamma - \alpha_a)_{opt}$ to be the net active region gain at the optimal carrier density, the optimal active region thickness, D_{opt}, is (from Equations 6 and 7),

$$D_{opt} = D_{mode} \frac{\alpha_m + \langle \alpha_w \rangle}{(\gamma - \alpha_a)_{opt}} \,. \tag{21}$$

For a separate confinement structure, where the mode width and non-active-region waveguide losses are roughly independent of the active region thickness, the active region can be optimized independently of the waveguide. For a double-heterostructure the mode width and active region width are closely coupled, and so it is necessary to balance optimization of the active region with optimization of the waveguide. This can be done by iterating the figure of merit calculation with the mode width calculation. The mirror loss of real cavities in current state of the art diodes, however, can be made so small that double heterostructure devices are rarely optimal.

2.2.1 Optimal Thickness of Bulk GaAs

Returning to the example of bulk GaAs, the threshold figure of merit permits an estimate of the optimal thickness for a set of cavity parameters. For a low-loss cavity of length 1 mm and a modal waveguide loss of 3 cm^{-1}, the non-active-region modal loss, $\langle \alpha \rangle = \alpha_m + \langle \alpha_w \rangle$, is 15 cm^{-1}. With an effective mode width of 0.4 μm

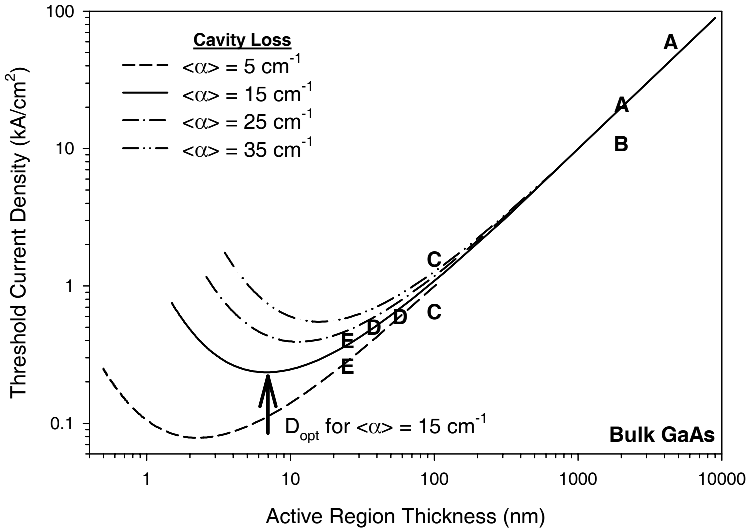


Figure 3. Calculated threshold current density for diodes based on bulk GaAs as a function of active region thickness. Each curve is for a specified non-active-region modal loss: $\langle\alpha\rangle = \alpha_m + \langle\alpha_w\rangle$. The optimal thickness for a cavity with 15 cm^{-1} modal loss is indicated with the *vertical arrow*. The *letters* indicate experimental results from the period 1969–1982. More information for the diode represented by each letter is provided in Table 1. Repetitions of the same letter in the figure represent different diodes described in the same reference

the optimal active region thickness is approximately 7 nm, which is obviously no longer a “bulk” material but rather a quantum well. The threshold current density can then be computed using Equation (11). For the device parameters listed above, the estimated threshold is approximately 240 A/cm^2.

Of course an optimal device design would have an active region thickness chosen so the active region operates at the optimal carrier density. Curves describing the threshold current density as a function of active region thickness, however, can be useful for evaluating the impact of a sub-optimal choice of active region thickness. Figure 3 shows the expected threshold current density calculated for GaAs as a function of active region thickness using Equation (11) and the recombination parameters described previously. The different curves correspond to different levels of non-active-region modal loss, $\langle\alpha\rangle$. A typical 1-mm low-loss cavity would have a modal loss around 15 cm^{-1}. Overlaid on these curves are experimental results from the period 1969–1982. The references corresponding to each letter on in the figure are listed in Table 1. Although the list is not exhaustive, it illustrates the general trend of improved threshold current density with decreasing active region width, which spans two orders of magnitude in threshold current density.

Table 1. Experimental results for early room temperature GaAs diodes. The letters representing each diode are the same as in Figure 3. Abbreviations: DH - Double heterostructure, SCH - Separate confinement heterostructure, GRIN - Graded index, SQW - Single quantum well

	Year	Author	Type of diode	Reference
A	1969	Hayashi	DH	[13]
B	1969	Kressel	DH	[14]
C	1973	Panish	SCH (non quantum-well)	[30]
D	1981	Tsang	GRIN-SCH SQW	[31]
E	1982	Hersee	GRIN-SCH SQW	[32]

2.3 Maximizing the Slope Efficiency

The previous discussion has focused on minimizing the threshold current density of a diode. Alternatively, the focus can be placed on maximizing the slope efficiency. It is common to increase the mirror losses by utilizing a short cavity or an anti-reflection coating on one of the facets in order to increase the differential quantum efficiency, which is proportional to the ratio of the mirror loss to total losses:

$$\eta_d \propto \frac{\alpha_m}{\alpha_m + \langle \alpha_w \rangle} \,. \tag{22}$$

However, Equation (10) shows that there is an additional factor that depends on the active region, which we refer to as the slope efficiency figure of merit of the material. Because the value of this ratio depends on the carrier density in the active region, the thickness of the active region can be chosen such that the diode lases where the ratio $(\gamma - \alpha_a)/\gamma$ is maximized. That there is a maximum in this ratio is a consequence of the fact that the gain, γ, has a minimum (threshold) carrier density before it becomes positive, whereas the loss term, α_a, typically does not. For example, intervalence band absorption contributions to α_a are often linearly dependent on the carrier density.

The slope efficiency curve for GaAs is shown in Figure 4. For GaAs, intervalence band absorption is not significant, so the slope efficiency figure of merit curve is not strongly peaked. However, the slope efficiency near the transparency density is small, and the slope efficiency for diodes with thick active regions is expected to be low.

2.3.1 Maximizing Efficiency for a Specific Optical Power

Often the intention is to minimize the operating current I corresponding to a desired output power P_0, instead of simply minimizing the threshold current density (corresponding to optimizing for low-power emission) or the slope efficiency (corresponding to optimizing for high-power emission). An expression for the required current can be obtained from Equation (11) as

$$I = J_{2D} A \,, \tag{23}$$

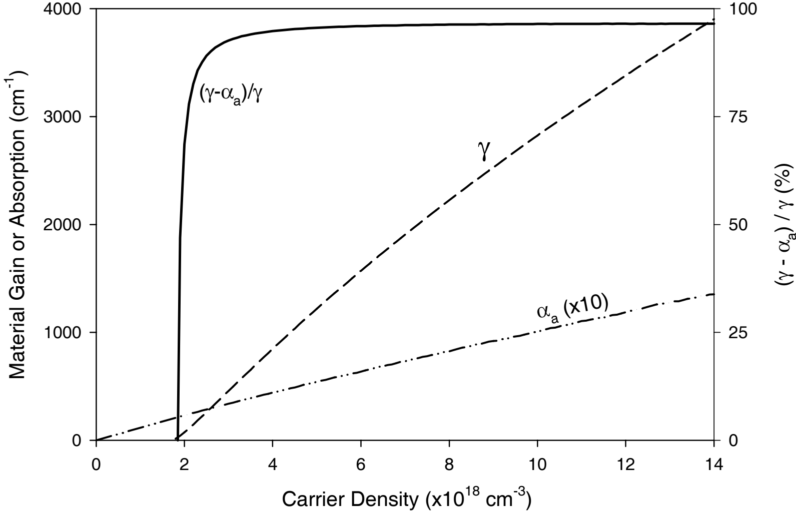


Figure 4. Slope efficiency figure of merit curve for bulk GaAs (*solid line, right axis*). Also shown are the calculated fundamental material gain (*dashed line, left axis*) and the calculated intersubband absorption at the energy of the peak gain (*dash-dot-dot line, left axis*). The intersubband absorption is multiplied by ten for clarity. The slope efficiency figure of merit indicates the impact of intersubband absorption in the active region on the slope efficiency of a diode. If a thicker than optimal active region were to be used such that the carrier density at threshold were less than 3×10^{18} cm^{-3}, the slope efficiency would be adversely affected

where A is the stripe area. It is possible to jointly optimize the cavity parameters (*e.g.*, active region thickness and mirror loss) for maximum efficiency at a given power output, and the optimal values depend on both the threshold and slope efficiency figure of merit curves.

For a cavity with a specific mirror loss, the optimal active region thickness can be determined by minimizing the required current with respect to the threshold carrier density:

$$\frac{\partial I}{\partial n} = 0 \Rightarrow \frac{\partial}{\partial n}\Big(\frac{J}{\gamma - \alpha_a}\Big) + \Big(\frac{e}{h\nu}\frac{P_0}{AD_m}\frac{1}{\alpha_m}\Big)\frac{\partial}{\partial n}\Big(\frac{\gamma}{\gamma - \alpha_a}\Big) = 0\,. \qquad (24)$$

If both the threshold and slope efficiency figures of merit curves are known, the optimal carrier density and gain can be obtained from Equation (24). Then the thickness of the active region is chosen so that threshold is reached when the carrier density of the active region has reached that point.

2.4 Cascaded Active Regions

The preceding section describes how to quantify the degree of optimization of an active region in a non-cascaded geometry; however, the discussion can be generalized

to allow for comparison with cascaded active regions [20]. The key advantage of cascading is that it is possible to reduce the impact of series resistance on the applied power in Equation (1) by increasing the slope efficiency. In a cascaded system, sections of active region are repeated in such a way that carriers traversing the system must go through each stage of the cascade. Conceptually, it is like arranging sections of active region in series rather than in parallel (as is the case with a single, thick active region). Cascading can be obtained by relying on tunneling between stages in an intersubband active region [33] or through the use of an Esaki tunnel junction [34,35] or a type-II semi-metal interface [36, 37] for interband active regions.

Cascading is better than simply connecting multiple laser diodes in series for two reasons. First, the cascaded active regions exist within the same waveguide, so that the emission from the stages in a cascade system is coherent. Second, cascaded systems avoid the extra sets of contacts and clads, which are normally responsible for most of a diode's series resistance. This assumes, of course, that more series resistance is not added by the transitions between stages than is saved by removing the extra clads and contacts.

We can begin by replacing the active region voltage in Equation (2) with

$$V_{AR} = N_{st} V_{st} \,, \tag{25}$$

where N_{st} is the number of stages in the cascade and V_{st} is the voltage drop across a single stage. In the expression for the total current, Equation (3), the active region thickness, D, is now explicitly the thickness of a single stage, D_{st}. If we assume that no carriers are lost during the traversal of the stages (*i.e.*, every carrier passes radiatively or non-radiatively through every stage of the device), the above threshold term in Equation (3) should be divided by the number of stages, giving

$$I = \frac{J_r A D_{st}}{\eta_i} + \frac{1}{N_{st}} \frac{1}{\eta_d} \frac{eP_0}{h\nu} \,, \tag{26}$$

with similar changes in the expression for current density Equation (5),

$$J_{2D} = \frac{J_r D_{st}}{\eta_i} + \frac{1}{N_{st}} \frac{1}{A} \frac{1}{\eta_d} \frac{eP_0}{h\nu} \,. \tag{27}$$

The definition of the effective mode width is still the *total* active region thickness divided by the confinement factor:

$$D_{mode} = \frac{N_{st} D_{st}}{\Gamma} \,. \tag{28}$$

Assembling Equations (26) and (27) with the threshold condition, Equation (6), which remains unchanged, gives

$$\begin{aligned} J_{2D} = {} & \frac{D_{mode}}{N_{st}} \frac{(\alpha_m + \langle\alpha_w\rangle)}{\eta_i} \left[\frac{J_r}{\gamma - \alpha_a} \right] \\ & + \frac{1}{N_{st}} \frac{1}{A} \frac{1}{\eta_i} \frac{eP_0}{h\nu} \frac{\alpha_m + \langle\alpha_w\rangle}{\alpha_m} \left[\frac{\gamma}{\gamma - \alpha_a} \right] \,, \end{aligned} \tag{29}$$

which is a generalization of Equation (11). The present result differs by the factor of $1/N_{st}$ in both terms.

There is one subtlety in this equation related to the definition of the effective mode width. The impact of this can be seen by considering the required current density as the number of stages becomes arbitrarily large: Equation (29) implies that the threshold current density goes to zero, which is unphysical. When considering large numbers of stages, it is better to write Equation (29) in terms of the confinement factor and per-stage active region thickness rather than in terms of the effective mode width:

$$J_{2D} = \frac{D_{st}}{\Gamma} \frac{(\alpha_m + \langle \alpha_w \rangle)}{\eta_i} \left[\frac{J_r}{\gamma - \alpha_a} \right] + \frac{1}{N_{st}} \frac{1}{A} \frac{1}{\eta_i} \frac{eP_0}{h\nu} \frac{\alpha_m + \langle \alpha_w \rangle}{\alpha_m} \left[\frac{\gamma}{\gamma - \alpha_a} \right] . \tag{30}$$

The subtlety is that the effective mode width, which is a useful tool for comparing structures with comparable active region thicknesses, is a weak function of the active region thickness. In the limit of an arbitrarily large number of stages, the ratio D_{mode}/N_{st} from Equation (29) approaches zero; whereas, in Equation (30), the ratio D_{st}/Γ approaches D_{st}, as it should. However, for a reasonable number of stages, the form of Equation (29) demonstrates the dependence on the number of stages more explicitly.

It is important not to end the discussion with the reduced current density because the operating voltage also depends on the number of stages. If we define the minimum current density that can be obtained for a single-stage device as J_{2D}^{single}, the current density required for an optimal cascaded device can be approximated as

$$J_{2D} \approx J_{2D}^{single}/N_{st} . \tag{31}$$

The the expression for the applied power density, Equation (1), can then be written in terms of J_{2D}^{single} as

$$P_{2D}^{in} = \frac{J_{2D}^{single}}{N_{st}} (N_{st} V_{st} + V_s) + \left(\frac{J_{2D}^{single}}{N_{st}} \right)^2 \rho_s \tag{32}$$

$$= J_{2D}^{single} V_{st} + \frac{J_{2D}^{single} V_s}{N_{st}} + \frac{\left(J_{2D}^{single} \right)^2 \rho_s}{N_{st}^2} . \tag{33}$$

Even though the external current applied to the device is decreased by the number of stages, the first term in Equation (33), which describes the power applied to the active region itself, is independent of the number of stages. This is because the decrease in current density is offset by a corresponding increase in the drive voltage. The benefit of cascading comes from reducing the power dissipated by the parasitic effects of series voltage and series resistance. This is especially important for high-power operation, where the series resistance term tends to dominate the applied power expression even for a small series resistance. A point that is sometimes overlooked,

however, is that any active region can in principle be cascaded, so it is important to start with an optimized active region and then cascade it as is necessary to suppress the power dissipated due to series resistance. For cascaded structures the figures of merit presented in the previous section remain the key quantities for measuring the optimization of a material and for comparing different materials.

Since their introduction in 1994, [33] intersubband quantum cascade lasers have come to dominate the long-wavelength infrared portion of the spectrum. Recent efforts have been directed at extending the operating range to wavelengths as short as 3 μm. Although the focus of this chapter is on interband mid-infrared diode lasers, a brief discussion of the relative figure of merit for quantum cascade lasers is in order. Intersubband quantum cascade lasers are based on two independent concepts: the use of a unipolar intersubband gain medium and active region cascading. A unipolar design operates over a very wide wavelength range with far less temperature dependence than conventional interband devices. The gain that can be obtained, however, is small due to short carrier lifetimes limited by phonon scattering.

The relationship between gain and carrier density within intersubband devices is characteristically different from that for interband devices. For intersubband devices the transparency density is nearly zero and the gain increases roughly linearly with density. This implies that the net gain to recombination current density ratio is nearly independent of current, whereas for interband devices there is an optimal carrier density that maximizes the figure of merit. For a constant gain to recombination current density ratio the number of states can be chosen arbitrarily. Frequently, many stages (*e.g.*, 25) are used to minimize the threshold current density, and reduce the power dissipation from series resistance and series voltage. For interband devices, however, the number of stages must be chosen so as to yield a threshold carrier density near the maximum of the figure of merit curve.

The gain produced by an intersubband transition is normally described as

$$\gamma = gJ_{2D} \,, \tag{34}$$

where g is referred to as the gain coefficient and conventionally reported in units of cm/kA. The gain coefficient can be estimated from [38]

$$g = \frac{4\pi e}{\varepsilon_0 n D_{st} \lambda(2\gamma)} \tau_u \left(1 - \frac{\tau_l}{\tau_{ul}}\right) z_{ul}^2 \,, \tag{35}$$

where n is the effective index of refraction of the mode, D_{st} is the thickness of one period of the cascade, (2γ) is the full-width at half maximum of the gain spectrum (normally estimated as the width of the electroluminescence spectrum), τ_u and τ_l are the total lifetimes of the upper and lower states in the lasing transition, τ_{ul} is the scattering time between these states, and z_{ul} is the matrix element between the lasing states. The gain coefficient can also be estimated from the threshold current density if the confinement factor and non-active-region losses are known:

$$g \approx \frac{\alpha_m + \langle \alpha_w \rangle}{\Gamma J_{2D}^{th}} \,. \tag{36}$$

The ratio of material gain to volumetric recombination current density is simply

$$\frac{\gamma - \alpha_a}{J_r} \approx \frac{\gamma}{J_r} = gD_{st}\,. \tag{37}$$

Calculated values for the threshold figure of merit for published intersubband quantum cascade devices are 6–10 $\mu\mathrm{m}^2/A$ in the 3–6 μm wavelength range [39–43] and 10–35 $\mu\mathrm{m}^2/A$ for wavelengths longer than 6 μm. [38,44–52] For interband devices, the threshold figure of merit in the mid-infrared is often much larger than these values, which suggests that interband devices have the potential to operate more efficiently in the 3–5 μm range than quantum cascade devices [20].

3 Band-edge Optimization of Active Region Materials

This section explores modifying the band-edge properties of a semiconductor to optimize its material properties for efficient laser diode operation. The next section will explore manipulation of states deeper in the conduction and valence bands that serve as final states for electronic transitions from the band edge. The primary goal of band-edge optimization is to minimize the carrier density required to achieve the desired amount of gain from the material. The principal way that this is accomplished is through balancing the densities of states of the valence and conduction bands near their band edges. In this section, we discuss why an imbalance in the conduction and valence band densities of states causes problems with efficient laser action, why the problems are much worse in the mid-infrared than the near-infrared, and what can be done to improve the situation.

3.1 Density of States Imbalance

In most direct-gap bulk semiconductors, the mass of the valence band greatly exceeds that of the conduction band; this mass inequality directly controls the density of states imbalance. Unfortunately this imbalance grows as the band gap decreases. Density of states imbalance can be illustrated by considering the electronic structure of InAs, which was the material used for the first mid-infrared diode laser [53]. The band structure of InAs is shown in Figure 5 along with the density of states derived from the band structure and the Fermi occupation functions for the density required to achieve transparency at room temperature (n = 5.5×10^{17} cm^{-3}). The density of states of the valence band exceeds that of the conduction band by approximately two orders of magnitude.

Because of the imbalance of the conduction and valence densities of states, the Fermi levels are shifted upward from the band edges. Almost all of the population inversion comes from the conduction band, which is highly degenerate at the band edge (labeled by “CB” in Figure 5(c)), while the valence band is still in the exponential tail of the Fermi distribution. One effect of this imbalance is a dramatic reduction of the differential gain. This can be understood simply from the slope of the Fermi functions in Figure 5(c). The differential gain will be proportional to the rate of change of the occupation functions with carrier density. For a given movement of the

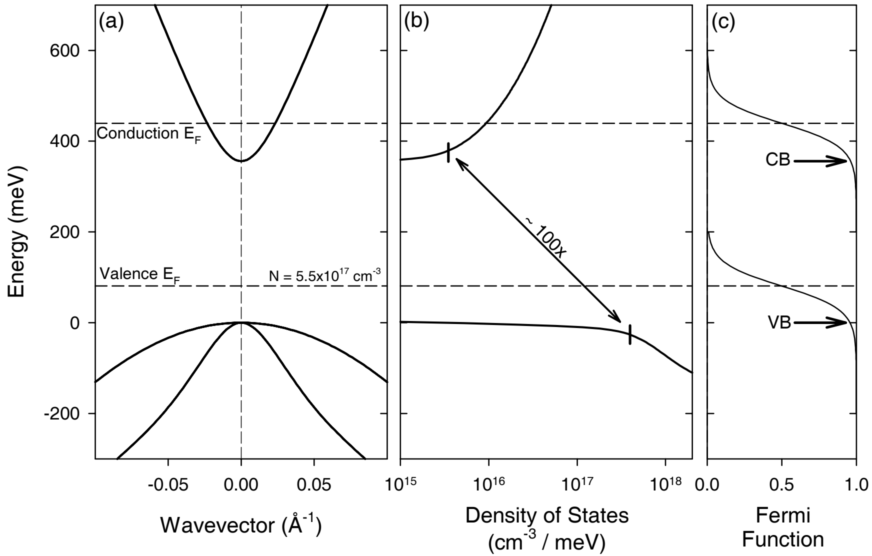


Figure 5. **(a)** Band structure, **(b)** density of states, and **(c)** Fermi occupation functions for bulk InAs at room temperature. The *dashed lines* in the figure are the quasi-Fermi energies of the conduction and valence bands at the transparency density ($n = 5.5\times10^{17}$ cm^{-3})

Fermi levels, the change in occupation function will be largest where where the slope of the Fermi function is large, that is, at the Fermi level. The slope of the occupation function falls off roughly exponentially with distance from the Fermi level. Because the Fermi levels in both the conduction and valence bands are far from the band edges, the differential gain near the band gap will be much smaller than in the case of balanced densities of states. In the ideal case of perfect balancing of the densities of states, the Fermi levels would be at the band edges at transparency, which would provide the maximum possible differential gain.

Imbalance in the densities of states arises from the nature of the electronic structure of zinc-blende materials. The origin of the imbalance can be seen most simply in the $\boldsymbol{k}\cdot\boldsymbol{p}$ picture, where the masses of the different bands arise from interactions with the other bands. The mass of the conduction band is primarily due to interactions with the light and split-off hole states. The nearest states above the conduction band which repel the conduction band from above are typically more than 3 eV away. Thus the conduction mass is light and becomes lighter as the band gap decreases. The masses of the light hole and split-off band are individually due to interactions with the conduction band, so their mass will be slightly heavier than that of the conduction band. The heavy-hole band does not interact with the conduction band or the other hole bands, so its mass will be large. The source of the heavy-hole mass is interactions with the antibonding p-like states several electron volts above the conduction band edge. This is illustrated in Figure 6, which shows the near-zone-center

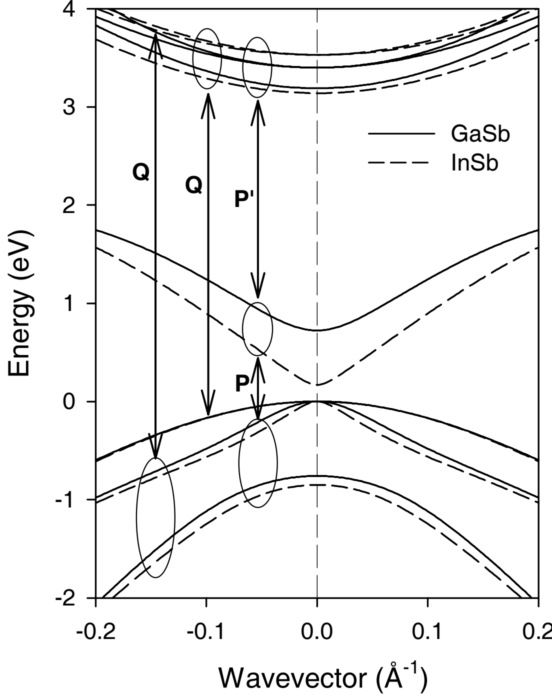


Figure 6. Electronic band structure of bulk GaSb (*solid lines*) and InSb (*dashed lines*) in a 14-band $\boldsymbol{k} \cdot \boldsymbol{p}$ model. The heavy-hole dispersions for the two materials are indistinguishable. The *vertical arrows* show the matrix elements responsible for coupling between states along the [001] direction. InSb has a lighter conduction band mass than GaSb because of the greater proximity of the conduction band to the light states of the valence band. Note that the conduction band ceases to be parabolic in shape at very small wavevectors

electronic structure of GaSb and InSb. The coupling between states is indicated by the matrix elements

$$iP = \langle s^c | p_x | x^v \rangle \text{ and cyclic permutations} \tag{38}$$

$$iP' = \langle s^c | p_x | x^c \rangle \text{ and cyclic permutations} \tag{39}$$

$$iQ = \langle x^v | p_z | y^c \rangle \text{ and cyclic permutations}\,, \tag{40}$$

where the superscript indicates conduction (c) or valence (v) states.

As the conduction band approaches the valence band edge, the interaction between the conduction and light valence band states repels the conduction band to produce a very light mass (approximately three times lighter for InSb than GaSb). In contrast, the heavy-hole mass of the two materials are comparable. Note also that non-parabolicity in the conduction and light-hole bands increases significantly as the band gap shrinks and the conduction band mass becomes lighter.

The mass of the conduction band, m_c, can be approximated in terms of the matrix element coupling the light band-edge states P, the energy gap E_g, and the electron free mass m_0 as

$$\frac{m_0}{m_c} \approx 1 + \frac{2P^2}{m_0 E_g} \,. \tag{41}$$

Because the magnitude of the momentum matrix element, P, is very similar for most zinc-blende semiconductors, ($2P^2/m_0 \approx 20$ eV), the conduction band mass can be written as

$$\frac{m_c}{m_0} \approx \frac{E_g}{20eV} \,. \tag{42}$$

Hence the conduction band mass decreases roughly linearly with the energy gap while the heavy-hole mass remains approximately constant.

Quantifying the imbalance between the valence and conduction band densities of states for a bulk system requires comparing the ratio of the band-edge masses. A simple single value of the mass cannot be assigned, however, to the conduction and valence states in the more complicated quantum systems that will be discussed later due to the presence of multiple valence subbands and significant non-parabolicity. Instead we estimate the ratio of the valence to conduction band densities of states using the deflection of the equilibrium Fermi energy, E_F, from the center of the energy gap. For bulk materials the equilibrium Fermi energy can be approximated as

$$E_F = E_g/2 + k_B T \ln\left[\left(\frac{m_v}{m_c}\right)^{3/2}\right] , \tag{43}$$

which can be inverted to give

$$\frac{\rho_v}{\rho_c} \propto \left(\frac{m_v}{m_c}\right)^{3/2} = \exp\left(\frac{2E_F - E_g}{k_B T}\right) . \tag{44}$$

The imbalances in the densities of states for the longer-wavelength direct gap III-V binary compounds, quantified using this measure, are shown in Figure 7. The equilibrium Fermi energy can be evaluated for any material once the electronic density of states is known. It is a measure of the density of states ratio at and near the band edge, integrated using a weighting function that falls off exponentially away from the band edges. For bulk InAs, which was shown in Figure 5, the ratio is 127, which is consistent with the value estimated from Figure 5(b).

A material need not have a large imbalance in the density of states simply because it is a mid-wavelength infrared gap semiconductor. The large density of states imbalance for these III-V and II-VI materials is driven by the symmetries of the conduction and valence states along with the narrow gap. The electronic structure of lead salt materials, however, differs critically from that of direct-gap III-V or II-VI semiconducting materials. Instead of a direct gap at the Γ point, theoretical calculations and experimental measurements of these materials indicate they have direct gaps at the L points of the Brillouin zone [54, 55]. The masses of the conduction and valence bands are very similar [56], which leads to comparatively long carrier recombination times in these materials [57]. The near equivalence of the masses of

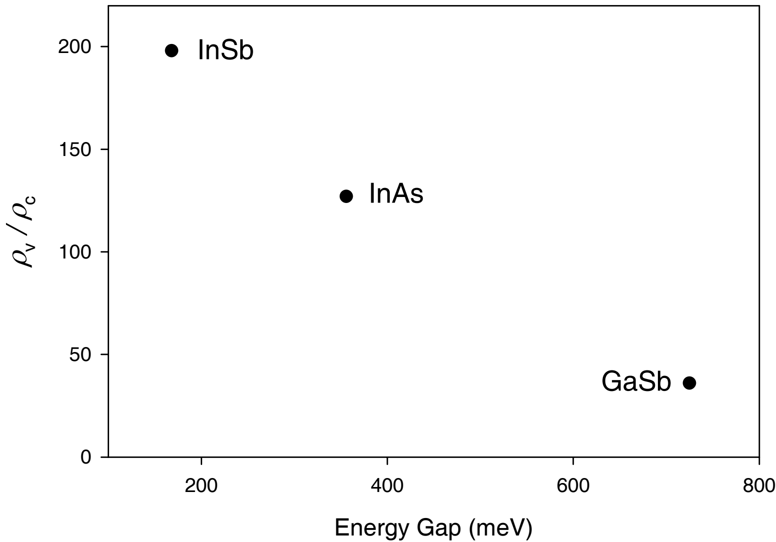


Figure 7. Ratio of the valence to conduction band densities of states as a function of energy gap for longer-wavelength bulk III-V materials. The ratio is determined using Equation (44) and the non-parabolic band structure calculated with a 14-band $\boldsymbol{k} \cdot \boldsymbol{p}$ model

the conduction and valence bands produces a very well balanced density of states, which should produce very impressive figures of merit. These have not yet been calculated, however, as the electronic structure models for lead salt materials [56] differ considerably from those of direct-gap III-V and II-VI semiconductors.

3.2 The Effects of Density of States Imbalance

Imbalance in the densities of states of the valence and conduction bands diminishes the performance of active region materials [5, 6]. Here we explore the origin of this degradation and the improvements that can be obtained by reducing the natural imbalance of a bulk semiconductor. One key strategy is to reduce the mass of the valence band. The valence band-edge properties of the material are the most amenable to modification through the use of strain and quantum confinement [5].

The impact of reducing the density of states imbalance becomes apparent even within a simple two-band model with bulk parabolic bands, and is shown in Figure 8. The model is based on an InAs-like material with a band gap of 350 meV, a conduction band mass of $m_c = 0.025\, m_0$, and a valence band mass of $m_v = 0.40\, m_0$. The purpose of calculating the gain as a function of valence band mass, with all other parameters held constant, is not to model precisely the real characteristics of InAs, but rather to illustrate the relative changes produced by reducing the valence band mass.

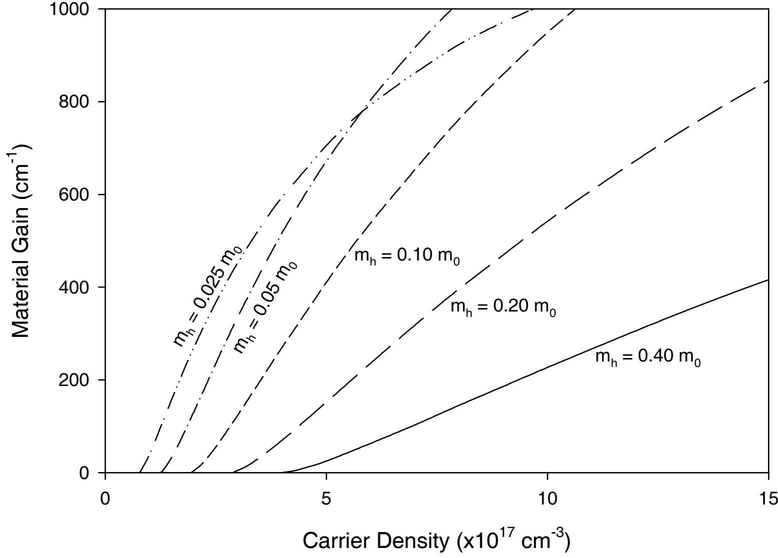


Figure 8. Calculated gain as a function of density for a bulk InAs-like material using a simplified two-band parabolic model. Reducing the valence band mass decreases the transparency density and dramatically increases the differential gain in spite of the fact that the joint density of states is reduced. The conduction band mass is assumed to be a constant 0.025 m_0 for these calculations

In this model, the gain as a function of photon energy, E, is given by

$$\gamma(E) = \frac{\pi e^2 \hbar}{\varepsilon_0 \mu(E) m_0^2 c E} \frac{P^2}{4} \rho_r(E) \Big[f_c(E) - f_v(E) \Big] , \tag{45}$$

where $\mu(E)$ is the index of refraction, P is the bulk momentum matrix element, ρ_r is the reduced density of states, and $f_c(E)$ and $f_v(E)$ are the Fermi occupation functions for the conduction and valence band states coupled by energy E. The occupation functions depend on the carrier density through the conduction and valence band quasi Fermi energies, which are estimated using an approximate form for the Fermi integral [58].

The curves shown in Figure 8 are for a range of valence band masses ranging from the nominal InAs mass $m_v = 0.40\, m_0 = 16\, m_c$ down to the limiting case of equal masses, $m_v = 0.025\, m_0 = 1\, m_c$. As the valence band mass decreases by a factor of 16, the transparency density decreases by approximately a factor of four. What is surprising, however, is that the differential gain increases in spite of the fact that the joint density of states decreases by almost a factor of two. The increase in differential gain is a consequence of obtaining a much larger contribution to the carrier population inversion from the valence band.

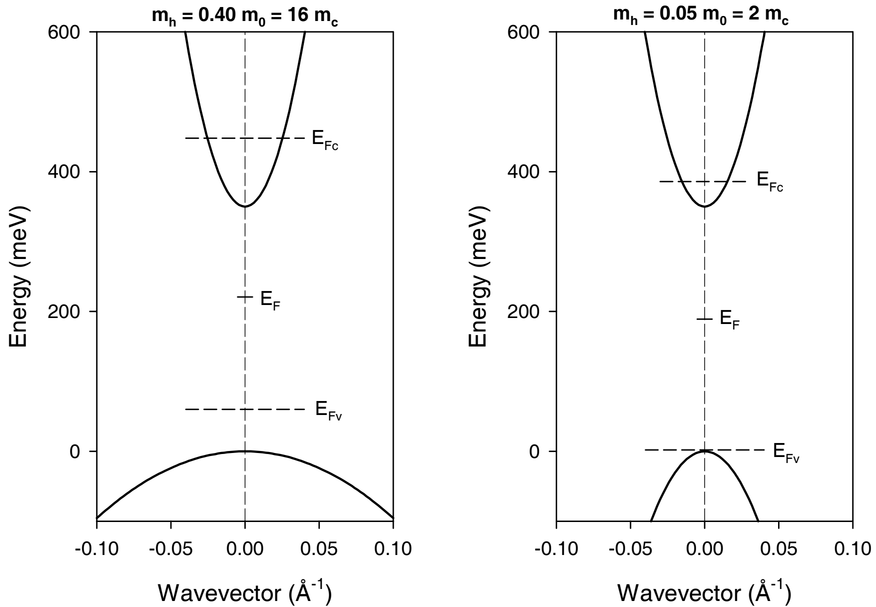


Figure 9. Simplified two-band model of an InAs-like bulk material: (*left*) a material where the valence band mass is 16 times that of the conduction band; (*right*) a material where the valence band mass is twice that of the conduction band. The *dashed lines* are the quasi-Fermi energies for a carrier density 1.5 times the transparency density. The material *on the right* will have a much larger differential gain because of the more balanced contribution from the valence band, which is just becoming degenerate

This is illustrated in Figure 9, which shows the band structure for two cases: one with the nominal valence band mass of InAs, $m_v = 0.40\, m_0 = 16\, m_c$ and another with a lighter valence band mass of $m_v = 0.05\, m_0 = 2\, m_c$. The quasi-Fermi levels corresponding to a carrier density 1.5 times that of the transparency density in each case are shown as the horizontal dashed lines in the figure. In the case of the nominal masses, the valence band Fermi level is still 60 meV away from the valence band edge, while for the reduced valence mass case, the valence band is just becoming degenerate. The increase in differential gain is explained by the fact that Fermi occupation function changes most dramatically at the Fermi energy rather than in the exponential tails. In the case of a lighter valence band mass, this occurs at the valence band edge.

Auger recombination, which becomes exponentially more important as the wavelength increases, is also influenced by density of states imbalance. The presence of a heavy valence band makes the simultaneous conservation of energy and momentum in the Auger process easier to satisfy. This can be seen schematically in Figure 10, which shows an example Auger transition in two bulk materials that are identical except for the mass of the heavy-hole. The transition involves a conduction electron

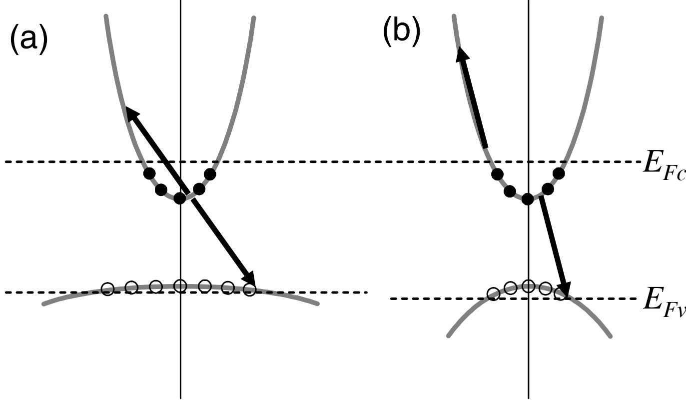


Figure 10. A large heavy-hole mass makes it easier to conserve energy and momentum in an Auger recombination process. In a material with a heavy valence band mass, carriers spread further in k-space. This makes it easier to conserve energy and momentum with a carrier near the band edge in the conduction band

that recombines with a heavy hole, and transfers the excess energy to another conduction electron. In Figure 10(a), where the heavy hole mass is large, the indicated transition can take place with carriers in heavily-occupied states. In Figure 10(b), it is not possible to complete the transition without using at least one carrier that is in the exponential tail of the distribution. Similar behavior can be expected for Auger processes in which the excess energy and momentum of a recombining electron-hole pair are transferred to a hole rather than an electron.

This can also be seen by comparing the activation energies for Auger processes. For example, the activation energy for the process shown in Figure 10 is

$$E_T = \frac{2m_c + m_h}{m_c + m_h} E_g \,. \tag{46}$$

The probability of a transition then varies as

$$P \sim n^2 p\, e^{-(E_T - E_g)/k_B T} \,, \tag{47}$$

or

$$P \sim n^2 p\, e^{-AE_g/k_B T} \,, \tag{48}$$

where

$$A = E_T/E_g - 1 = \frac{m_c}{m_c + m_h} \,. \tag{49}$$

For the two cases considered in Figure 9, $m_h = 16\, m_c$ and $m_h = 2\, m_c$, the A coefficients are $1/17$ and $1/3$, respectively. The reduction in the probability of Auger events then depends on

$$\frac{P_{m_h = 2m_c}}{P_{m_h = 16m_c}} \sim e^{-(1/3 - 1/17)E_g/k_B T} \,. \tag{50}$$

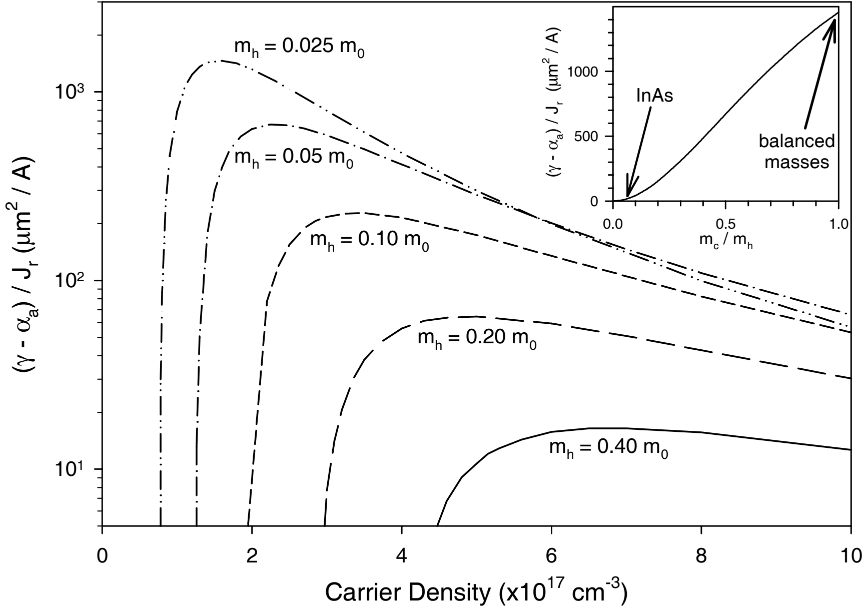


Figure 11. Threshold figure of merit curves for the simplified two-band parabolic model of an InAs-like material with varying valence band mass. The conduction band mass is 0.025 m_0 in all cases. As the valence band mass becomes lighter, the transparency density shifts to lower values and the peak value of the figure of merit curve increases significantly (note the log scale). The *inset* shows the peak value of the figure of merit curves as a function of m_c/m_h

For a material with a band gap of 350 meV at room temperature, the suppression of the Auger rate only from reducing the mass of the valence band is expected to be more than a factor of 40.

These are simple arguments for the impact of density of states imbalance on Auger recombination. Proper treatment of Auger recombination is complicated, and the predictive nature of simple models such as that used for the activation energy (assuming, for instance, Boltzmann statistics and parabolic bands) are limited. The impact of density of states imbalance on Auger recombination, however, is significant, and is seen in the more detailed quantitative calculations described in Section 6. This Auger suppression is independent of, and in addition to, suppression that comes from removing states at resonant energies one band gap below the valence band edge, which will be addressed in Section 4.

The threshold current density figure of merit curve, shown in Figure 11, summarizes the impact of changes to the transparency density and differential gain within this two-band model. In order to estimate the volumetric recombination current density as a function of carrier density, a density-independent (monomolecular) recombination lifetime of 5 ns and an Auger coefficient of C=1.1×10^{-26} cm^6/s (the mea-

sured value for InAs [59]) were used. The same Auger coefficient is used for all of the cases, in spite of the fact that the Auger probability decreases as the density of states imbalance is reduced. That makes the results shown in Figure 11 a conservative estimate of the improvements that come from balancing the densities of states. The peak value of a figure of merit curve is inversely proportional to the threshold current density required to operate an optimized laser based on that material. The figure of merit for $m_v = 0.05\, m_0 = 2\, m_c$ is more than 40 times larger than that for nominal masses. Even a factor of two reduction in the mass leads to an improvement of approximately a factor of four in the threshold figure of merit.

3.3 Reducing Density of States Imbalance using Strain

One avenue for decreasing the impact of the density of states imbalance is to p-dope the active region material in order to give the carrier density in the valence band a head-start. Although this can be used to increase the differential gain, its use is limited in the mid-infrared because of Auger recombination rates and optical losses (due to inter-valence band and free carrier absorption) that depend directly on the carrier densities.

There are two additional ways to reduce the valence band density of states, thereby reducing the density of states imbalance. The first is through the incorporation of biaxial strain [5]. While the net effect of hydrostatic strain is to change the band gap of the material, biaxial strain causes a break in the symmetry of the crystal and leads to a splitting of the heavy and light hole bands. Use of either compressive or tensile strain can reduce the density of states by removing the degeneracy of the light and heavy hole states at zone center. The maximum thickness of a strained layer is limited due to problems with lattice relaxation. This makes strain a more useful tool in a quantum well or superlattice, where individual layers tend to be thin, than in a bulk material.

An example of the change in electronic structure due to strain is shown in Figure 12, which shows the band structure of $Ga_{0.80}In_{0.20}Sb$ without strain and coherently strained to GaSb ($\Delta a/a_0$ = 1.2%). In the in-plane direction, the heavy and light hole states are split apart by 83 meV and the masses are modified significantly. The heavy hole state becomes very light and the light-hole state becomes heavy until the bands approach each other, at which point they return to their unstrained curvature. The in-plane mass of the heavy hole at zone-center approaches that of the conduction band ($m_h = 0.04\, m_0$). The density of states thus appears modified by the placement at the band edge of a 40 meV deep region which has a density of states comparable to that of the conduction band. The ratio of the valence to conduction band density of states is therefore reduced from 53 in the unstrained material to 12 in the strained material.

3.4 Reducing Density of States Imbalance using Quantum Confinement

The disparity in the density of states can also be reduced with quantum confinement. There are two distinct benefits to using a quantum confined system: splitting

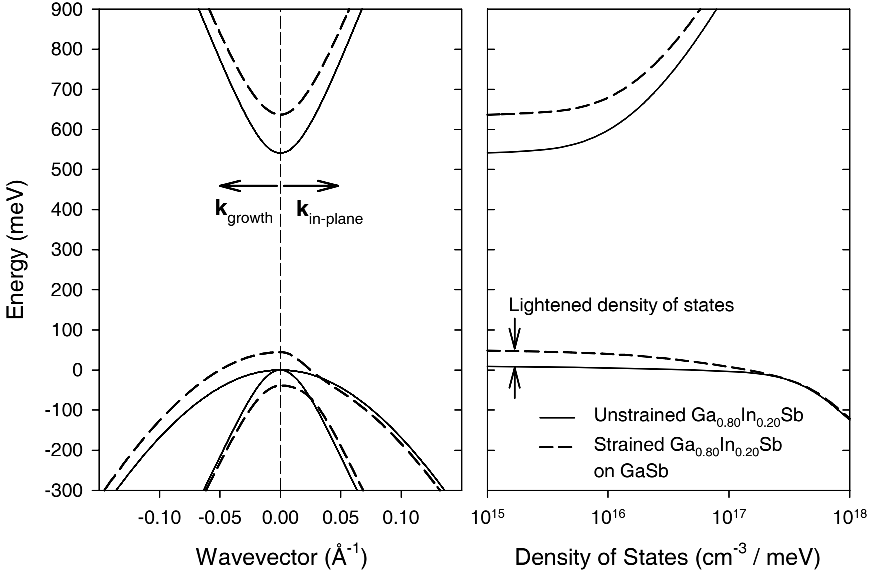


Figure 12. Band structure of unstrained (*solid lines*) and strained (*dashed lines*) $Ga_{0.80}In_{0.20}Sb$ on GaSb. One of the effects of strain is to split the degeneracy of the valence band edge. For compressive strain, which is shown here, the heavy hole moves upward in energy and becomes very light in the in-plane direction. The light hole state moves downward and becomes heavier in the in-plane direction. Both bands regain their normal curvature after they approach each other near the unstrained energy. On the *right* is shown the density of states on a logarithmic scale. The critical modification for laser operation is the region of small density of states at the band edge. The valence band density of states becomes identical for the two systems at energies below that of the unstrained valence band edge

the light-hole and heavy-hole states due to their differing effective masses and modifying the density of states from a three-dimensional character to a two-dimensional character. The differential effect of confinement on the light hole and heavy hole comes from the mass dependence of the confinement energy. For a particle in an infinite one-dimensional box,

$$E_{confinement} = \frac{\pi^2 \hbar^2}{2m^* L^2}, \tag{51}$$

where m^* is the effective mass and L is the width of the box. Because the light hole effective mass is much smaller than that of the heavy hole, the light hole state has a larger confinement energy, which splits the degeneracy of the heavy and light states, much like compressive strain.

The second benefit from quantum confinement is in the very different nature of a two-dimensional density of states compared to the normal three-dimensional density of states of bulk materials. A two-dimensional density of states has a step-

like shape rather than a square-root shape. The maximum gain that can be obtained at the band edge of a quantum well is almost always large enough to overcome cavity losses, while for a bulk material, the maximum gain that can be obtained near the band edge is small because the density of states vanishes at the band edge. This can be seen in the blue-shifting of the peak of the gain spectra shown earlier for bulk GaAs in Figure 1. Because of this blue shift, devices based on bulk materials operate at energies above the band edge. The states below the operating energy in the joint density of states are detrimental since they must be filled but do not contribute to lasing. The strict step-like density of states only exists in structures where the carriers are truly confined. Without adequate band offsets, the bands (especially in the conduction band) have some amount of dispersion. This miniband dispersion broadens the step-like density of states by the miniband width.

3.5 Development of 2–3 μm GaInAsSb Diodes

The development of diodes based on GaInAsSb in the 2–3 μm wavelength range provides a good example of the improvements that can be obtained through band edge optimizations. The development of room temperature diodes in this wavelength range can be understood almost entirely in terms of band-edge optimization of the

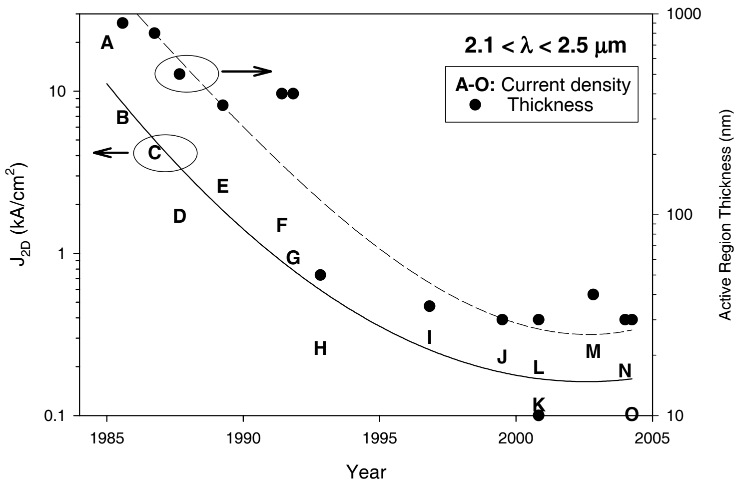


Figure 13. History of the development of GaInAsSb lasers in the 2.1–2.5 μm wavelength range. The *letters* indicate experimental threshold current densities (*left axis*). The *circles* show the thickness of the active region (*right axis*). For quantum well materials, the active region thickness is the number of wells times the well width. Additional information on each device is provided in Table 2. The *lines* are merely a guide to the eye for the threshold current density (*solid line*) and the active region thickness (*dashed line*). The threshold current density tracks the reduction in active region thickness over nearly two orders of magnitude, which covers both bulk double-heterostructures and quantum well diodes

Table 2. Experimental results for room-temperature GaInAsSb lasers in the 2.1–2.5 μm wavelength range. The letters representing each diode are the same as in Figures 13 and 15

	Year	Author	Wavelength (μm)	Strain ($\Delta a/a_0$, %)	Reference
A	1985	Bochkarev	2.3	Unstrained	[60]
B	1985	Caneau	2.2	Unstrained	[61]
C	1986	Chiu	2.2	Unstrained	[62]
D	1987	Caneau	2.2	Unstrained	[63]
E	1989	Zyskind	2.2	Unstrained	[64]
F	1991	Choi	2.3	Unstrained	[65]
G	1991	Choi	2.2	Unstrained	[66]
H	1992	Choi	2.1	Unstrained	[67]
I	1996	Baranov	2.36	Unstrained	[68]
J	1999	Garbuzov	2.3	1.5-2.3	[69]
K	2000	Mermelstein	2.26	1.4	[70]
L	2000	Mermelstein	2.26	1.4	[70]
M	2002	Kim	2.5	1.6	[71]
N	2004	Salhi	2.3	1.9	[72]
O	2004	Salhi	2.4	1.4	[73]

quaternary active region. The historical development of diodes between 2.0–2.5 is shown in Figure 13. The trend of decreasing current density with active region thickness is striking. The reduction of the reported threshold current densities is linear with active region thickness over more than an order of magnitude of variation. Much of this reduction can be understood from the figure of merit curve for bulk GaInAsSb.

The figure of merit curve for bulk $Ga_{0.84}In_{0.16}As_{0.15}Sb_{0.85}$ (from [61–65, 67]) is shown as a function of carrier density in Figure 14. The volumetric recombination rate, J_r, is estimated from the carrier density assuming a monomolecular recombination lifetime (Shockley-Read-Hall and surface recombination) of 5 ns, a calculated radiative recombination rate of 8×10^{-11} cm^3/s, and an Auger coefficient of 1×10^{-28} cm^6/s [74–78]. The curve reaches a maximum of 273 μm$^2/A$ at a carrier density of 1.4×10^{17} cm^{-3}, which corresponds to a material gain of 311 cm^{-1}.

Areal threshold current densities can then be calculated from estimated cavity properties. The effective mode width is taken to be 1 μm based on a normal double-heterostructure or broadened-waveguide width of 0.8 μm. The non-active-region cavity loss is estimated from the mirror losses, α_m, and modal losses in the waveguide, $\langle\alpha_w\rangle$, using

$$\langle\alpha\rangle = \alpha_m + \langle\alpha_w\rangle \; . \tag{52}$$

For a low-loss cavity, $\langle\alpha_w\rangle$ can be as small as 5–10 cm^{-1}. Mirror losses are typically 12 cm^{-1} for a 1 mm long device with uncoated facets, but can vary from 6–40 cm^{-1} for typical devices based on the cavity length and facet coatings. For the calculations that follow, we will use a family of modal losses including 10 cm^{-1}, which represents the limit of a 2 mm long low-loss device with uncoated facets, 20 cm^{-1}, which represents a more conventional 1 mm cavity, and 30 cm^{-1}, which represents

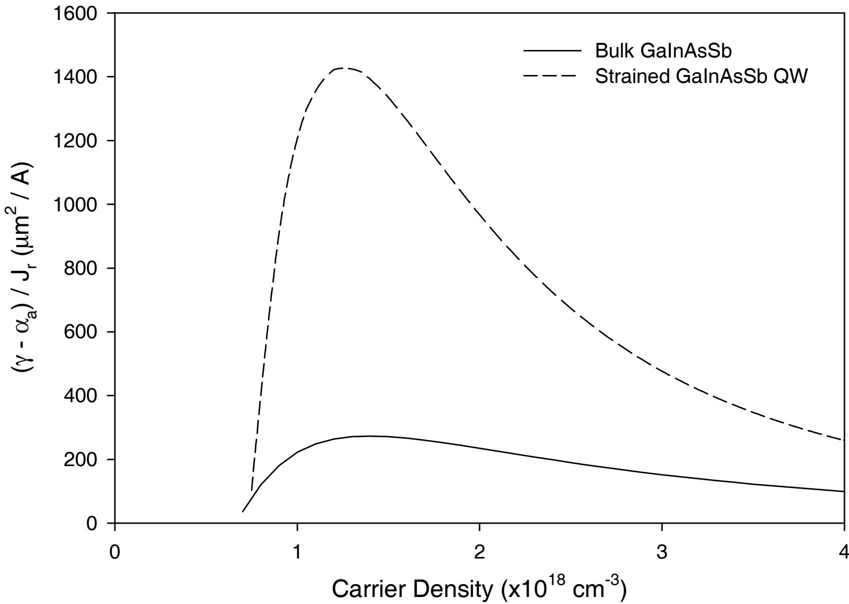


Figure 14. Figure of merit curves for bulk $Ga_{0.84}In_{0.16}As_{0.15}Sb_{0.85}$ and the compressively-strained multiple quantum well structure of [70]. For the quantum well diode, the carrier density shown is the density in the well

a device with larger waveguide losses or a short cavity length. The results are shown as solid lines in Figure 15 with the experimental values from Figure 13. The minimum current density for bulk GaInAsSb in a low-loss cavity occurs for a thickness of approximately 30 nm. Much of the reduction in current density comes simply from reducing the thickness of the active region, although the results indicated "H"–"O" lie below the 10 cm^{-1} curve for bulk GaInAsSb.

The discrepancy between the results labeled "H"–"O" and the solid lines in Figure 15 is not surprising given that the solid lines are for unstrained bulk GaInAsSb and do not incorporate the benefits that arise from strain and quantum confinement. The agreement is better when comparing the results "H"–"O" with the current density for a strained GaInAsSb quantum-well (as found in the diodes "K" and "L" from [70]), which is shown as the dashed line in Figure 15. The improvement over bulk for the same modal loss is approximately a factor of five. The optimal thickness for a low-loss cavity in this case is 10 nm, which is a single quantum well.

The improved performance comes from modifications to the band-edge electronic structure of the materials due to quantum confinement and strain. To illustrate this, the band structures of unstrained bulk GaInAsSb and the strained GaInAsSb quantum well of [70] are compared in Figure 16. The splitting seen in Figure 16 is

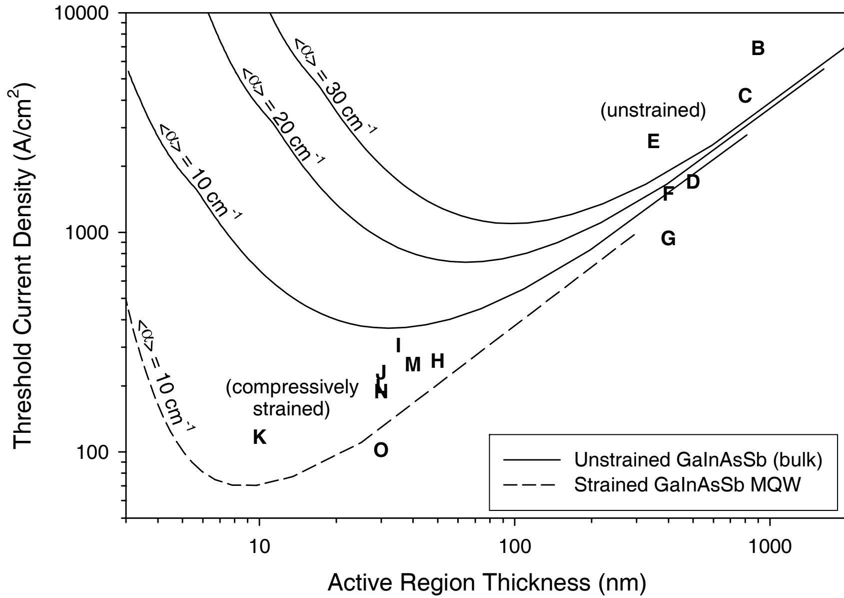


Figure 15. Calculated threshold current density as a function of active region thickness for unstrained $Ga_{0.84}In_{0.16}As_{0.15}Sb_{0.85}$ (*solid lines*) for a range of non-active-region modal loss values. Also shown is the threshold current density for the compressively strained multiple quantum well structure of [70] (*dashed line*). The *letters* indicate experimental threshold current densities as in Figure 13. More information on the diodes represented by each of the letters is available in Table 2

due to the differential impact of quantum confinement on the light and heavy band edge states as well as the further splitting of these states by the compressive strain. The splitting significantly reduces the valence band density of states at the band edge, shown in Figure 17. The conduction band density of states is also modified by the strong conduction band quantum confinement. Note that the first shoulder in the valence band density of states is at approximately 1×10^{16} cm^{-3}/meV, which is within a factor of two of that of the conduction band-edge density of states. This is consistent with the highly improved ratio of the valence and conduction band in-plane masses, 7.1, which is almost an order of magnitude greater than the ratio of 60 for unstrained bulk GaInAsSb. More improvement than this is not obtained because this shoulder in the valence band density of states does not extend more than 30 meV into the valence band (corresponding roughly to the zone-center splitting of the top two heavy hole subbands). Hence near room temperature the carrier density remains sensitive to the valence band density of states deeper than 30 meV within the valence band. In Section 3.6 the ability to further tailor the band-edge states in type-II structures will be exploited to achieve density of states ratios much closer to unity.

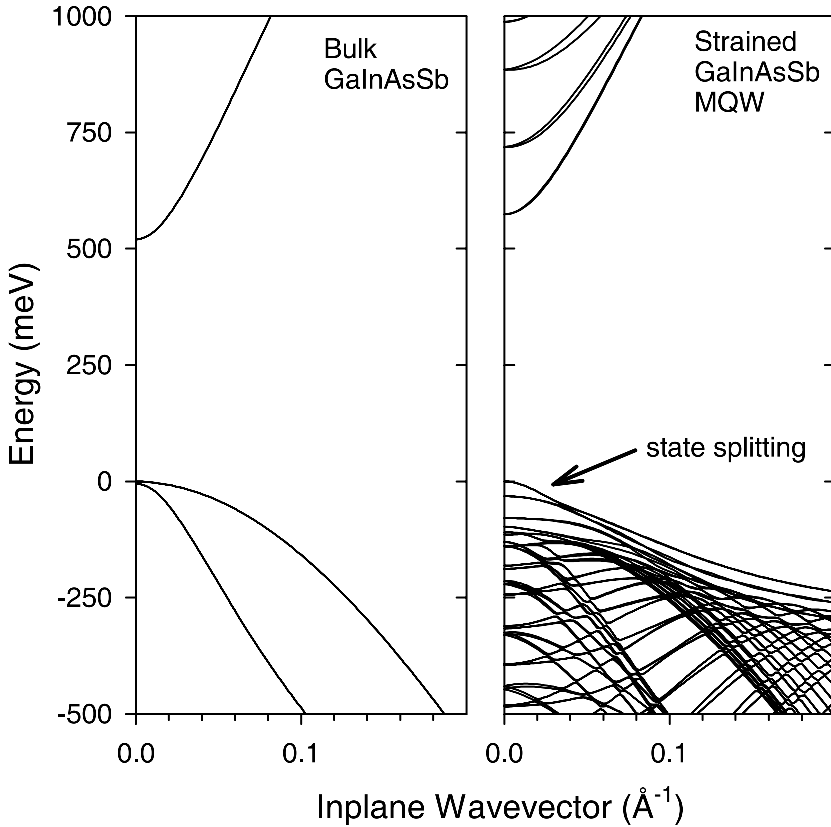


Figure 16. Comparison of band structures of bulk $Ga_{0.84}In_{0.16}As_{0.15}Sb_{0.85}$ and the compressively strained quantum well of [70]. The key difference is the reduction of the mass of the valence band-edge state due to the splitting of the zone-center valence states caused by quantum confinement and strain

The net benefit of quantum confinement and strain is illustrated in Figure 14, which shows the figure of merit for the strained quantum well of [70] (using the same recombination parameters as for bulk) in addition to the curve for bulk GaInAsSb. The reduction in the band-edge ratio of the densities of states produces a peak figure of merit for the strained material of 1420 $\mu m^2/A$, which is a factor of five larger than for the unstrained material. It is even a factor of 5.6 times the value for bulk GaAs.

For wavelengths longer than 2.5 μm, optimization of the active region thickness is not in itself sufficient to reduce the threshold current density to a practical level. As is shown in Figure 18, until 1999 room temperature threshold current densities

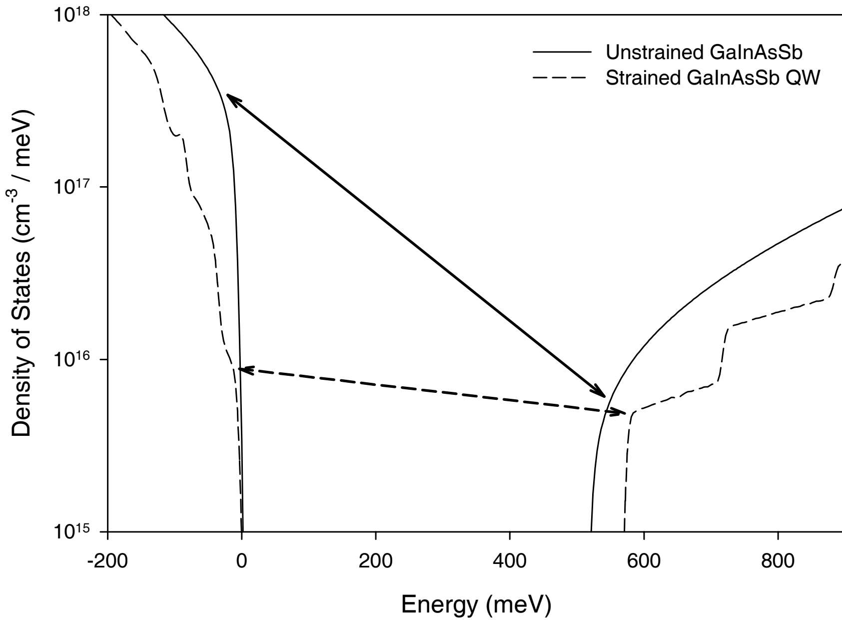


Figure 17. Comparison of the density of states for bulk $Ga_{0.84}In_{0.16}As_{0.15}Sb_{0.85}$ and the compressively strained quantum well of [70]. The splitting of the valence band induced by strain and quantum confinement creates a small shoulder in the valence band density of states. The magnitude of the valence band density of states at this shoulder is only twice that of the conduction band. However, this reduction of the valence band density of states relative to bulk only extends 30 meV into the valence band. The step-like structure in the conduction band density of states is due to quantum confinement in the growth direction

were still around 10 kA/cm^2. The major improvement in design strategy in the late 1990s was the introduction of significant levels of strain into the active region.

Development at wavelengths greater than 2.5 μm lagged that at shorter wavelengths because of a number of factors, all of which will continue to be issues as the wavelength increases beyond 3 μm. The optical mode naturally spreads as the wavelength increases, which necessitates a thicker active region and increases the threshold current density. Waveguide losses also increase because of free carrier and intersubband absorption, particularly in the p-doped clad. As noted previously, Auger recombination increases exponentially with wavelength. For example, Garbuzov estimated from device results that the Auger coefficient at room temperature increases by a factor of three between 2.3 and 2.6 μm [87]. Overcoming these detrimental factors to produce useful diodes requires even greater levels of strain.

To illustrate the improvements that were obtained by incorporating significant levels of strain, we can compare the 2.8 μm wavelength materials from [80] and [84], which correspond to symbols “Q” and “W” in Figure 18 and Table 3. Figure 19

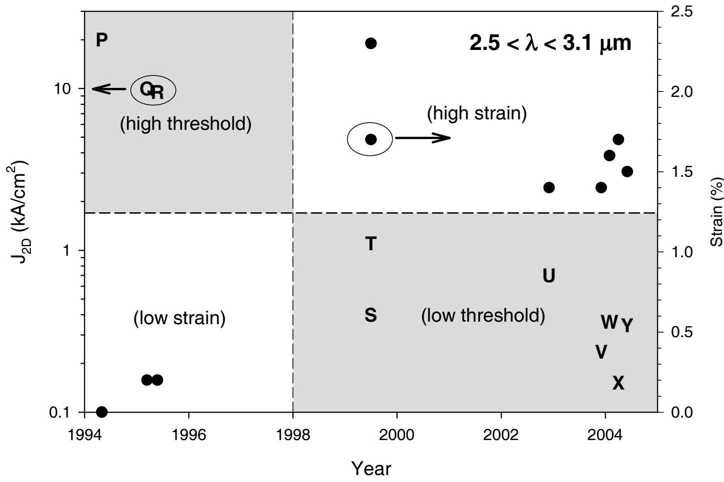


Figure 18. History of the development of room-temperature GaInAsSb lasers in the 2.5–3.1 μm wavelength range. The *letters* correspond to threshold current density (*left axis*) and the *circles* in this case indicate the level of strain in the active region (*right axis*). The introduction of $>1.2\%$ strain after 1998 is the most significant reason for the reduction in threshold current density, even for diodes with wavelengths as long as 3.04 μm. More information on the diodes represented by each of the letters is available in Table 3

Table 3. Experimental results for room-temperature 2.5–3.1 μm GaInAsSb quantum well diodes (except for diode "P", which only operated up to 255K). The letters representing each diode are the same as in Figure 18

	Year	Author	Wavelength (μm)	Num. QW	QW Width (nm)	Reference
P	1994	Choi	3.0	Double het.	900	[79]
Q	1995	Lee	2.78	4	10	[80]
R	1995	Lee	2.78	4	10	[81]
S	1999	Garbuzov	2.6	2	10–20	[69]
T	1999	Garbuzov	2.7	2	10–20	[69]
U	2002	Grau	2.72	3	20	[82]
V	2003	Grau	2.96	2	20	[83]
W	2004	Grau	2.81	3	20	[84]
X	2004	Salhi	2.61	2	14	[85]
Y	2004	Lin	3.04	2	20	[86]

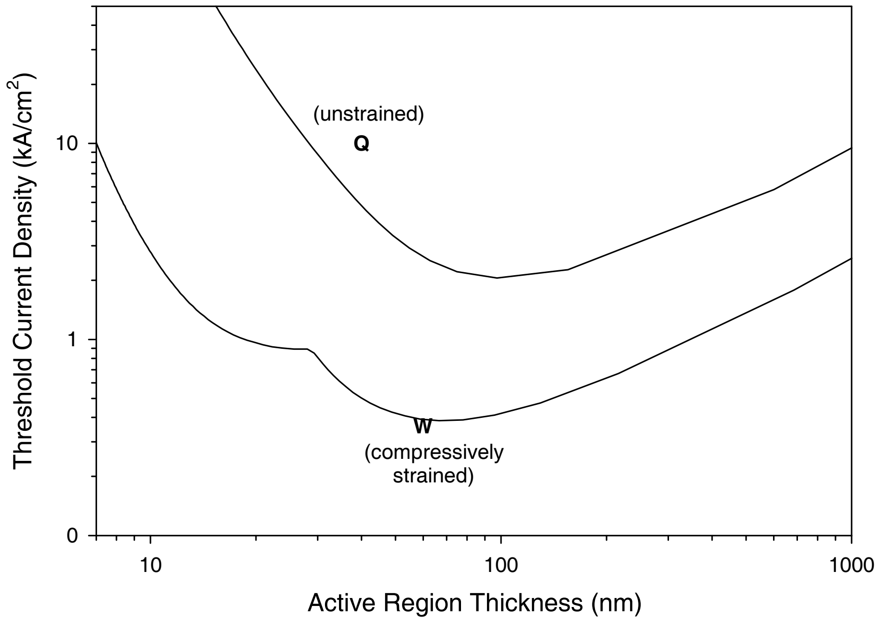


Figure 19. Calculated threshold current density for the 2.8 μm unstrained quantum well of [80] and the 2.8 μm strained quantum well of [84] as a function of active region thickness (number of wells times the width of each well). Also shown are *letters* indicating the experimental thresholds. The letters represent the same diodes as in Figure 18 and Table 3. The change in slope for the strained well for total thicknesses below 30 nm is due to the second level transition becoming dominant at higher carrier densities

shows the estimated threshold current densities for the two quantum well materials assuming a mode width of 1.2 μm and modal waveguide losses of $\langle \alpha_w \rangle$ = 10 cm^{-1}. Values for the mirror losses are based on the devices used: 13 cm^{-1} for the unstrained quantum well (500 μm cavity with facets coated for 90% and 30% reflectivity) [80] and 6 cm^{-1} for the strained quantum well (2 mm long uncoated device) [84]. The Auger coefficient at 2.8 μm is taken as 5×10^{-28} cm^6/s. In both cases, the estimated values show reasonable agreement with the corresponding device result, although it appears that a reduced threshold current density could have been achieved with unstrained quantum wells by increasing the number of quantum wells.

As with the structures in the 2-2.5 μm wavelength range, the differences between these strained and unstrained quantum wells are a consequence of reducing the valence band density of states. Figure of merit curves for these diodes are shown in Figure 20. The ratios of valence to conduction band densities of states are 52 for the unstrained quantum well and 17 for the strained wells. The amount of strain and quantum confinement that can be used is limited in this case by the need for wide

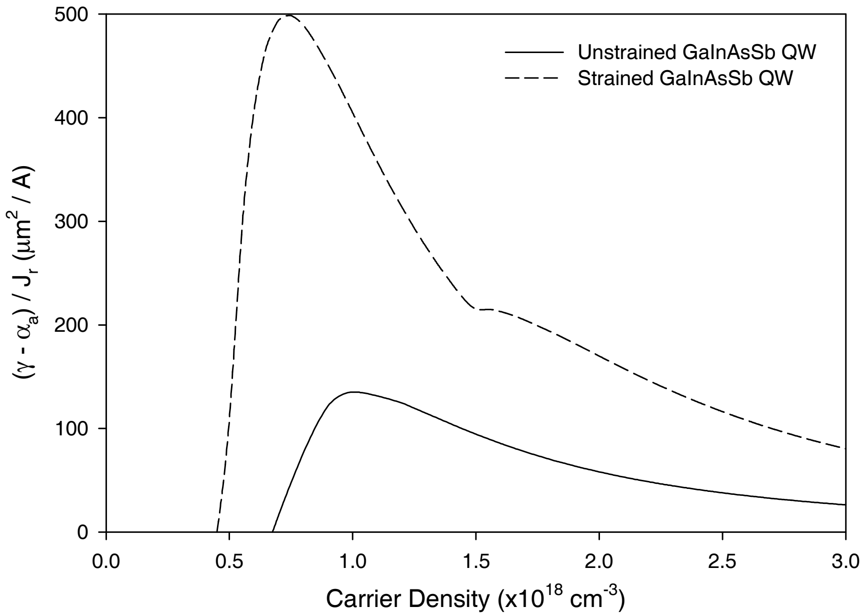


Figure 20. Calculated threshold figure of merit curves for the unstrained and compressively strained 2.8 μm quantum wells of [80] and [84]. The peak value for the strained quantum well is nearly five times that of the unstrained quantum well. The primary reason is the reduced density of states at the valence band edge due to the use of strain

quantum wells to keep the confinement energy from increasing the wavelength. Thus the desire for longer wavelength and greater band edge optimization are at odds here.

3.6 Optimizing Band Edge Properties Using InAs/GaSb/AlSb Materials

The approaches described so far have been able to take diode technology in the 2–3 μm wavelength range to high-power, room-temperature operation. The current state of interband laser diodes beyond 3 μm is far less satisfactory. As described previously, a significant source of the challenge is the increase in density of states imbalance and Auger recombination as the band gap decreases. Thus, to reach high-power room-temperature operation, the conduction band and valence band densities of states would have to be even closer to perfect balance in the >3 μm wavelength range. Taking the strategies described for the 2–3 μm wavelength range to greater extremes, however, brings us to material limitations. If the strain is increased beyond 3%, then the critical thickness becomes very small. If the wells are made narrower in order to increase quantum confinement, then several undesirable consequences are possible: for example, the optical emission wavelength can drop below 3 μm, or the

lowest conduction and valence band states can approach the energies of the barriers, turning carrier escape out of the quantum wells into a significant problem.

Superlattices based on the "broken-gap" semiconductor materials InAs/GaSb [88] and InAs/GaInSb [89] were first proposed for use as long-wavelength detectors. The energy gap of these superlattices can be tuned from 0.8 eV to < 0 meV (*i.e.*, semimetallic, rather than semiconducting behavior) based on the thicknesses of the constituent layers. Calculations of threshold current densities for InAs/GaInSb superlattices [90, 91] suggested promising device performance, although the structures considered were not lattice matched to GaSb or InAs. In these materials much larger hole confinement is possible than in InAsSb/AlAsSb structures [92]. The first lasers demonstrated based on InAs/GaInSb materials consisted of short segments of these superlattices between strain-compensating barriers [93–95]. These InAs/GaInSb superlattice multiple quantum well diodes lased at wavelengths across the mid-infrared spectral range, including up to 255K pulsed at 3.2 μm. It was only discovered later that when short segments of the superlattice were separated by barriers that the electrons resided predominately in the center of the superlattice segment, whereas the holes resided predominately near the barriers, reducing the optical overlap. Furthermore, at typical lasing densities only a single conduction subband was occupied, whereas each GaInSb layer in the superlattice segment created an occupied hole miniband, creating a highly unbalanced conduction and valence density of states [96].

An improvement to this approach eliminated the edge effects by surrounding a GaInSb hole well by a pair of InAs electron wells, which were in turn surrounded by AlSb layers [97]. This material geometry has become known as the "W"-laser based on the profile of the conduction and valence band edges. Because of the large and indirect conduction band of the AlSb barriers, and concern about carrier transport through these barriers, W-lasers were originally used only as active regions for optically pumped lasers. A similar system is that shown in Figure 21, which utilizes a direct-gap quinternary alloy for the strain compensating barrier, simplifying the process of electrical injection. The large band gap of the AlSb barrier in the W-laser designs renders adjacent InAs/GaInSb/InAs wells independent, whereas the design of [96] permits coupling of the electron states in adjacent wells and has become known as a four-layer superlattice. The reduction of the valence band-edge density of states is substantial for both the W-laser design and the four-layer superlattice.

The InAs/GaSb/AlSb type-II material system provides unique opportunities for efficiently using strain and quantum confinement. In the structure shown in Figure 21, the holes reside predominately in the GaInSb region and the electrons in the InAs regions. The AlGaInAsSb quinternary provides a barrier to both electrons and holes. It is only important to have compressive strain in the region where the holes reside, thus only the GaInSb region is so strained. By limiting the compressive strain to a region of less than 30 Å thickness, and using a quinternary with tensile strain, it is possible to include a much higher level of strain (40% indium) than proposed for the W-laser (30% indium). When grown on a GaSb substrate, the InAs regions provide some compensating tensile strain, and the AlGaInAsSb quinternary can provide the rest. Thus the structure shown in Figure 21 is strain-compensated over one unit

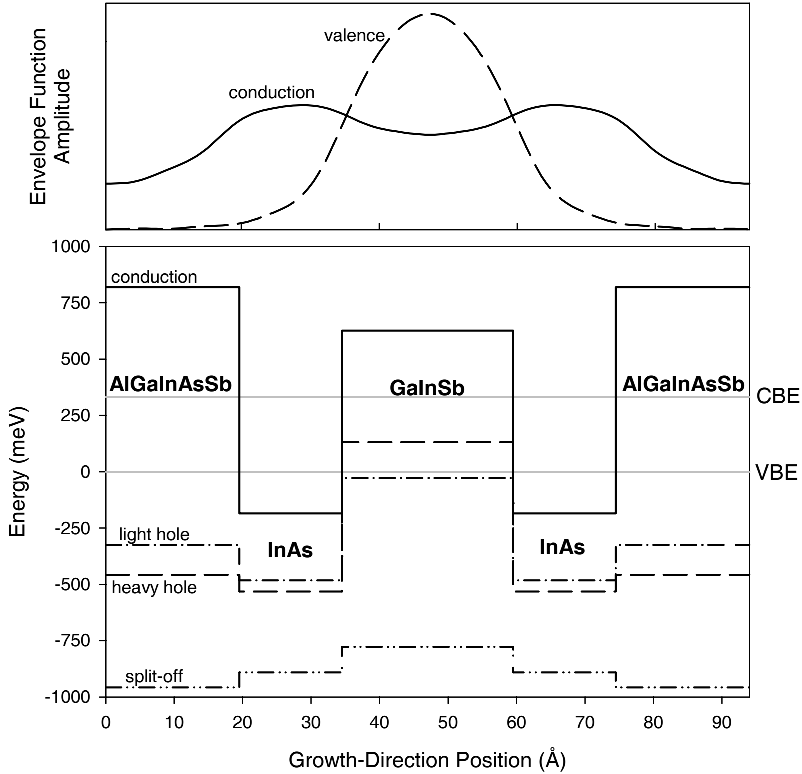


Figure 21. Band edge diagram and envelope functions for a four-layer superlattice consisting of 13.8Å InAs/24Å $In_{0.4}Ga_{0.6}Sb$/13.8Å InAs/40Å $Al_{0.3}In_{0.28}As_{0.5}Sb_{0.5}$. As the first conduction state is composed of envelope functions from the bulk conduction, light-hole and split-off bands, the magnitude of the envelope function vector is shown. The band energies are measured relative to the superlattice's top hole state

cell. Greater localization of the top hole state than the lowest conduction state also reduces the density of states imbalance between the conduction and valence bands.

The band edges shown in Figure 21 indicate that the range of conduction band edges spans 1 eV, and the range of valence band edges spans nearly 0.75 eV. Both of these numbers are more than twice the band gap of the superlattice material, suggesting that the electronic structure of the superlattice can be modified deep within both the conduction and valence bands. For a comparable type-I structure (Figure 22), such as a 150 Å $InAs_{0.85}Sb_{0.15}/In_{0.90}Al_{0.10}As_{0.90}Sb_{0.10}$ quantum well [98] the valence band offsets are considerably smaller ($\sim$70 meV), and thus much smaller improvements of the electronic structure are possible by placing the bulk material in a quantum well.

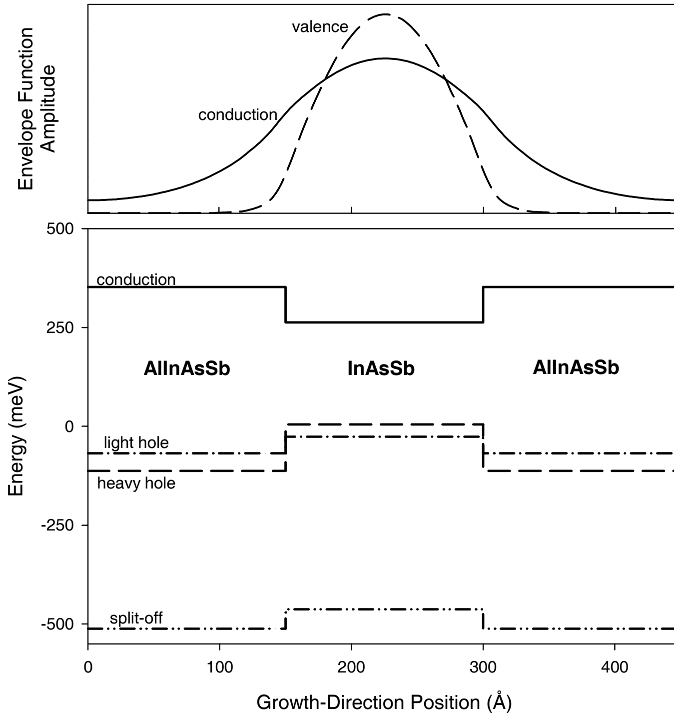


Figure 22. Band edges and envelope functions for a type-I, 150 Å $InAs_{0.85}Sb_{0.15}/In_{0.90}Al_{0.10}As_{0.90}Sb_{0.10}$ quantum well [98]. The well region is compressively strained at 0.5%

The large quantum confinement and strain splitting possible in a type-II structure produces benefits that can be clearly seen by comparing the band structure and density of states of the four-layer superlattice material (Figure 23) to those of a conventional type-I quantum well (Figure 24). Only three valence subbands are apparent in the first 400 meV below the valence maximum for the four-layer superlattice, and the splitting between the first and second valence subband is 250 meV, which is comparable to the splitting between the first and second conduction subbands. In contrast, there are a very large number of subbands near the valence maximum of the type-I quantum well. A comparison of the conduction and valence dispersions in the four-layer superlattice and type-I quantum well indicate much closer conduction and valence masses in the four-layer superlattice of Figure 23 than in the type-I quantum well of Figure 24. Figure 25 shows the consequence of these differences in band structure on the densities of states for the two structures. The resulting density of states ratio is 41.7 for the type-I quantum well and an order of magnitude less (4.1) for the four-layer superlattice.

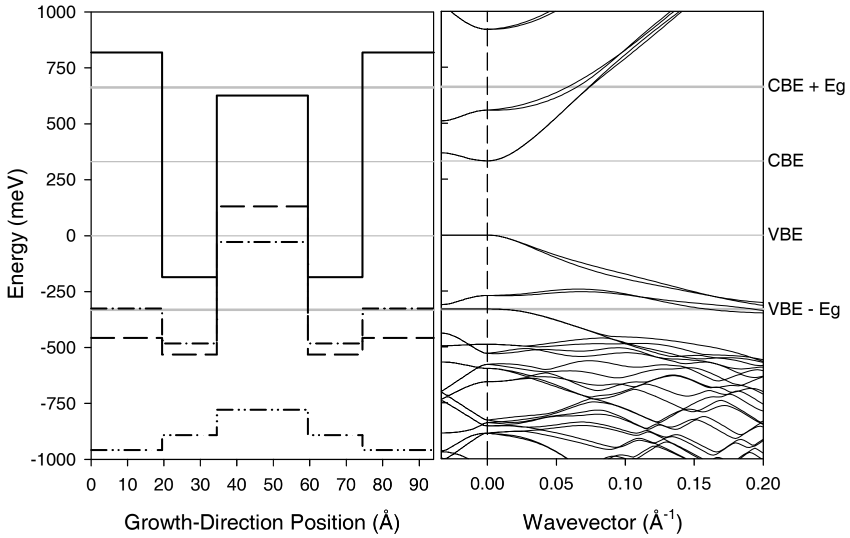


Figure 23. Band edges and dispersion relations for the four-layer superlattice of [96]. The energies of the conduction band edge (CBE), valence band edge (VBE), and states one energy gap above the CBE and one energy gap below the VBE are shown

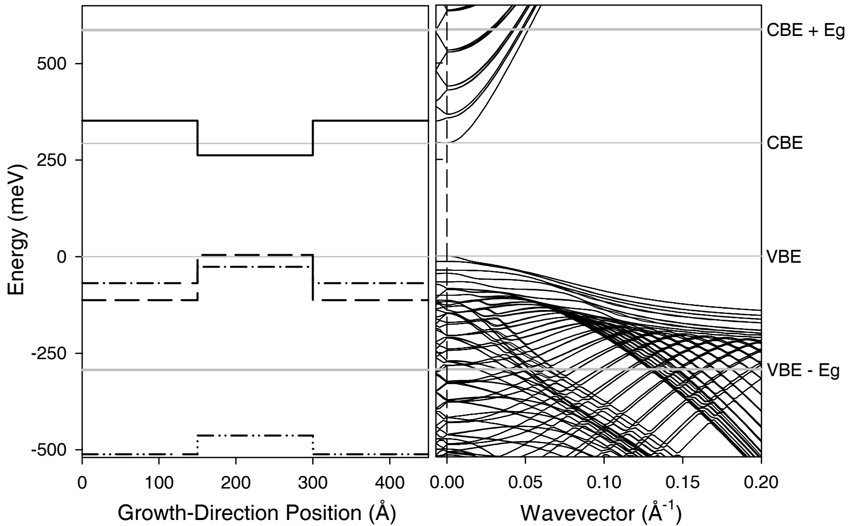


Figure 24. Band edges and dispersion relations for the type-I quantum well of [98]. The energies of the conduction band edge (CBE), valence band edge (VBE), and states one energy gap above the CBE and one energy gap below the VBE are shown

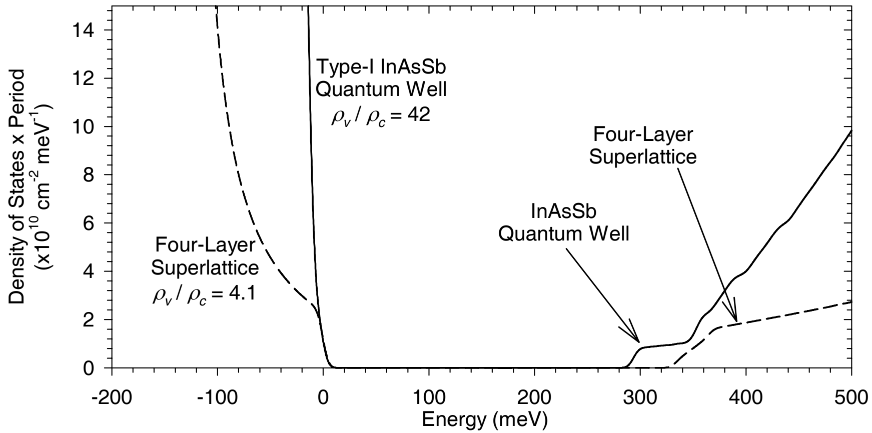


Figure 25. Comparison of the density of states for the InAsSb type-I quantum well and the four-layer superlattice of [96]. The densities of states have been multiplied by the quantum well and superlattice periods for this comparison. The four-layer superlattice derives most of its improved performance from the dramatic reduction in the valence band density of states as compared to the type-I quantum well

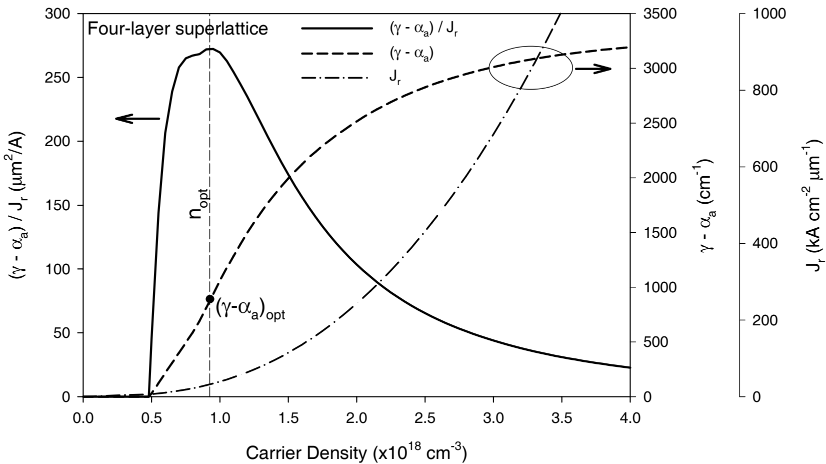


Figure 26. Figure of merit curve for the four-layer superlattice of [96]. The optimal density for active region performance, n_{opt} is indicated on the diagram, along with the net gain at the optimal carrier density. The recombination parameters used to calculate the volumetric recombination current density are a Shockley-Read-Hall time of 5 ns and an experimental Auger rate taken from [27]

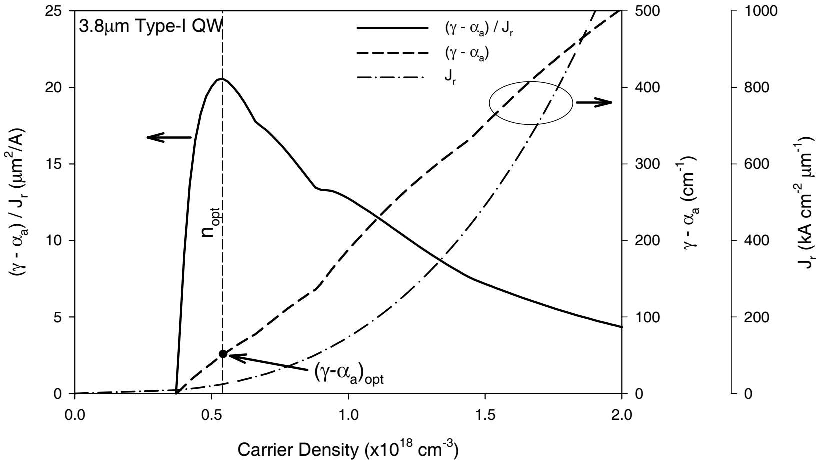


Figure 27. Figure of merit curve for the type-I quantum well of [98]. The optimal density for active region performance, n_{opt} is indicated on the diagram, along with the net gain at the optimal carrier density. The recombination parameters used to calculate the volumetric recombination current density are a Shockley-Read-Hall time of 5 ns and an experimental Auger coefficient of 9.0×10^{-27} cm^6/s taken from [99]

The benefits that flow from having large band offsets and thin layers come at a cost. The primary consequence of using type-II materials is that the overlap between electron and hole wavefunctions is reduced, and thus the optical matrix elements are reduced. The conduction and valence wavefunctions of the type-I quantum well (Figure 22) overlap very strongly, whereas those of the four-layer superlattice overlap less (Figure 21). In the four-layer superlattice the holes are strongly confined to the GaInSb layer, and the electrons are extended throughout the structure. Thus the C1-HH1 matrix element is reduced by approximately a factor of two relative to the type-I quantum well.

The net effect of this tradeoff between density of states optimization and the matrix element strength can be seen by comparing the figure of merit curves for the four-layer superlattice and the type-I quantum well, which are shown in Figures 26 and 27. For both cases, experimental Auger rates [27,99] and a 5-ns Shockley-Read-Hall lifetime were used to calculate the volumetric recombination current density. Even though the optical matrix element for the four-layer superlattice is half that of the type-I quantum well, the more balanced densities of states in the four-layer superlattice yield a figure of merit ten times larger than for the type-I quantum well.

4 Final State Optimization of Active Region Materials

Manipulating electronic states deep in the conduction or valence band is more challenging than modifying the band-edge states, yet the potential performance enhancement can be substantial, both for reducing internal loss in laser active regions and for minimizing the non-radiative recombination rate. Non-radiative (Auger) recombination adversely affects the threshold figure of merit; intersubband absorption, which contributes to the internal loss, affects both the threshold and the slope efficiency figures of merit. In the intersubband absorption and non-radiative recombination processes, either a nonequilibrium electron near the band edge is excited to an energy at least one energy gap above the conduction minimum or a nonequilibrium hole near the band edge is excited to an energy more than one energy gap below the valence maximum. Thus both of these processes depend on the initial states, which are near the band edge, and the final states, which are high in the band. The principal method of reducing these processes by controlling the initial states is to reduce the threshold carrier density. This can have the undesirable effect of reducing the gain as well; the previous two sections described how to optimize all these quantities using the two figures of merit. Manipulation of the final states, however, has no negative consequences for gain, as the gain only involves the band-edge states.

The strategies that can be used to modify the electronic structure of the final states to suppress these processes are similar strategies to those used to modify the band-edge states. A simple rule to evaluate whether it is possible to control the final states is to consider whether the available band offsets of the material system span the states which need to be restructured. For example, Figure 23 shows the band edges of the constituent materials of the four-layer superlattice, as well as the energies one band gap above the conduction minimum and one band gap below the valence maximum. The highest conduction band edge shown is $\sim$150 meV above the conduction band energy to control, and the lowest valence band edge shown is $\sim$150 meV below the valence band energy to control, thus manipulation of these states is possible. This is an unusually large range of band edges, and is considerably greater than that available in ordinary type-I quantum wells, such as that shown in Figure 24. A close examination of the band structure on the right hand side of Figure 23 indicates that the effects of valence subband quantization are highly pronounced for energies above the valence band edge of the material with the lowest valence band edge. Below that energy the subband spacing becomes very small, and no clear gaps are apparent in the bands.

This guideline for modifying final states suggests that as the band gap of a material diminishes, it becomes easier to structure the final states. Initial proposals to control final states in order to suppress nonradiative recombination focused on detector materials with band gaps in the 8–12 μm range — a range where all the higher-energy subbands of InAs/GaInSb superlattices can be forced to lie well over a band gap away from the band edge [100, 101].

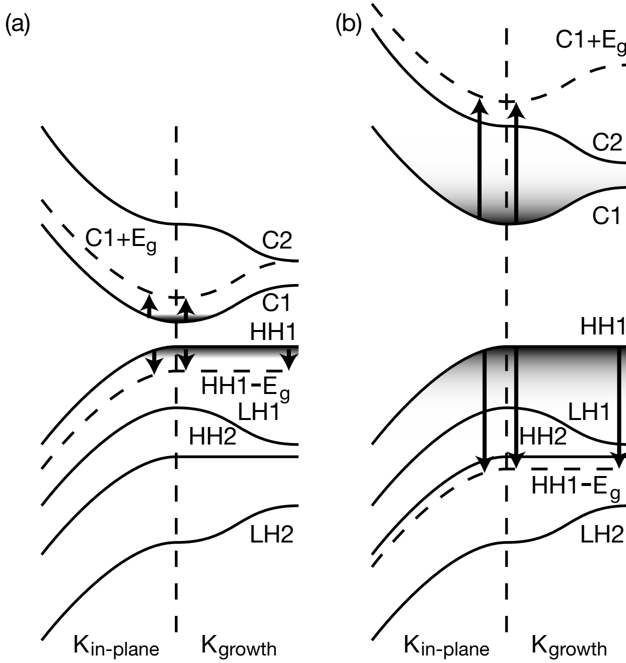


Figure 28. Band structures and relevant energy ranges for final states of intersubband absorption for a narrow-gap semiconductor, similar to that found in **(a)** long-wavelength infrared materials and **(b)** mid-wavelength infrared materials. The *dashed lines* indicate the energies of potential final states for intersubband absorption at the lasing photon energy within the conduction subbands and valence subbands

4.1 Suppression of Intersubband Absorption

In the intersubband absorption process (shown schematically in Figure 28), a photon excites the nearly vertical ($\Delta k \sim 0$) transition of an initial, band-edge state to a final state one photon energy away from the band edge. The principal source of intersubband absorption in these materials is the optical excitation of holes between valence subbands, as there are many more valence subbands near one energy gap away from the band edge and as the masses of valence subbands are heavier than the masses of conduction subbands. Thus the most important initial states are located near the valence band maximum at zone center. The nearly vertical nature of the intersubband absorption transition makes intersubband absorption optimization considerably simpler than Auger recombination, as it tends to involve final states nearer to zone center.

Figure 28 illustrates the potential for reducing intersubband absorption by eliminating potential final states for (a) a long-wavelength infrared material and (b) a mid-wavelength infrared material. The top two valence subbands are split due to strain and quantum confinement. Nonequilibrium carriers are located in the upper valence band, as indicated by the dark shading. The lasing photon energy is approx-

imately the band gap of the material, and any optically-induced transition exciting nonequilibrium holes from the upper valence band deeper into the valence subbands will be a vertical transition of energy comparable to the band gap (indicated by the arrows). If the splitting between the upper valence subband and the other subbands is larger than the energy gap (as shown in Figure 28(a)), then this lower band cannot provide final states for intersubband absorption, and the entire process is eliminated. The two lowest conduction subbands are also split by quantum confinement, usually by considerably more than the valence subband splitting, and therefore it is not possible to excite nonequilibrium electrons from the lowest conduction subband to the next conduction subband. Although only a slice of the band structure is shown in Figure 28, the same lack of final states occurs for all momenta with significant nonequilibrium carrier occupation.

In the mid-infrared, however, such a substantial splitting has not been achieved in a real material system. A more realistic situation is that shown in the schematic Figure 28(b), where an effective gap near the zone center is located at the energy relevant for intersubband absorption. As shown in Figure 28(b), as the intersubband absorption transitions are vertical, the uppermost valence band can be shifted down by the lasing photon energy, and if that shifted band does not intersect other valence bands then intersubband absorption is suppressed. A study of the dependence of intersubband absorption (and thus internal loss) on the shifting of these bands indicates that intersubband absorption is very sensitive to the precise location of these bands [102]. Accurate prediction of the structures that will have reduced intersubband absorption relies on accurate calculations of the band structure an energy gap above the bottom of the conduction band and an energy gap below the top of the valence band.

Shown in Figure 29 is the structure of the intersubband absorption, and the consequences for gain curves, for a type-I InAsSb quantum well and the four-layer superlattice. For the type-I quantum well the large absorption feature peaking at 220 meV is due to vertical transitions of holes from the top heavy-hole subbands to light-hole subbands. Transitions involving the spin-orbit band begin at transition energies comparable to split-off energy, which for this material is 450 meV. They are just becoming visible at the right edge of the figure. A very similar intersubband absorption pattern is also present for bulk materials. For materials with wavelengths longer than $\sim$4 μm, the band gap starts to move into the region of strong intersubband absorption and performance is adversely affected. For example, in Figure 29(c), the net gain for the quantum well near the band edge is only 100 cm^{-1}, which is a third of what it would be without intervalence band absorption. The slope efficiency figure of merit, $(\gamma - \alpha_a)/\gamma$, is $\sim$30% at these carrier densities and the differential quantum efficiency will be low no matter what type of device geometry is used. This problem becomes worse at longer wavelengths and there is little that can be done about it with these type-I materials.

With the large band offsets available with the InAs/GaInSb material system, however, it is possible to structure the intervalence band absorption and clear out a spectral region between 150–250 meV, as is shown for a 5.6 μm four-layer superlattice in Figure 29(b). The difference in intervalence band absorption between the type-I material and the type-II material at 200 meV is close to 600 cm^{-1}.

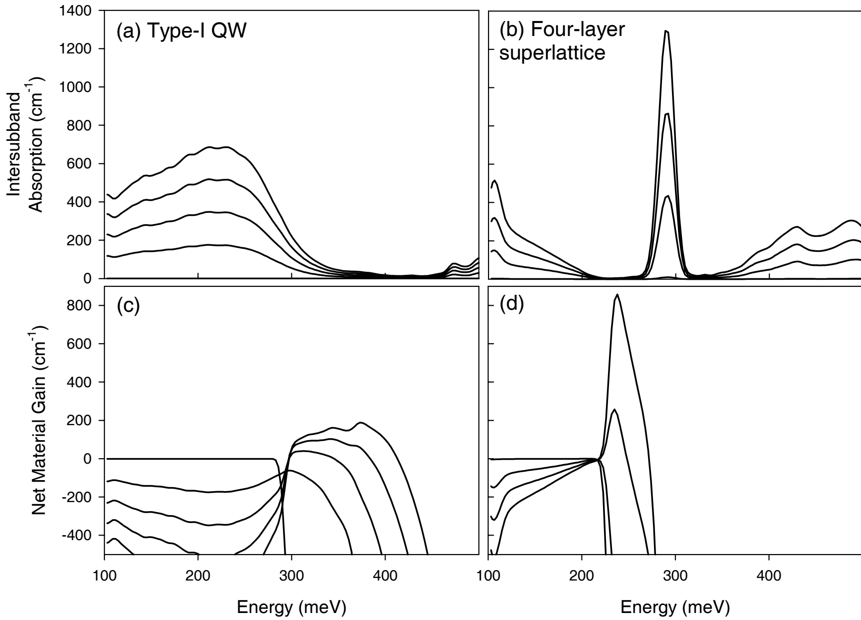


Figure 29. Intersubband absorption **(a,b)** and gain **(c,d)** for a 4 μm type-I quantum well ([98]–(a),(c)) and a 5.2 μm four-layer superlattice ([103]–(b),(d)). Curves are for n=p=0, 2.5, 5.0, 7.5, 10.0$\times 10^{17}$ cm^{-3} for the type-I quantum well and n=p=0, 2, 4, 6$\times 10^{17}$ cm^{-3} for the four-layer superlattice. A broad intervalence band absorption peaking around 200 meV in (a) is common for bulk and type-I structures. For the superlattice, however, the region between 150–250 meV is clear of intervalence absorption

From a design perspective, the presence of these large gaps near zone center in the intersubband absorption spectrum means that it is possible to reduce the intersubband absorption to zero, *if the location of these gaps can be predicted accurately*. In Section 6 the degree of accuracy in calculating electronic structure and the derived quantities such as optical absorption will be described. A larger gap in the intersubband absorption spectrum generally reduces the sensitivity of the optimized active region performance to errors in calculation and growth.

4.2 Suppression of Auger Recombination

Suppression of Auger recombination by controlling the final states of the process in an optimized material has many similarities to suppression of intersubband absorption. For Auger recombination, however, the transitions are not restricted to be vertical. Any Auger process involves the recombination of a nonequilibrium electron and a hole, and the transfer of both the excess energy and excess momentum to a third nonequilibrium carrier, which is excited high into the conduction band (for an electron) or deep into the valence band (for a hole). Figure 30(a) indicates schematically

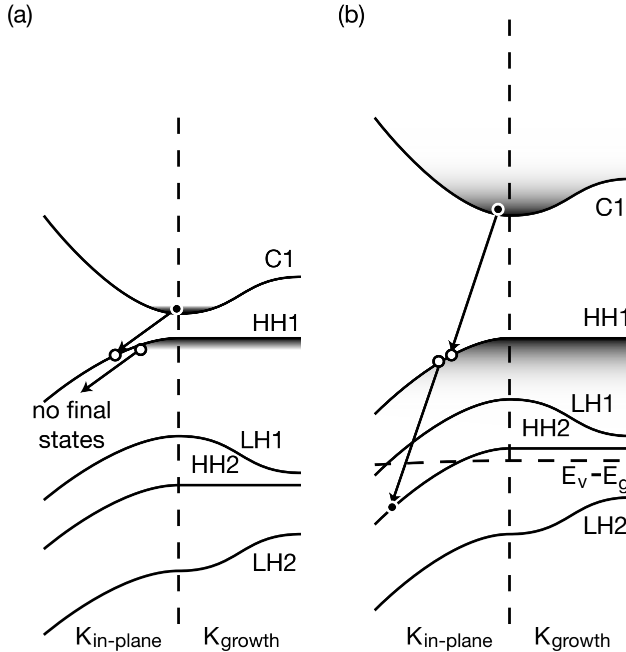


Figure 30. Schematic of Auger recombination processes where the energy and momentum of an electron-hole pair is transferred to a hole. **(a)** If all other subbands besides the band-edge valence subband have been pushed lower by an energy much greater than the band gap, then there are nearly no Auger recombination processes that conserve energy and momentum. This is the typical situation in optimized long-wavelength infrared materials based on InAs/GaInSb. **(b)** If the band gap is larger, then it is still possible to position the valence subbands so there is a gap near zone center. Here processes are still allowed that involve initial states farther from the zone center, and hence are less occupied than those at the valence maximum. This is the typical situation in mid-wavelength infrared materials based on InAs/GaSb/AlSb

the band structure and relevant energies for final-state Auger suppression in narrow-gap semiconductors. If the second valence subband is split from the first valence subband by significantly more than the band gap, then there are no possible final states for Auger transitions, or at least none that involve initial nonequilibrium hole states that have substantial occupation. Figure 30(b) shows that this set of conditions is much more difficult to satisfy for wider band gap semiconductor superlattices. Although there is no possible transition precisely at zone center that satisfies energy and momentum, by considering states *near* zone center, it is possible to conserve energy and momentum. Thus final-state optimization for Auger recombination is effective at low temperature in the four-layer structures of [28, 96, 97], but only minimally effective at room temperature [104]. A design that permits significant suppression of Auger recombination at room temperature will be described later in Section 5.

Figure 31 is a typical electronic structure for a type-II mid-infrared laser active region [96]. In order to identify the regions of greatest importance for Auger re-

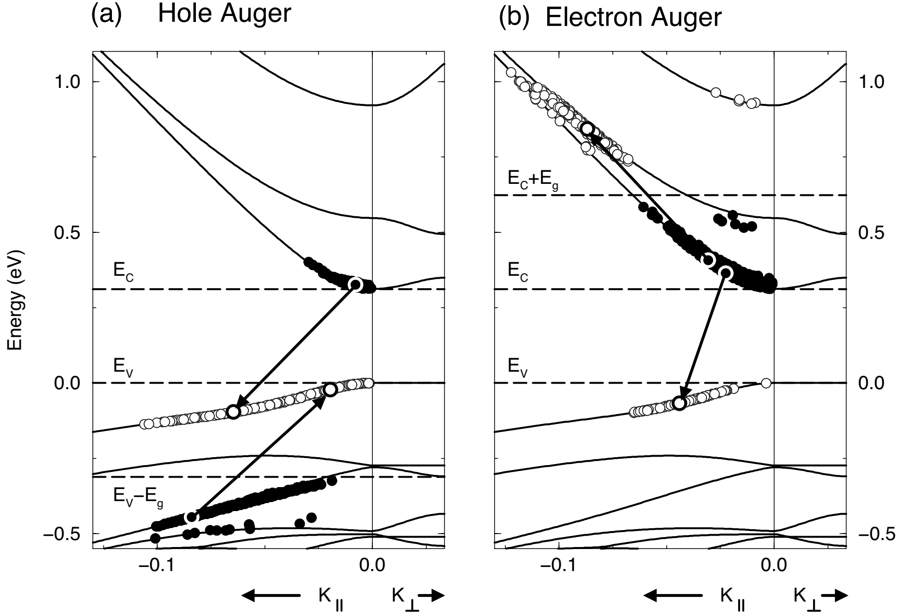


Figure 31. Band structure of the four-layer superlattice of [96] indicating the band edges and the resonance energies (*dashed lines*) as well as the most probable transitions (99% of the rate) at 300 K and $n = 5\times10^{17}$ cm^{-3} for **(a)** hole Auger and **(b)** electron Auger processes, plotted as a function of *in-plane* momentum only. The single most probable transition is shown (*with arrows*) as a schematic Auger process with holes indicated by the *larger, dark empty circles* and electrons indicated by *filled circles*. Because only a slice of the three-dimensional band structure is shown, momentum transfer vectors are projected onto the slice and the schematic scattering process appears not to conserve momentum. This is a misleading result of projecting the three-dimensional band structure onto a line plot. $K_\perp$ is the momentum along the growth direction, while $K_{||}$ is the momentum parallel to the interfaces. Copyright (1999) by The American Physical Society [27]

combination, calculations of the Auger recombination rate were performed using a coarse grid for the 12-dimensional integration over momentum space. This resulted in a sampling of the most probable Auger transitions in the material for a given temperature and carrier density. The initial and final energies and wavevectors for these transitions for $n = 5 \times 10^{17}$ cm^{-3} at 300K are shown overlaid on the four-layer superlattice's electronic structure for (a) processes where the excess energy and momentum is transferred to a nonequilibrium hole, and (b) processes where the excess energy and momentum is transferred to a nonequilibrium electron. The calculations of the total Auger recombination rate are in excellent agreement with experiment, as reported in [27], and shown in Section 6. The "hole Auger" process dominates in these materials when there are equal numbers of electrons and holes.

The minimum energy for these final states is an energy gap from the band edges, and is indicated in Figure 31 by a dashed line. Although a large gap is evident in the electronic structure near that dashed line at zone center, there is no gap present

by the time the in-plane momentum reaches ~ 0.01Å in the valence band (~ 0.03Å in the conduction band). Even further from zone center there are many permitted transitions, distributed predominantly around $K_{||} \sim 0.05$Å in the hole states. Consequently, there is little final-state suppression in the four-layer superlattice at room temperature.

Because of the large gap in final states at zone center, it was originally supposed that there would be large fluctuations in the Auger rate as states moved into and out of resonance with the band gap. However, careful consideration of theoretical calculations [104] and experimental measurements [27] of the Auger recombination rate of the four-layer superlattice of Figure 21 indicate that only about a factor of 2 suppression of the Auger recombination rate at room temperature can be attributed to final state optimization. These results are now understood to be due to significant diminishing of the final-states gap away from zone center.

5 Cavity-integrated Active Region Design for Optimal Laser Performance

The previous sections have described strategies for optimizing laser active regions in cavities that have specific known internal losses. It is reasonable to consider the compatibility of a given active region with the cavity — this set of topics includes the effectiveness of carrier injection into the active region, which depends on the conductivity of the surrounding cavity as well as the vertical transport properties of the active region material, and also the optical confinement in the cavity. In this section we present our "best case" integrated design of active region, electrical transport, and optical cavity [105]. This design has not yet been fabricated in a device, but illustrates the possibilities inherent in mid-infrared laser diode design through electronic structure engineering.

5.1 Optimization of the Active Region

The potential of suppressing intersubband absorption and Auger recombination, and the particular opportunities available in type-II structures, were explored in Sections 3 and 4. Suppression of intersubband absorption at room temperature is possible, but might be sensitive to growth and calculation errors. Suppression of Auger recombination from final-state optimization appears to be much less than desirable. Successful suppression of Auger recombination at room temperature requires a strategy for extending the final-states gap that is present at zone center to larger in-plane wavevectors without compromising the band-edge optimizations. Such a strategy has been developed and used in the design of an active region that is optimized for significant final-state Auger suppression at room temperature.

The designed structure is shown in Figure 32. The optimization strategy has been refined to focus on the band edges and the final-state energies entering into the intersubband absorption and Auger processes. Instead of using a high barrier from a

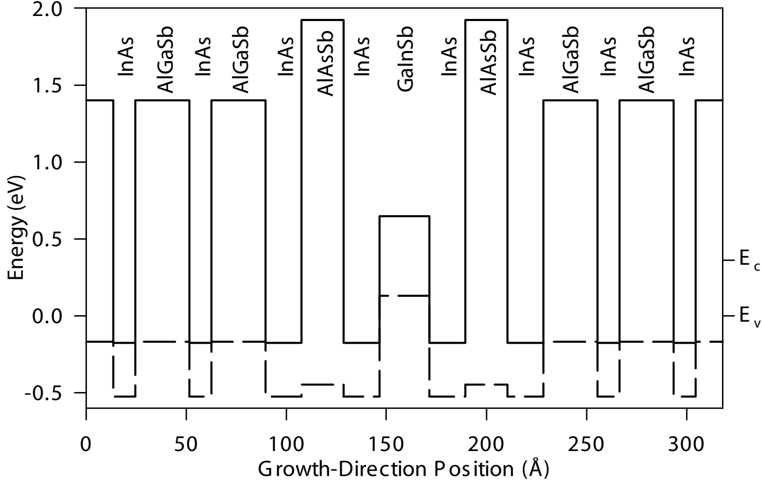


Figure 32. Bulk band edge energies as a function of growth direction position for a 3.4 μm broken-gap quantum well. The *solid line* is the bulk conduction band energy and the *dashed line* is the bulk heavy-hole energy. The conduction and valence band energies of the superlattice are shown on the right of the figure. Reprinted with permission from [105]. Copyright 2001, American Institute of Physics

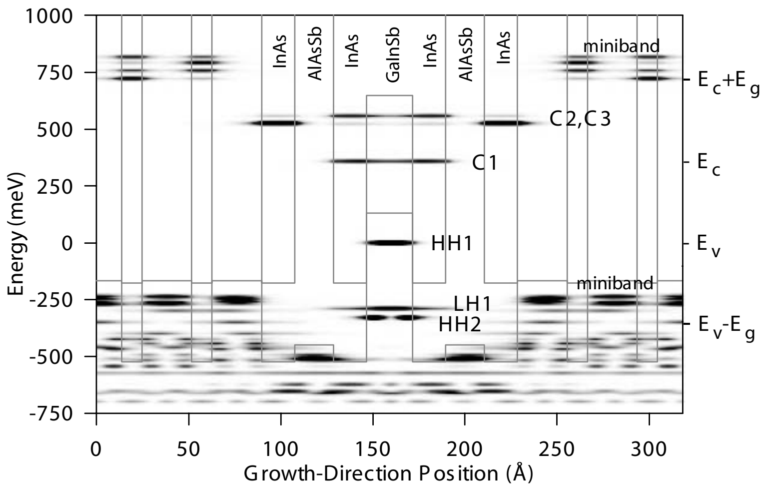


Figure 33. Spatial distribution of the zone-center superlattice states. The darkness of each point is proportional to the norm of the envelope function at that position. There is a spatial gap in the valence band structure in the central layer (below HH2) that is responsible for final-state optimization. The LH1 state in the middle is resonant with the band edge states of the superlattice, enhancing vertical transport [106]. Reprinted with permission from [105]. Copyright 2001, American Institute of Physics

bulk material (AlGaInAsSb in the four-layer superlattice), this new structure uses a two-layer superlattice based on alternating layers of InAs and AlGaSb. This superlattice barrier has minibands and minigaps, and the final-state optimization process involves placing a minigap one energy gap below the top of the valence band. Vertical transport can be enhanced by placing a miniband in between the band edge and the region of intersubband absorption and Auger recombination final states.

Figure 33 shows the norm of the zone-center envelope functions as a function of energy and growth-direction position. The first conduction band state (C1) is localized across the central InAs/GaInSb/InAs layers while the first valence band state (HH1) is strongly localized in the GaInSb layer. Below the third valence band well state (HH2), there is a 150 meV gap in the zone-center electronic structure in the central well layers, even though there are barrier states at those energies. There is nearly no spatial overlap between the initial state for intersubband absorption (the HH1 state) and the final states at the correct energy, which are localized in the barrier. Because of this, the matrix elements for transitions from the top of the valence band to these states are negligible. This "spatial gap" in the electronic structure in the spectral region one lasing energy below the top heavy hole state in the well is the origin of final-state intersubband absorption optimization in this material. Figure 34 shows the intersubband absorption and gain for this structure. The sizable gap in the intersubband absorption spectra has been tuned to lie at the energies where the largest gain is produced by this active region. The resulting slope efficiency figure of merit for this material is 99%.

Whereas the electronic structure features seen in Figure 33 are zone-center features, the spatial gap exists in the electronic structure at finite in-plane momentum as well. The key for reducing Auger recombination at room temperature is extending the gap in the final states for Auger processes further out in the in-plane direction. The final-states gap in the electronic structure at in-plane wavevectors of ~ 0.05 Å^{-1} is approximately 60 meV, which is more than twice as large as in the four-layer superlattice. In addition, the states on the lower bound of the gap (around -480 meV) are only weakly coupled to the top of the valence band due to the states' spatial isolation from each other. Figure 35 shows the electronic structure of the four-layer superlattice and the broken-gap quantum well. The darkness of each point in Figure 35 is related to the strength of the coupling to states near the top of the valence band.

The size of the final-states gap, which is greater than the energy of an optical phonon, suggests that phonon-assisted Auger recombination will be suppressed as well. The most probable Auger transitions for this material involve valence band continuum states 500 meV from the valence band edge. The net result is that the Auger recombination rate is reduced by almost two orders of magnitude with respect to the four-layer superlattice at the optimal densities for lasing.

The calculated room temperature Auger recombination rate at the optimal density for lasing is 0.06 ns^{-1}. Because this rate is much smaller than the typical 0.3 ns^{-1} rate due to Shockley-Read-Hall (SRH) processes, the system is expected to be SRH-limited at room temperature. The total carrier lifetime at the optimum density for

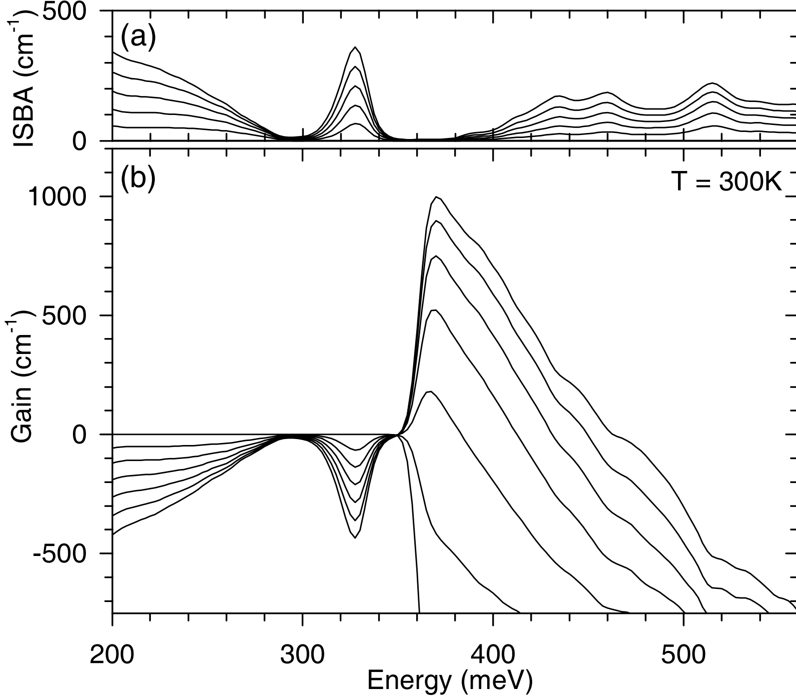


Figure 34. **(a)** Intersubband absorption and **(b)** gain spectra for the 3.4 μm broken-gap quantum well. Curves are shown for carrier densities of 1, 2, 3, 4, and 5×10^{17} cm^{-3}. There is significant suppression of intersubband absorption at photon energies near the peak of the gain spectrum, which results in a large slope-efficiency figure of merit for this active region. Reprinted with permission from [105]. Copyright 2001, American Institute of Physics

lasing is 2.5 ns, which is six times longer than for the four-layer superlattice of [27, 28].

Assuming a 3.0 ns SRH lifetime, the figure of merit for this broken-gap quantum well is 2900 $\mu\text{m}^2/A$ at 300K. This value far exceeds that of other mid-infrared materials ($40-75$ for type-I quantum wells and 260 for type-II superlattices). Because the threshold current density is inversely proportional to this figure of merit, lasing thresholds for this structure should be more than an order of magnitude smaller than current interband diodes. The large net material gain (625 cm^{-1}) at the optimal density suggests that a single quantum well active region should be sufficient for lasing in a suitably designed cavity.

5.2 Transport and Optical Cladding

An optimized cavity designed for this active region is shown in Figure 36. An essential feature of this cavity is a separate confinement region (SCR) that reduces the mode overlap with the highly doped clads and has good vertical hole mobility. The

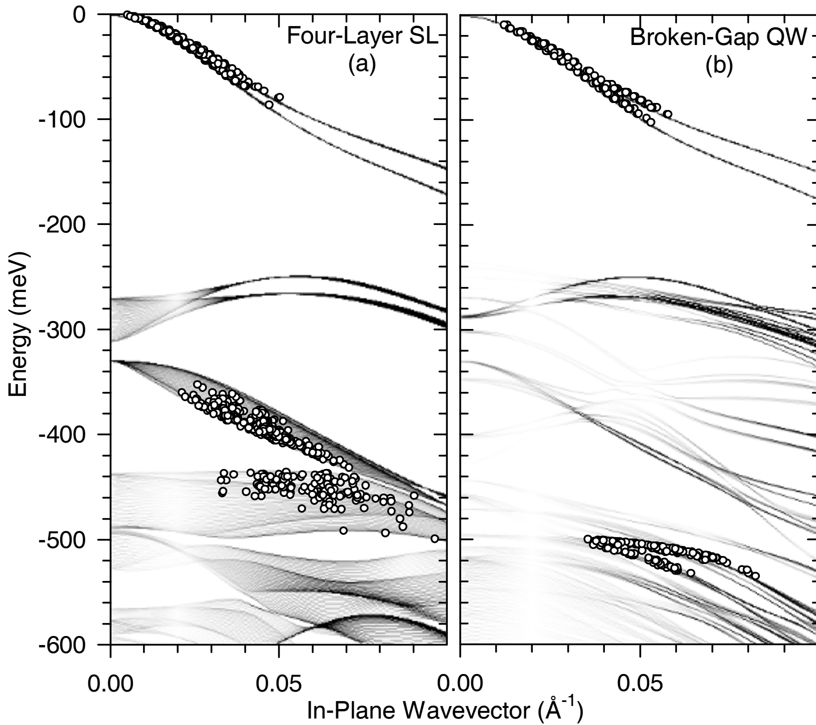


Figure 35. Initial and final energies and in-plane wavevectors for the most probable hole-hole Auger transitions (*open symbols*) overlaid on the electronic structure of the valence band of **(a)** the four-layer superlattice of [27] and **(b)** the broken-gap quantum well of [105]. In this half of the Auger process (the recombination of a conduction band electron with a hole is not shown), an electron is promoted from a state deep in the valence band to an unoccupied (hole) state near the top of the valence band. The strength of coupling to the highest-energy heavy hole band is indicated by the *darkness* of each point on the band structure

SCR of the cavity is 0.4μm thick on the n side and 0.6μm thick on the p side. Recent work in near-infrared diode development $(2.0 - 2.7\mu$m) has demonstrated the importance of such SCRs for high-temperature and high-power operation [69, 107, 108]. The material used for the SCR is the same InAs/AlGaSb superlattice that is used as the quantum-well barrier in the active region. It is lattice matched to the GaSb substrate and has a larger index of refraction ($n \sim 3.45$) than the doped clads.

The spatial gap in the valence band structure, described above as the source of intersubband absorption and Auger recombination suppression, is a consequence of the InAs/AlAsSb/InAs layers surrounding the GaInSb well. These potential barriers isolate the GaInSb well states from those of the superlattice barrier down to energies around 500 meV. The exception to this is the first light-hole state (LH1), which is localized primarily in the GaInSb layer but extends significantly into the barrier

superlattice. LH1 is an extended state that facilitates capture from the superlattice barrier into the GaInSb well, as proposed in [106]. The presence of a distributed state in the midst of a spatial gap is analogous to the conduction band structure of a resonant tunneling diode.

The vertical mobility of holes in this superlattice has been calculated using the linear Boltzmann equation, assuming an energy independent momentum scattering time, but without additional approximations to the calculated electronic structure. Scattering due to interface fluctuations in layer thicknesses is expected to dominate the momentum scattering time, and that scattering is energy independent. In the four-layer superlattice a 100 fs momentum scattering time reproduced the measured in-plane ambipolar diffusion of optically excited carriers. Assuming that the momentum scattering time for electrons and holes is the same we find a vertical hole mobility of 160 cm^2/Vs for a doping level of 1×10^{17} cm^{-3}. This vertical hole mobility, which is comparable with that of bulk InAs and is two orders of magnitude larger than that of the four-layer superlattice, ensures good vertical transport between the clads and the active region. Because the SCR is a superlattice, it is straightforward to grade its band edges by varying the superlattice period. This can be used to match the band-edges of the SCR to the injection level of the clads or to localize injected carriers near the active region.

The $AlAs_{0.08}Sb_{0.92}$ p clad has a small index of refraction ($n \sim 3.23$), a large confining potential for electrons, and a well-matched hole injection level with the

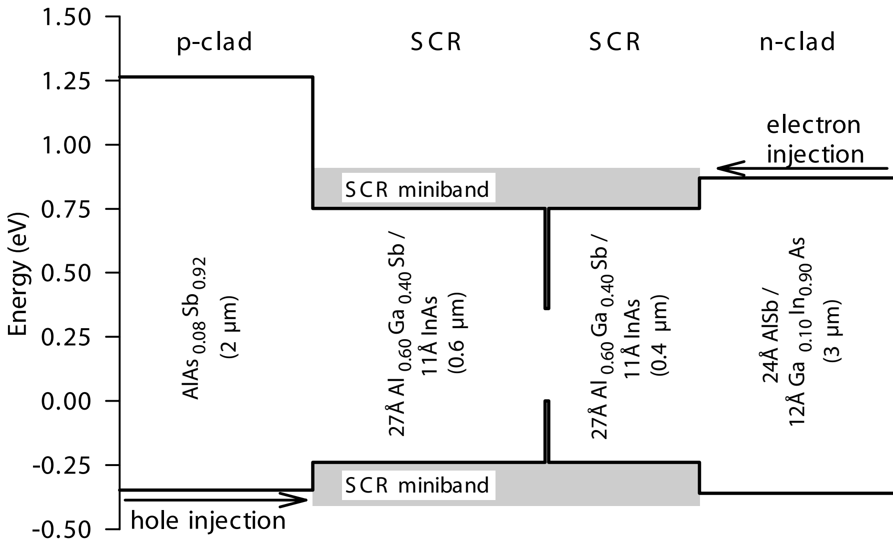


Figure 36. Schematic of the laser diode structure that includes a single quantum well active region, lightly-doped separate confinement regions (SCR), and heavily-doped clads. The estimated mode width and non-active-region modal losses for this structure are 1.3 μm and 20 cm^{-1}, respectively. Reprinted with permission from [105]. Copyright 2001, American Institute of Physics

superlattice clad. The 24Å AlSb / 12Å $Ga_{0.10}In_{0.90}As$ superlattice n clad ($n \sim 3.33$) provides electron injection matched to the miniband of the separate confinement superlattice and allows n-doping with silicon.

Based on optical mode calculations for a 2 mm long cavity with uncoated facets, we estimate the sum of the mirror losses and non-active-region modal waveguide losses for the cavity described above to be 20 cm^{-1}. The optical confinement factor is 2.1% for a single quantum well in this cavity. Because active region losses due to intersubband absorption are included in the calculation of the net active region gain, they are not included in the non-active-region cavity loss value. The calculated internal threshold current density for a single quantum well diode is 100 A/cm^2. The use of a small number of quantum wells (or perhaps even just one) significantly reduces the problem of non-uniform carrier distribution in the active region.

6 Calculating the Electronic and Optical Properties of Semiconductors

One of the most appealing aspects of mid-infrared active region design theory is that the material properties limiting performance are primarily intrinsic. Limitations due to material quality are not usually amenable to improvement from theoretical considerations. Limitations due to the fundamental electronic structure of the material, however, can be overcome with improved theoretical understanding and creative structural design.

The crucial parameters needed to understand the physical properties of semiconductor system (bulk materials, quantum wells, and superlattices) as they relate to laser active regions are the electronic structure's state energies $E_i(\boldsymbol{k})$ and momentum matrix elements between states $\boldsymbol{P}_{ij}(\boldsymbol{k})$. Here i and j label the individual bands, and $\boldsymbol{k}$ the crystal momentum. The state energies determine the densities of states, and by implication the chemical potentials associated with particular carrier densities. The momentum matrix elements determine the oscillator strength of optical transitions between subbands, whether among conduction subbands or valence subbands, or between a conduction and valence subband. Thus they directly enter into the calculations of optical gain, loss, and radiative lifetimes. They also provide a convenient way to perturbatively evaluate the nonradiative (Auger) recombination rates. Thus all the material properties that enter into the laser threshold figure of merit and the differential efficiency figure of merit derive from these two quantities.

We begin by outlining the methods for calculating the electronic state energies and momentum matrix elements, and then detail how all the essential optical and electronic properties that enter the figures of merit can be obtained from these quantities. We conclude with some example comparisons with experiments.

6.1 Electronic Structure and Momentum Matrix Elements

Superlattice electronic structure techniques for calculating $E_i(\boldsymbol{k})$ and $P_{ij}(\boldsymbol{k})$ have been quite varied, and have in general included all of the approaches taken to

calculate bulk electronic structure. These range from atomistic *ab initio* methods to *ad hoc* methods applicable to only one material system, and include tight-binding calculations [109–113], pseudopotential calculations [114–124], and most commonly envelope-function approximation restricted basis set calculations ($\boldsymbol{k} \cdot \boldsymbol{p}$) [91, 125–129]. Often a considerable computational advantage can be obtained by restricting the calculation to a narrower range of energies than the full atomic range. One common bulk approach that takes this approach is the pseudopotential method, which usually assumes a simplified frozen model for the core electrons and then calculates the valence and higher-energy states. The method we follow here takes the low-energy limit to an even greater extreme. We begin by constructing an effective low-energy Hamiltonian for the electrons in the artificially-structured semiconductor. The quantities that enter into this low-energy Hamiltonian are the band energies of constituent bulk semiconductors, a small number of bulk matrix elements between states, and valence band offsets between constituent materials. These quantities are highly physical, and could in principal be obtained from a theoretical calculation of bulk quantities via density functional theory [130] or pseudopotential [128] approaches. Typically the range of energies involved is only a few volts, extending from the highest-lying valence states (bonding p) to the lowest-lying conduction states (antibonding s), and possibly even to the next set of conduction states (antibonding p). In GaAs the energy separation of these states is less than 5 eV.

The physical quantities which enter into this low-energy theory can be obtained from *ab initio* calculations or from experimental measurements. The choice of which to use is a philosophical and practical one. As the only additional approximations that enter into the low-energy Hamiltonian are the restriction to a specified energy range, if *ab initio* calculations are used for the parameters then we have an effective *ab initio* method of calculating the electronic structure of artificially-structured semiconductors. The method is fast and efficient, and was initially put forward in detail in prescient work by Smith and Mailhiot [128]. The *ab initio* calculations, however, can at best be 100% accurate, in which case they will reproduce the experimentally-measured values for these quantities. Thus we can decouple our theory from the additional error inherent in *ab initio* calculations by taking parameters from experimental measurements.

It is sometimes the case that atomistic *ab initio* calculations will provide insight into a physical phenomenon that does not emerge from current low-energy heterostructure electronic structure calculations. One such case is the effect of interface bonds on band gaps in InAs/GaSb superlattices [131]. This indicates that an effect which was ignored in constructing the low-energy Hamiltonian is significant, and can lead to a new low-energy Hamiltonian [132]. The one concern that cannot be ignored for those working with low-energy models using empirical parameters, is that as the model becomes more complex, with more parameters to be taken from experiment or *ab initio*, the *ab initio* calculations can begin to have an advantage. Quantities which are very difficult to disentangle experimentally can often be extracted simply from *ab initio* calculations, and thus the experimental error can exceed the calculational error in current *ab initio* calculations. Very often there are some quantities that are very well known experimentally, such as band gaps, and

some quantities that are very poorly known experimentally, such as interface terms mixing heavy and light holes in InAs/GaSb superlattices. Here it is worth contemplating hybrid models, where the band gaps and momentum matrix elements are taken from experiment, but other quantities of greater experimental uncertainty are taken from calculations.

The low-energy models go under the general name of $\boldsymbol{k} \cdot \boldsymbol{p}$ methods, as this was the form of a perturbation Hamiltonian in some of the earliest exemplars of this technique. The $\boldsymbol{k} \cdot \boldsymbol{p}$ name does not emphasize, unfortunately, the physical approximations which enter into current incarnations of the theory. These are the restriction to low energies, and therefore the restriction to a limited number of bulk states in the electronic structure basis. Most current theories consider the restricted basis set in a fully non-perturbative fashion, so the name "$\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory" is no longer appropriate. In the following section, we provide a brief introduction to the use of $\boldsymbol{k} \cdot \boldsymbol{p}$ methods for bulk materials and heterostructures.

6.1.1 $\boldsymbol{k} \cdot \boldsymbol{p}$ Methods for Bulk Materials

The $\boldsymbol{k} \cdot \boldsymbol{p}$ formalism is an approach for calculating the electronic structure and momentum matrix elements throughout the Brillouin zone in terms of the band gap and the momentum matrix elements at zone center. Rather than trying to solve for the energy levels and wavefunctions directly from an approximation of crystal potentials, in $\boldsymbol{k} \cdot \boldsymbol{p}$ approaches one starts with experimental values for the energy levels and the momentum matrix elements at zone center. The only necessary experimental parameters are the band edge energy gaps and momentum matrix elements, which can be readily deduced from experimental studies (*e.g.*, photoluminescence, absorption, and effective-mass measurements). The energy level structure and matrix elements throughout the zone follow directly from the known symmetry of the band-edge states and the Schrödinger equation. Many detailed references for the application of the $\boldsymbol{k} \cdot \boldsymbol{p}$ formalism to bulk materials are available elsewhere (*e.g.* [133]).

The starting point for $\boldsymbol{k} \cdot \boldsymbol{p}$ is the time-independent, single-particle Hamiltonian in the position basis

$$H(\boldsymbol{r}) = \frac{\boldsymbol{p}^2}{2m} + \frac{\hbar}{4m^2c^2}[\boldsymbol{\sigma} \times \nabla V(\boldsymbol{r})] \cdot \boldsymbol{p} + V(\boldsymbol{r}), \tag{53}$$

where $\boldsymbol{p}$ is the momentum operator, $\boldsymbol{\sigma}$, are the Pauli matrices, and $V(\boldsymbol{r})$ is the crystal potential. The second term in Equation (53) represents the spin-orbit interaction. The spin-orbit coupling can be combined with the crystal potential term into an effective potential, $\overline{V}(\boldsymbol{r})$, so that

$$H(\boldsymbol{r}) = \frac{\boldsymbol{p}^2}{2m} + \overline{V}(\boldsymbol{r}) \tag{54}$$

Applying the Hamiltonian to a wavefunction produces

$$H(\boldsymbol{r})|l\boldsymbol{k}\rangle = E_l(\boldsymbol{k})|l\boldsymbol{k}\rangle \tag{55}$$

where $|l\boldsymbol{k}\rangle$ is the superlattice wavefunction for a particle in band l with wavevector $\boldsymbol{k}$. $E_l(\boldsymbol{k})$ is the energy of the state.

The wavefunctions with non-zero wavevector, written in terms of the zone-center wavefunctions $|n'0\rangle$,

$$|l\boldsymbol{k}\rangle = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{n'} c_{ln'}(\boldsymbol{k})|n'0\rangle, \tag{56}$$

form a complete orthonormal basis. Here $c_{ln'}(\boldsymbol{k})$ are the expansion coefficients and the summation is over all zone-center states. We will never need to solve for or otherwise determine the bulk zone-center wavefunctions. The requirements to proceed with $\boldsymbol{k}\cdot\boldsymbol{p}$ theory are the zone-center energy levels and momentum matrix elements, which can be deduced from experimental measurements.

The expansion in Equation (56) can be substituted into the eigenvalue equation, Equation (55). Expanding the momentum operator yields

$$\begin{aligned}\sum_{n'} c_{ln'}(\boldsymbol{k}) \Bigg[\Big(\frac{p^2}{2m}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Big)|n'0\rangle + \Big(\frac{\boldsymbol{p}}{m}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Big) \cdot \big(\boldsymbol{p}|n'0\rangle\big) \\ + \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\big(\frac{p^2}{2m} + \overline{V}(\boldsymbol{r})\big)|n'0\rangle \Bigg] = E_l(\boldsymbol{k})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{n'} c_{ln'}(\boldsymbol{k})|n'0\rangle.\end{aligned} \tag{57}$$

Multiplying both sides on the left with $\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\langle n, 0|$ reduces the expression to

$$\begin{aligned}\sum_{n'} c_{ln'}(\boldsymbol{k}) \left[\frac{\hbar^2 k^2}{2m}\delta_{nn'} + \frac{\hbar \boldsymbol{k}}{m} \cdot \langle n0|\boldsymbol{p}|n'0\rangle + \langle n0|\frac{p^2}{2m} + \overline{V}(\boldsymbol{r})|n'0\rangle \right] \\ = E_l(\boldsymbol{k}) \sum_{n'} c_{ln'}(\boldsymbol{k})\delta_{nn'}.\end{aligned} \tag{58}$$

If eigenstates of the total angular momentum are used as a basis for the bulk wavefunctions, then the bulk energy term is a diagonal matrix whose diagonal terms are the energies of the bulk states at zone center. In this basis, Equation (58) becomes

$$\sum_{n'} c_{ln'}(\boldsymbol{k}) \left[\Big(\frac{\hbar^2 k^2}{2m} + E_n(0)\Big)\delta_{nn'} + \frac{\hbar \boldsymbol{k}}{m} \cdot \boldsymbol{P}_{nn'}(0) \right] = E_l(\boldsymbol{k}) c_{ln}(\boldsymbol{k}), \tag{59}$$

where

$$E_n(0) = \langle n0|\frac{p^2}{2m} + \overline{V}(\boldsymbol{r})|n0\rangle, \tag{60}$$

$$\boldsymbol{P}_{nn'}(0) = \langle n0|\boldsymbol{p}|n'0\rangle. \tag{61}$$

Note that the properties of the zone-center wavefunctions only appear via the zone-center energies and momentum matrix elements. The more detailed structure of the wavefunctions themselves need not be known.

Equation (59) can be written more compactly as

$$\sum_{n'} M_{n'n}(\boldsymbol{k}) c_{n'l}(\boldsymbol{k}) = E_l(\boldsymbol{k}) c_{nl}(\boldsymbol{k}), \tag{62}$$

where the matrix M is the transpose of the term in square brackets in Equation (59). This matrix is referred to as the $\boldsymbol{k} \cdot \boldsymbol{p}$ matrix. Its eigenvalues are the state energies at wavevector $\boldsymbol{k}$ and its eigenvectors are the mixing coefficients used to describe any finite-$\boldsymbol{k}$ state in terms of the zone-center states.

To this point, the solution is exact. To make the problem tractable, a finite set of zone-center states is used to approximate the infinite basis set. This embodies the low-energy approximation involved in the $\boldsymbol{k} \cdot \boldsymbol{p}$ approach. For narrow-gap III-V materials, a minimal set of states includes the conduction band, heavy-hole, light-hole, and split-off band. Because each of these is degenerate at zone center, a total of eight bands are utilized and this approximation is referred to as 8-band $\boldsymbol{k} \cdot \boldsymbol{p}$. Unfortunately, the electronic structure generated by these eight bands has spherical symmetry (*i.e.*, $E(\boldsymbol{k})$ depends only on $|\boldsymbol{k}|$) rather than having the proper cubic symmetry, and it is also inversion symmetric (*i.e.* $E_{\uparrow}(\boldsymbol{k}) = E_{\downarrow}(\boldsymbol{k})$). Inversion asymmetry and the reduction from spherical to cubic symmetry are consequences of the coupling of the eight band-edge states with the six antibonding p-like states. This shortcoming is sometimes corrected by incorporating some of the effects of the antibonding p-states using additional terms in the 8-band $\boldsymbol{k} \cdot \boldsymbol{p}$ matrix. Some special cases of these terms have specific names, such as the Luttinger parameters that provide cubic symmetry for the holes, or the Dresselhaus term which provides a perturbative expression for the inversion asymmetry in the conduction band. It is preferable, however, to include the antibonding p-states in the $\boldsymbol{k} \cdot \boldsymbol{p}$ calculation directly by expanding the set of basis states to 14.

Shown in Figure 37 is the band structure of InAs (the dominant source of the conduction state in InAs/GaSb materials) calculated in 8-band $\boldsymbol{k} \cdot \boldsymbol{p}$, 14-band $\boldsymbol{k} \cdot \boldsymbol{p}$, and non-local pseudopotential [134] models. Whereas $\boldsymbol{k} \cdot \boldsymbol{p}$ models agree well with experiment near the zone-center (by choice of parameters), the error increases as the zone boundary is approached. The pseudopotential results have a roughly equal level of error throughout the zone (note the incorrect heavier mass at Γ). Comparison does give a guide to the degree of accuracy of the $\boldsymbol{k} \cdot \boldsymbol{p}$ models farther out in the zone. The 14-band bulk basis set that we use for our heterostructure calculations produces a better description of the conduction band, most notably out to half the zone boundary, and the light hole rapidly becomes heavier (correctly) as one moves away from the zone center. Ideally a model should retain the extreme accuracy of $\boldsymbol{k} \cdot \boldsymbol{p}$ at the zone center, but also achieve accuracy similar to or better than pseudopotential calculations near the zone boundary. This could be done by including even more states in the basis. Full-zone calculations of the energy level structure of Si and Ge used a basis set consisting of 20 states [135], and an initial attempt was made to apply it to superlattices [136]. Recently an extension to eighteen bands has been implemented in a superlattice formalism [137]. As additional states are included in the $\boldsymbol{k} \cdot \boldsymbol{p}$ basis set, the method becomes mathematically equivalent to atomistic approaches [128].

6.1.2 $\boldsymbol{K} \cdot \boldsymbol{p}$ Methods for Heterostructures

If we wish to control the properties of mid-infrared wavelength laser active regions by designing new, artificially-layered semiconductors, then we must extend the $\boldsymbol{k} \cdot \boldsymbol{p}$

formalism to quantum wells, superlattices, and other heterostructures. Fortunately, the $\boldsymbol{k} \cdot \boldsymbol{p}$ formalism can be straightforwardly expanded to cover these systems. To draw the closest analogy to bulk materials, it is easiest to examine materials that are characterized by an additional translational periodicity in addition to the bulk periodicity. For the purposes of this Chapter, we assume that the additional periodicity exists in one spatial dimension. This case covers multiple quantum wells and superlattices. Most non-periodic structures can still be treated with this approach by enclosing non-periodic portions within a larger-scale periodicity. For example, an isolated quantum well can be treated as the limiting case of a multiple quantum well structure with increasingly wider barriers. The term "superlattice" will refer to any type of periodic structure, including both multiple quantum well systems and true superlattices with strong inter-well coupling.

In order to distinguish between superlattice and bulk states, we use upper case variables (*e.g.*, L, N, and $\boldsymbol{K}$) for superlattice parameters and lowercase variables (*e.g.*, l, n, and $\boldsymbol{k}$) for bulk parameters. For a one-dimensional periodicity the struc-

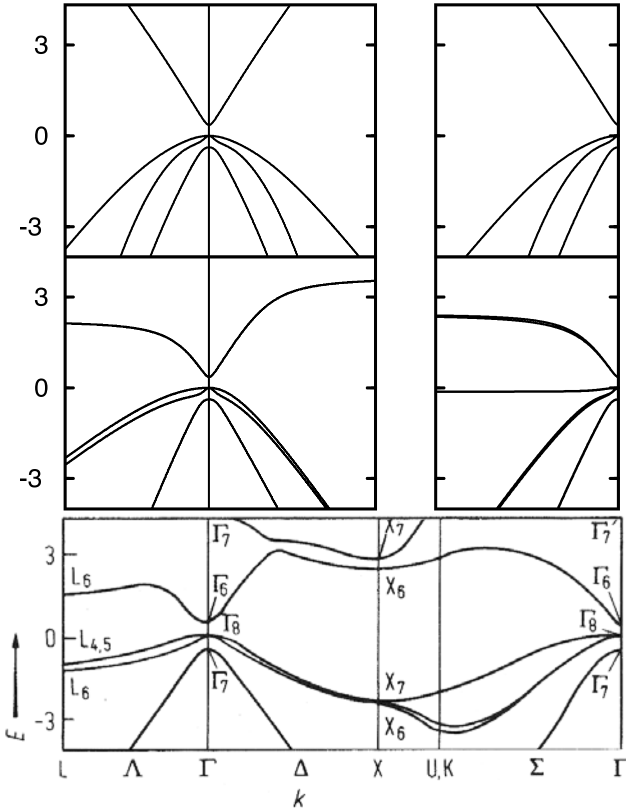


Figure 37. Band structure of InAs calculated in a (*top*) 8-band model, a (*middle*) 14-band model, and (*bottom*) non-local pseudopotential calculation. Copyright 1976, The American Physical Society [134]

tural composition varies along the growth direction, which is denoted $\hat{z}$. The translational symmetry in the in-plane directions (spanned by k_x and k_y) is not modified, so the bulk wavevectors in the in-plane direction remain good quantum numbers (*i.e.*, k_x is equivalent to K_x and k_y is equivalent to K_y).

Applying the Hamiltonian of Equation (54) to a superlattice wavefunction gives

$$H(\boldsymbol{r})\langle \boldsymbol{r}|L, \boldsymbol{K}\rangle = E_L(\boldsymbol{K})\langle \boldsymbol{r}|L, \boldsymbol{K}\rangle, \tag{63}$$

where $|L, \boldsymbol{K}\rangle$ is the superlattice wavefunction for a particle in band L with superlattice wavevector $\boldsymbol{K}$ and $E_L(\boldsymbol{K})$ is the state energy.

We start by writing the zone center superlattice wave functions in terms of the zone-center bulk wavefunctions, which form a complete orthonormal basis

$$\langle \boldsymbol{r}|L, \boldsymbol{K} = 0\rangle = \sum_{n'} F_{n'}(L, \boldsymbol{r})\langle \boldsymbol{r}|n', 0\rangle, \tag{64}$$

where $F_{n'}(L, \boldsymbol{r})$ are the expansion coefficients, and the summation is over all zone-center bulk states. For superlattice $\boldsymbol{K} \cdot \boldsymbol{p}$, the expansion coefficients are functions of position, and are referred to as envelope functions. Applying the Hamiltonian to the expanded zone center superlattice wavefunction yields

$$\begin{aligned}\sum_{n'} \Bigg[\Big(\frac{p^2}{2m} F_{n'}(L, \boldsymbol{r})\Big)\langle \boldsymbol{r}|n'0\rangle + \Big(\frac{\boldsymbol{p}}{m} F_{n'}(L, \boldsymbol{r})\Big) \cdot \big(\boldsymbol{p}\langle \boldsymbol{r}|n'0\rangle\big) \\ + F_{n'}(L, \boldsymbol{r})\Big(\frac{p^2}{2m} + \overline{V}(\boldsymbol{r})\Big)\langle \boldsymbol{r}|n'0\rangle \Bigg] \\ = E_L(0) \sum_{n'} F_{n'}(L, \boldsymbol{r})\langle \boldsymbol{r}|n'0\rangle \,. \end{aligned} \tag{65}$$

The approach simplifies considerably if the envelope functions vary slowly over the spatial scale of an atomic monolayer. Making this approximation permits a simple construction of the eigenvalue equation by operating on it from the left with

$$\int \mathrm{d}^3 r \langle n0|r\rangle, \tag{66}$$

and averaging over the unit cell, producing

$$\begin{aligned}\sum_{n'} \Bigg[\frac{p^2}{2m} F_{n'}(L, \boldsymbol{r})\delta_{nn'} + \Big(\frac{\boldsymbol{p}}{m} F_{n'}(L, \boldsymbol{r})\Big) \cdot \langle n0|\boldsymbol{p}|n'0\rangle \\ + F_{n'}(L, \boldsymbol{r})\langle n0|\frac{p^2}{2m} + \overline{V}(\boldsymbol{r})|n'0\rangle \Bigg] \\ = E_L(0) \sum_{n'} F_{n'}(L, \boldsymbol{r})\delta_{nn'} \,, \end{aligned} \tag{67}$$

which can be rewritten as

$$\sum_{n'} \left[\left(\frac{p^2}{2m} + E_n^{bulk}(\boldsymbol{r}) \right) \delta_{nn'} + \boldsymbol{P}_{nn'}^{bulk}(\boldsymbol{r}) \cdot \frac{\boldsymbol{p}}{m} \right] F_{n'}(L, \boldsymbol{r}) \\ = E_L(0) F_n(L, \boldsymbol{r}) \,, \tag{68}$$

where

$$E_n^{bulk}(\boldsymbol{r}) = \langle n0 | \frac{p^2}{2m} + \overline{V}(\boldsymbol{r}) | n0 \rangle, \tag{69}$$

$$\boldsymbol{P}_{nn'}^{bulk}(\boldsymbol{r}) = \langle n0 | \boldsymbol{p} | n'0 \rangle \,. \tag{70}$$

Just as in bulk $\boldsymbol{k} \cdot \boldsymbol{p}$ theory, in heterostructure $\boldsymbol{K} \cdot \boldsymbol{p}$ theory the bulk wavefunctions do not appear explicitly in Equation (68) for the superlattice energies and envelope functions. The separation of bulk and superlattice spatial scales using envelope functions reduces the problem to that of a continuum system with modulated bulk band edges (with an additional term that mixes the different bulk bands). In a single-band approximation, the bulk momentum matrix is zero and the problem reduces to the Krönig-Penney model. The requirement that the envelope functions vary slowly with respect to an atomic monolayer does not mean that the bulk *potentials* must vary slowly. It is normally assumed that the bulk band edges change discretely as one moves from layer to layer in the superlattice. The slowly varying envelope approximation only begins to break down when considering superlattice states away from the band edge with relatively large band indices. The validity of the approximation can be readily checked by examining the spatial gradient of the envelope functions after they are calculated. The envelope function itself approach can still be preserved, even when the envelope functions are rapidly varying, by focusing on solving the Hamiltonian exactly within the restricted basis [138]. Unfortunately in this case it is often important to know additional characteristics of the bulk basis functions beyond the zone-center energies and zone-center momentum matrix elements.

With the substitution

$$\boldsymbol{p} \to \frac{\hbar}{\mathrm{i}} \nabla \to \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial z} \hat{z} \,, \tag{71}$$

Equation (68) becomes an infinite set of coupled differential equations. Just as the restricted basis set of bulk bands simplified the bulk $\boldsymbol{k} \cdot \boldsymbol{p}$ problem into a matrix of finite size, a restricted basis of bulk bands limits the number of differential equations to a tractable number. These equations are solved at the zone center of the superlattice ($\boldsymbol{K} = 0$) to yield the zone center envelope functions and superlattice energies. This can be done using finite-difference, finite-element, Fourier transform, or transfer matrix techniques. There is a strong propensity for some of these approaches, particularly the transfer matrix approaches, to encounter spurious solutions due to numerical instabilities. Our strategy here is to Fourier transform the set of coupled differential equations, which converts the coupled differential equations into a higher-rank coupled matrix equation whose eigenvalues are straightforward to find.

When the equations are treated as continuum equations, then there are an infinite number of zone center superlattice states, which span the Hilbert space generated by the product of the bulk basis and the possible values of the real-space coordinate. The equations are not truly continuum, however, for the approximations underlying them

break down if the envelope functions vary over length scales shorter than the unit cell length. Thus it is appropriate to discard any wavelengths entering the envelope functions that are shorter than the unit cell length. This is natural to do when the continuum equations have been Fourier transformed. Simply discarding the Fourier terms corresponding to these short wavelengths is straightforward. When this is done the number of zone center superlattice states is finite, and the calculation can proceed. When a transfer matrix approach is attempted, the number of zone center superlattice states has to be truncated in another way. Typically this is done either by discarding superlattice states that only couple very weakly to the band-edge superlattice states, or by discarding all superlattice states a predetermined energy away from the band edges.

Once a restricted basis of zone-center superlattice states is known, this becomes the new basis, similar to the restricted bulk basis underlying calculations of wavefunctions and energies of finite wavevector states in bulk $\boldsymbol{k}\cdot\boldsymbol{p}$. The superlattice states at finite wavevector can then be expressed as a sum of coefficients multiplying the elements of the basis (the zone-center superlattice states), as in [127]

$$\langle \boldsymbol{r}|L\boldsymbol{K}\rangle = \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} \sum_{N'} c_{LN'}(\boldsymbol{K})\langle \boldsymbol{r}|N'0\rangle, \tag{72}$$

where $c_{LN'}(\boldsymbol{K})$ are the expansion coefficients. Inserting this again into the eigenvalue equation yields

$$\begin{aligned}\sum_{N'} c_{LN'}(\boldsymbol{K})\bigg[\Big(\frac{p^2}{2m}\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\Big)\langle \boldsymbol{r}|N'0\rangle + \Big(\frac{\boldsymbol{p}}{m}\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\Big)\cdot\Big(\boldsymbol{p}\langle \boldsymbol{r}|N'0\rangle\Big)\\ + \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\Big(\frac{p^2}{2m} + \overline{V}(\boldsymbol{r})\Big)\langle \boldsymbol{r}|N'0\rangle\bigg] = E_L(\boldsymbol{K})\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} \sum_{N'} c_{LN'}(\boldsymbol{K})\langle \boldsymbol{r}|N'0\rangle.\end{aligned} \tag{73}$$

Operating on the left with

$$\int \mathrm{d}^3r\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\langle N0|\boldsymbol{r}\rangle \tag{74}$$

gives

$$\begin{aligned}\sum_{N'} c_{LN'}(\boldsymbol{K})\bigg[\Big(\frac{\hbar^2K^2}{2m} + E_N(0)\Big)\delta_{NN'} + \frac{\hbar\boldsymbol{K}}{m}\cdot\boldsymbol{P}_{NN'}(0)\bigg]\\ = E_L(\boldsymbol{K})c_{LN}(\boldsymbol{K}),\end{aligned} \tag{75}$$

where

$$E_L(0) = \langle N0|\frac{p^2}{2m} + \overline{V}(\boldsymbol{r})|N0\rangle, \tag{76}$$

$$\boldsymbol{P}_{NN'}(0) = \langle N0|\boldsymbol{p}|N0\rangle, \tag{77}$$

are the energy levels of and momentum matrix elements coupling the zone center superlattice states.

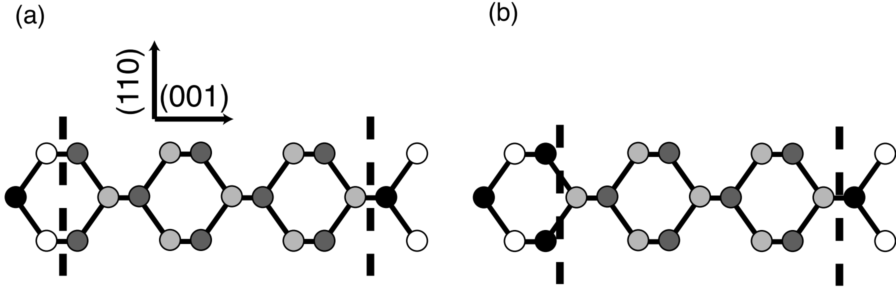


Figure 38. Asymmetric bonding character at the two interfaces of an InAs/GaSb superlattice. The shades indicate, in order of darkness, As, Sb, Ga, and In. *Dashed lines* indicate the location of the interface bonds. **(a)** two InSb-bond interfaces, **(b)** one InSb bond and one GaAs bond

The term in brackets in Equation (75) is the superlattice $\boldsymbol{K} \cdot \boldsymbol{p}$ matrix. This can be diagonalized at any point in the Brillouin zone and the set of eigenvalues comprise the superlattice energy levels. The eigenvectors are the coefficients that rotate the zone center states into the finite $\boldsymbol{K}$ states. If we define the matrix $C(\boldsymbol{K})$ to be the matrix of eigenvectors, the momentum matrix element at any point in the zone can be calculated from the zone-center momentum matrix elements by a unitary transformation:

$$\begin{aligned} \boldsymbol{P}_{NN'}(\boldsymbol{K}) &= \langle N\boldsymbol{K}|\boldsymbol{p}|N'\boldsymbol{K}\rangle \\ &= C^{\dagger}(\boldsymbol{K})\,\boldsymbol{P}_{NN'}(0)\,C(\boldsymbol{K})\,. \end{aligned} \tag{78}$$

6.1.3 Influence of Interface Bonds

The presence of unusual bonds at the interface between two layers, as shown in Figure 38(a,b) for InAs/GaSb superlattices, is a characteristic of mixed anion systems. The bonds in a (001)-grown superlattice can either progress from InAs to GaSb by including a transitional InSb bond or a GaAs bond. Neither of these interface bonds are present in the bulk of the superlattice. As the interface bonds do not have tetrahedral symmetry [they have C_{2v} symmetry, as shown in Figure 38(a,b)], these bonds do more than change the local potential for electrons and holes; they can also mix the heavy and light valence states at zone center. This mixing changes both the band gap and the valence subband splittings, although the size of the effect is controversial.

A restricted basis set $\boldsymbol{K} \cdot \boldsymbol{p}$ model can account for this effect by introducing off-diagonal terms mixing heavy and light states at zone center (first suggested in [140]). Although the original proposal of [140] provided an estimate of these values, these estimates were poor for InAs/GaSb. The value of the interface terms for InAs/GaSb can be better determined through calculations associated with the two experimental quantities of electron spin relaxation and band-gap trends for superlattices of varying layer thickness. Electron spin relaxation depends sensitively on inversion asymmetry, and the spin relaxation in a short-period InAs/GaSb superlattice is dominated by

the inversion asymmetry introduced by the interfaces [25]. In conjunction with measurements on the band gaps of several InAs/GaSb superlattices [139] the interface matrix elements mixing heavy and light holes at the two superlattice interfaces can be fully constrained [132]. Figure 39 shows the results for band-gap trends including interface terms, as well as their influence on the material absorbance.

These interface terms also significantly influence the intersubband absorption in InAs/GaSb/AlSb-based laser active regions [132], and hence should be considered carefully when optimizing the active region material. Figure 40 shows the differing valence band structures for the W-laser structure of [97], with and without interface terms. The particular spectral region of most interest for the internal loss of a laser is the gap near 230 meV. The single intersubband absorption peak apparent in the calculation without interface terms splits into two peaks, and the region near 230 meV is filled by one of the peaks. The splitting occurs because the spin-splitting in the second valence subband is dramatically increased by the interface terms, and thus one peak corresponds to transitions to one of the spin-split valence subbands, whereas the other peak corresponds to transitions to the other spin-split valence subband.

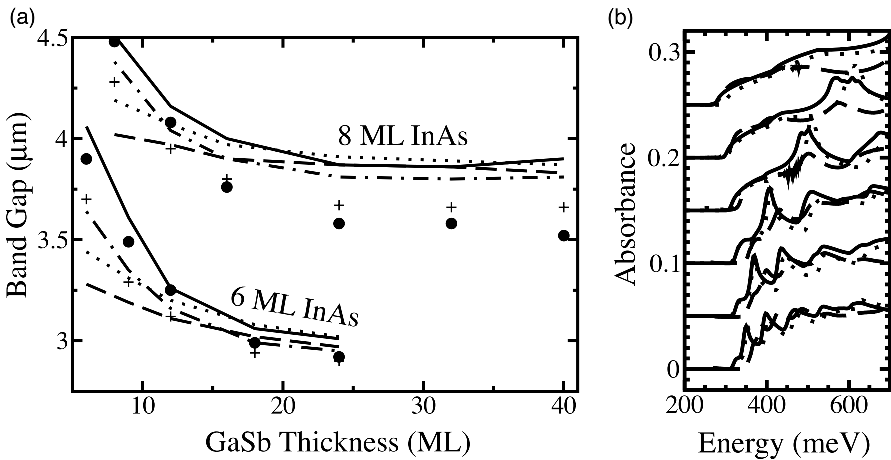


Figure 39. **(a)** Band gaps for InAs/GaSb superlattices measured by photoluminescence (*filled circles*) and calculated with the empirical pseudopotential method (EPM, +) in [131]. Restricted basis $\boldsymbol{K} \cdot \boldsymbol{p}$ calculations are shown for structures including compositional gradients, without (*dotted line*) and with interface terms (*solid line*) [132]. Also shown are restricted basis $\boldsymbol{K} \cdot \boldsymbol{p}$ calculations not including compositional gradients, with (*dot-dashed line*) and without (*dashed line*) interface terms. **(b)** Normal incidence absorption at $T = 77$K for 8ML InAs/xML GaSb, where x is (*from top down*) 8, 12, 16, 24, 32, 40. Experimental results from [139] (*dashed line*), and theoretical calculations from [132] both without (*dotted line*) and with (*solid line*) interface terms. Reprinted with permission from [132]. Copyright 2002, American Institute of Physics

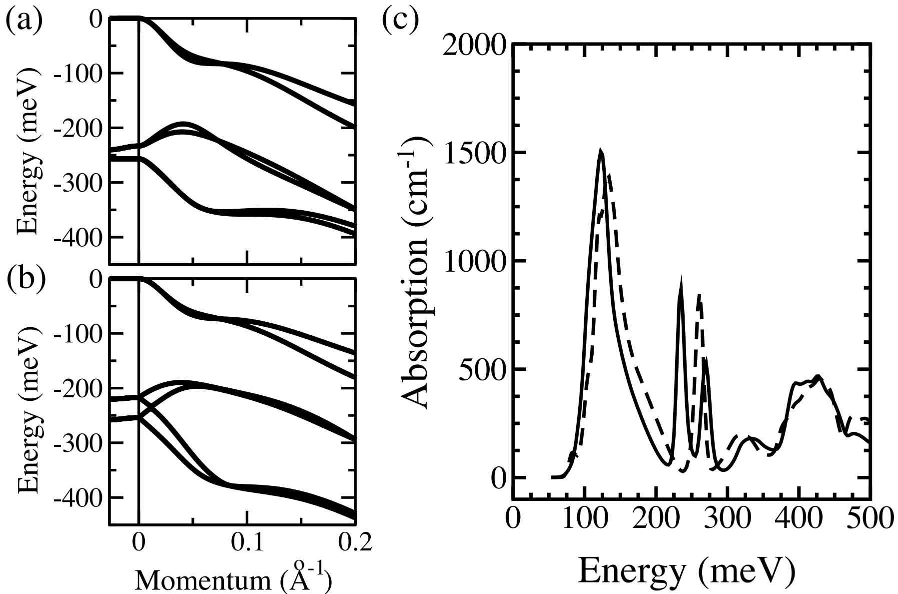


Figure 40. Band structures for the W-laser active region of [97] without **(a)** and with **(b)** interface terms. **(c)** Intersubband absorption (relevant for internal loss) calculated for normal incidence for the structure at $n = 10^{18}$ cm^{-3} and $T = 300$K neglecting interface terms (*dashed line*) and including interface terms (*solid line*). Reprinted with permission from [132]. Copyright 2002, American Institute of Physics

6.2 Comparison with Experiment

Many of the most interesting active-region materials are quantum mechanical structures that involve a number of binary, ternary, and quanternary compounds utilizing strain, quantum confinement, inter-well coupling, and band mixing effects. These factors make the analysis of these materials complex. Some of the design strategies discussed in the previous chapter rely on fine details of the electronic structure of the materials. For example, final-state Auger optimizations rely on precise placement of valence subbands within the electronic structure more than one energy gap below the top of the valence band. With any calculation method there is uncertainty, especially when there may be significant uncertainty in the input experimental parameters. So it is important to combine calculations (which can be performed quickly for many structures) with experimental techniques (which are difficult, but provide accurate information). We provide here a few examples of the comparison between calculated and measured electronic and optical material properties.

6.2.1 Nonequilibrium Absorption

The first example is calculation of the density-dependent change-in-absorption spectrum for a structure consisting of a short GaInSb/InAs superlattice segment bounded

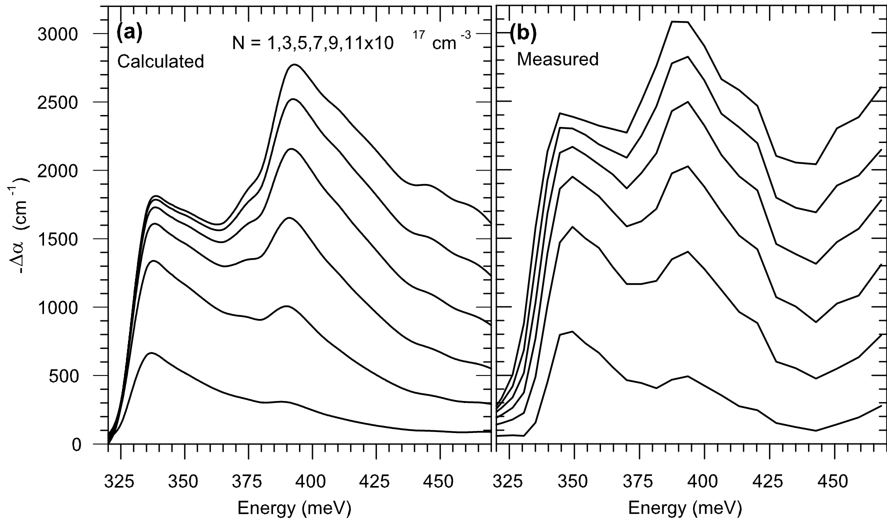


Figure 41. Measured and calculated change-in-absorption spectra for a structure grown by molecular beam epitaxy on an undoped GaSb substrate and consisting of ten 225Å quantum wells. Each well is composed of a short segment of a type-II GaInSb/InAs superlattice consisting of five 33Å $Ga_{0.75}In_{0.25}Sb$ layers and four 15Å InAs layers. The wells are separated by 400Å $Al_{0.2}Ga_{0.8}Sb$ barriers. The multiple quantum well structure is terminated at both ends by 1000Å $Al_{0.2}Ga_{0.8}Sb$ layers and capped with 30Å of GaSb

by a quaternary barrier [29]. Similar structures have been used successfully as the active region in 3–4 μm diode lasers [93]. Given the electronic structure, $E_i(\boldsymbol{k})$, and the momentum matrix elements, $P_{if}(\boldsymbol{k})$, of the superlattice material, carrier-density-dependent absorption (or gain) spectra can be calculated, as described in Section A2. Such spectra, along with density-dependent carrier lifetimes, are essential material properties for diode laser modeling and optimization. The density-dependent absorption spectra in this structure was measured using time-resolved differential transmission and a broadly-tunable, mid-infrared, optical parametric oscillator [141, 142]. As shown in Figure 41, the measured and calculated spectra are in good agreement over a wide range of carrier densities up to those expected for diode laser operation and over a wide spectral range that includes contributions from higher lying conduction subbands. These results indicate that the calculated band structure and the band filling model are sufficiently accurate to provide the basis for optimizing these structures for diode laser active regions. These particular calculations were done with an 8-band superlattice model, although subsequent calculations with a 14-band model indicated little difference.

The measured change in absorption spectra are shown for densities from 0.5–12.0×10^{17} cm^{-3}. The density axis is obtained using a capture efficiency of 73%. The spectra reveal strong bleaching of the absorption over a wide spectral range. It is also clear that the absorption change near the band edge is strongly saturated,

indicating nearly complete filling of the conduction band states in this energy range (note that, due to the extremely large density of states, valence band filling should be relatively insignificant). The spectra also illustrate a step-like behavior characteristic of quasi-two-dimensional systems. We associate the lowest energy peak near 340 meV with the blocking of transitions to the first conduction subband and the higher energy peak near 390 meV with the blocking of transitions to the second conduction subband.

The calculations, however, consistently underestimate the change in absorption. There are three possible sources for this error: 1) underestimation of the linear absorption, 2) overestimation of intersubband absorption, and 3) neglect of the bleaching of the Sommerfeld enhancement. Of these three, the latter is the most probable. Calculations of the Sommerfeld enhancement factor for InGaAs/AlInAs superlattices suggest that it is approximately 20% of the bare carrier absorption [143], which, if fully bleached, would account for most of the underestimation of the change in absorption

The second example is another comparison of a density-dependent change-in-absorption spectra, but now for a four-layer superlattice. This is the same superlattice

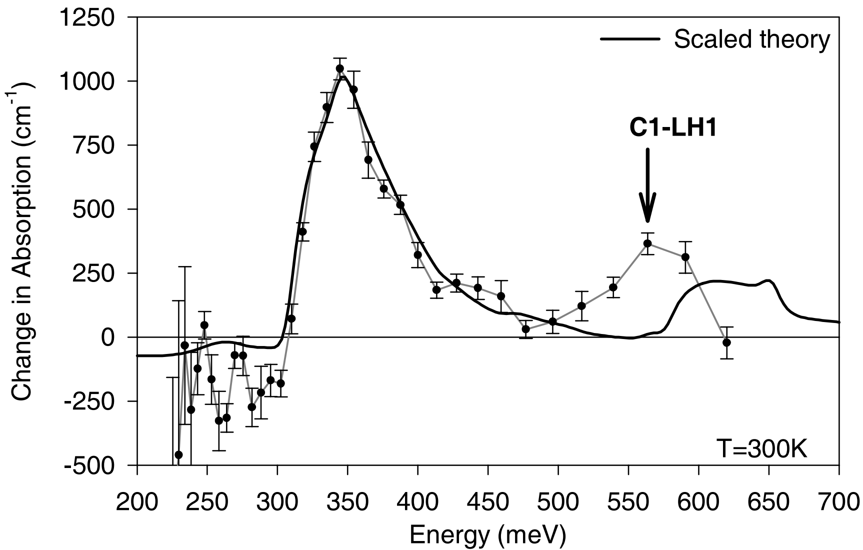


Figure 42. Measured and calculated change-in-absorption spectrum for the four-layer superlattice described in Section 3.6 and [27]. The theoretical curve, shown as the *heavy black line*, has been shifted to match the experimental band gap and scaled in amplitude. The feature at 550 meV is due to transitions from the first conduction subband to the top light-hole state. This measurement gives a measure of the heavy-hole–light-hole splitting in the material, which is a key parameter for Auger and intersubband absorption final-state optimizations. In this case, the theoretical modeling appears to overestimate the splitting by 50 meV

material that was described previously in Section 3.6. A key quantity measured here is the location of the feature around 550 meV. This feature corresponds to transitions from the first conduction band to the highest light-hole state in the valence band. Because of strain and quantum confinement there is a large splitting between the top heavy-hole and light-hole states in this material. The magnitude of the splitting is critically important to Auger suppression in the design. It can be seen in Figure 42 that, although the shape of the dominant feature matches the measured value very well, the predicted position of the conduction to light-hole transition is off by 50 meV. This discrepancy could be handled by trying to identify an improved set of input material parameters that would bring the two into agreement. A better solution, given the potential complexity of the interplay among parameters, is to be aware of the error and compensate for it in further modeling.

6.2.2 Auger Recombination

A comparison of theoretical and experimental results for four structures, shown in Figure 43, illustrates the current level of accuracy for Auger calculations based on 14-band electronic structure models [23]. The four structures are tabulated below in Table 4. Structure (a) is a "W"-laser grown on a GaSb substrate at the Naval Research

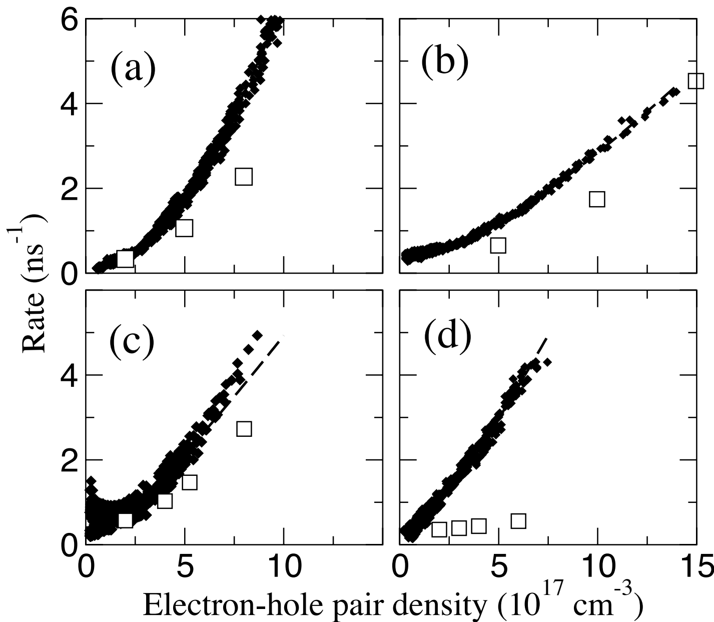


Figure 43. Experimental (*solid diamonds*) and computed (*hollow squares*) total recombination rates per electron-hole pair for the four systems listed in Table 4. The *dashed lines* are fits of the experimental data to $r = A + Bn + Cn^2$, where n is the nonequilibrium density. Reprinted with permission from [23]. Copyright 2002, American Institute of Physics

Table 4. Description of four structures considered in a comparison of theoretical and experimental Auger rates [23]

	Bandgap (μm)	Structure	Ref.
(a)	4.1	17Å InAs/26Å $Ga_{0.74}In_{0.26}Sb$/17Å InAs/43Å AlSb	[147]
(b)	3.8	15Å InAs/25Å $Ga_{0.6}In_{0.4}Sb$/15Å InAs/ 39Å $Al_{0.30}In_{0.28}Ga_{0.42}As_{0.5}Sb_{0.5}$	[27]
(c)	4.1	33Å $Ga_{0.65}In_{0.35}Sb$/20Å InAs/40Å AlSb	
(d)	3.4	18Å InAs/21Å $AlAs_{0.15}Sb_{0.85}$/18Å InAs/25Å $Ga_{0.6}In_{0.4}Sb$/ 18Å InAs/21Å $Al_{0.15}Sb_{0.85}$/18Å InAs (27Å $Al_{0.6}Ga_{0.4}Sb$/11Å $InAs)_{4.5}$	[105]

Laboratories [144]. Time resolved photoluminescence measurements [28, 145] were performed to obtain the recombination rates. This structure has been successfully employed as the active region of optically pumped lasers emitting in the mid-wavelength infrared [144, 146]. Structures (b) and (d) were grown on a GaSb substrate with a GaSb cap layer at the University of Iowa [27] and structure (c) was grown at the University of Houston. The carrier recombination rates for (b), (c) and (d) were determined via differential transmission measurements [142] as described in [27]. Each of the structures is a type-II heterostructure which should show some suppression of Auger recombination relative to bulk structures because of the improved matching of conduction and valence band-edge densities of states (or, approximately, effective masses). The first three structures, however, show only marginal suppression of Auger recombination at 300K through final-state optimization, whereas the fourth structure is predicted to have significant suppression of Auger recombination at 300K. Agreement between theory and experiment is excellent for the first three structures, but not for the fourth, indicating that final-state Auger suppression remains sensitive to theoretical or growth uncertainties.

The agreement between theoretical and observed Auger rates at the optimal carrier density for lasing is no worse than a factor of two for the first three systems, suggesting the validity and accuracy of the band structure and Auger rate computational methodologies. However, for the fourth structure there is significant disagreement. There are several possible sources of the discrepancy between theory and experiment for structure (d). The presence of Auger suppression in the predicted electronic structure of (d) is sensitive to the precise layer thicknesses of the constituent materials [105]. Hence errors either in the theoretical calculations of the band structure, or errors in growth may lead to less suppression than expected. The multitude of bands involved in the sophisticated thick structure (d) make convergence of the Monte Carlo calculations for the Auger rates delicate. Also, other Auger processes which are typically of less importance in less optimized structures may now emerge to dominate, including defect-assisted Auger or phonon-assisted Auger. Detailed studies of trends in carrier lifetimes for the broken-gap quantum well of structure (d) would be required to clarify the origin of the discrepancy.

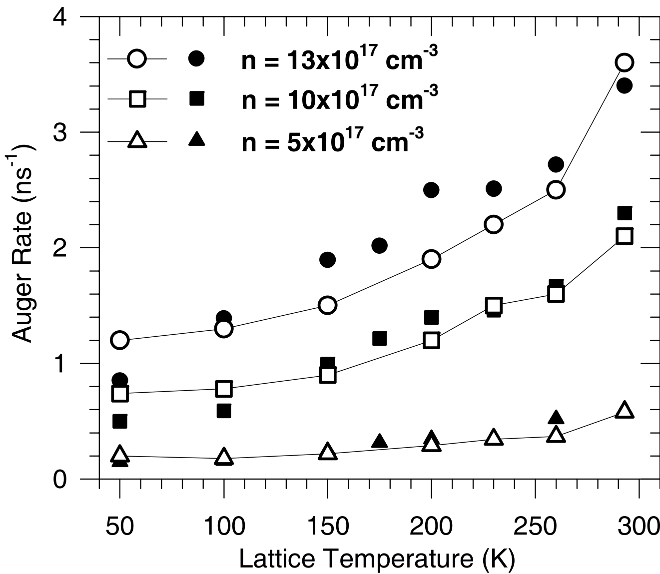


Figure 44. Measured (*solid symbols*) and calculated (*hollow symbols*) Auger recombination rates versus temperature. The calculation and measurements are described in [28]

Measurements of Auger rates have also been performed for a single structure as a function of temperature using a time-resolved photoluminescence upconversion experiment, as is shown in Figure 44 [28]. The structure being studied is a 3.8 μm four-layer superlattice. Calculated and measured Auger rates are in very good agreement for a series of three carrier densities over the temperature range from 50–300K.

7 Concluding Remarks

Sections 3–6 describe the complete path from electronic structure calculations of layered narrow-gap semiconductor materials to the threshold current density and slope efficiency of laser diodes. A wide variety of material quantities, including threshold carrier density, differential gain, intersubband absorption, and carrier recombination rates have been weighed into two figures of merit that describe the threshold current density and slope efficiency of a laser diode with a given optical loss (Section 2). Insight from these figures of merit suggest that laser diodes with very thin strained active regions are likely to perform best. The history of improvements in laser diode performance in the 2–3 μm wavelength range is substantially related to progress in making the active regions of the lasers thinner and more strained. Even though we have emphasized here the predictive nature of electronic structure calculations, it is also possible to use experimental measurements of the constituent quantities to

construct empirical figures of merit for the threshold current density and slope efficiency for a given material. These empirical values can be used to compare active region materials fairly head-to-head even in the absence of accurate theoretical calculations.

Several other quantities that can have an effect on laser performance, but are not included in these figures of merit, such as the vertical conductivity of electrons and holes and the susceptibility to filamentation, have also been calculated. Here the design criteria are qualitative — an optimized laser diode design, such as in Section 5, will have conductivities no lower than the typical conductivities of highly-performing laser diodes in other wavelength ranges (such as the near infrared).

A similar approach applies to filamentation. The susceptibility to filamentation depends on the linewidth enhancement factor, which can be calculated directly from the density-dependent optical absorption using the Kramers-Krönig relations. The linewidth enhancement factors for most type-II structures, including [27,96,97,105], are small ($\sim$1) compared to type-I active regions [15]. For these structures the linewidth enhancement factor at the optimal gain for the threshold figure of merit is quite close (typically within 5%) of the minimum linewidth enhancement factor calculated for these structures. Therefore laser active regions grown near the optimal active region thickness for minimizing the threshold current density should be near the optimal value of the linewidth enhancement factor as well.

Additional material quantities that will affect laser performance include the series resistance, the series voltage, and the thermal conductivity of various regions in the diode. As discussed previously, the importance of the series resistance can be reduced by cascading active regions. Series voltage usually comes from contact problems, and can be rendered negligible by properly preparing the contacts.

The least is known about the thermal conductivity of the various regions of the laser diode, as there are no reliable quantitative calculations of thermal conductivity through complex layered structures such as these laser diode structures. Some studies of the thermal conductivity of superlattices suggest that superlattices may have lower thermal conductivity than bulk materials [7,8]. This might play a role in the behavior of superlattice separate confinement regions, such as those in the design of Section 5. The diode lasing results of [93, 95], however, were obtained with superlattice clads similar to those presented here, and are still competitive with recent diode work in output power and operating temperature.

We would like to acknowledge conversations with T. F. Boggess, C. H. Grein, and J. Prineas and support for this work from the National Science Foundation and the National Institutes of Health.

Appendix. Derivation of Expressions for the Electronic and Optical Properties of Semiconductors

Once the electronic structure of a material and the momentum matrix elements have been calculated, a number of important properties relevant to active region operation can be deduced from them.

A1 Density of States

The density of states per unit energy for a particular band, i, can be derived directly from the electronic structure by counting the number of $\boldsymbol{k}$-space states with energy E:

$$N_i(E) = \sum_{\boldsymbol{k}} \delta(E - E_i(\boldsymbol{k}))\,, \tag{A1}$$

In the normal thermodynamic limit (crystal volume $V \to \infty$), the summation over $\boldsymbol{k}$ becomes

$$N_i(E) = V \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\, \delta(E - E_i(\boldsymbol{k}))\,, \tag{A2}$$

or

$$\rho_i(E) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\, \delta(E - E_i(\boldsymbol{k}))\,, \tag{A3}$$

where $\rho_i(E)$ is the number of states per unit volume per unit energy. The total density of states for the conduction and valence bands is then

$$\rho_c(E) = \sum_{i \in c} \rho_i(E)\,, \tag{A4}$$

$$\rho_v(E) = \sum_{i \in v} \rho_i(E)\,, \tag{A5}$$

where the sum is over all bands in the conduction or valence manifolds.

The density of states determines the relationship between the carrier density and the Fermi level. The electron and hole concentrations in a material with a Fermi level E_f is

$$n = \int_{-\infty}^{\infty} \rho_c(E) f(E)\, \mathrm{d}E\,, \tag{A6}$$

$$p = \int_{-\infty}^{\infty} \rho_v(E)[1 - f(E)]\, \mathrm{d}E\,, \tag{A7}$$

where $f(E)$ is the Fermi occupation function:

$$f(E) = \frac{1}{1 + \mathrm{e}^{(E - E_F)/k_B T}}\,. \tag{A8}$$

A2 Optical Properties

The absorption coefficient can be found from the state energies and momentum matrix elements using Fermi's Golden Rule. The rate of transitions from state i to f induced by an optical field with frequency ω is

$$r_{if} = \frac{2\pi}{\hbar} |\langle \psi_f | H_0 | \psi_i \rangle|^2 \delta(\hbar\omega - (E_f - E_i))\,, \tag{A9}$$

where r_{if} is in units of inverse seconds, ψ_i and ψ_f are the initial and final wavevectors, and H_0 is the stationary part of the interaction Hamiltonian

$$H(t) = H_0(\omega)\mathrm{e}^{-\mathrm{i}\omega t}\,. \tag{A10}$$

Here

$$H_0(\omega) = \frac{e}{m}\boldsymbol{A}(\omega)\cdot\boldsymbol{p}\,, \tag{A11}$$

where $\boldsymbol{A}(\omega)$ is the vector potential and $\boldsymbol{p}$ is the momentum operator. For a plane wave with polarization $\hat{\boldsymbol{e}}$, the vector potential per photon is (in MKS units)

$$\boldsymbol{A}(\omega) = -\mathrm{i}\sqrt{\frac{\hbar}{2\varepsilon_0\mu(\omega)\bar{\mu}(\omega)\omega}}\hat{\boldsymbol{e}}\cdot\boldsymbol{p}\,\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}\,, \tag{A12}$$

where $\mu(\omega)$ is the index of refraction, $\bar{\mu}(\omega)$ is the group index, and $\boldsymbol{q}$ is the wavevector of the optical field. The total rate of absorptive transitions is

$$r_{if} = \frac{\pi e^2}{\varepsilon_0\mu(\omega)\bar{\mu}(\omega)m^2\omega}|\langle\psi_f|\hat{\boldsymbol{e}}\cdot\boldsymbol{p}\,\mathrm{e}^{i\boldsymbol{q}\cdot\boldsymbol{r}}|\psi_i\rangle|^2\delta(\hbar\omega-(E_f-E_i))\,. \tag{A13}$$

The square of the inner product term can be simplified by replacing the initial and final wavevectors with their Bloch form,

$$\psi_f(\boldsymbol{r}) = u_f(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{A14}$$

and using the dipole approximation ($\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \to 1$). The total rate then becomes

$$\begin{aligned} r_{if} &= \frac{\pi e^2}{\varepsilon_0\mu(\omega)\bar{\mu}(\omega)m^2\omega}\int\frac{\mathrm{d}^3k}{(2\pi)^3}|\langle u_f|\hat{\boldsymbol{e}}\cdot\boldsymbol{p}|u_i\rangle|^2\delta(\hbar\omega-(E_f(\boldsymbol{k})-E_i(\boldsymbol{k})))\\ &= \frac{\pi e^2}{\varepsilon_0\mu(\omega)\bar{\mu}(\omega)m^2\omega}\int\frac{\mathrm{d}^3k}{(2\pi)^3}|\hat{\boldsymbol{e}}\cdot\boldsymbol{P}_{if}(\boldsymbol{k})|^2\delta(\hbar\omega-(E_f(\boldsymbol{k})-E_i(\boldsymbol{k})))\,. \end{aligned} \tag{A15}$$

The preceding expression is the bare rate for an electron in state i making a transition to an empty state f. This needs to be modified to account for the occupation probabilities of electrons in states i and f. The probability of a state being occupied is given by the Fermi function, $f(E)$. Modifying the expression for the transition rate by the probability of an electron being present in the initial state and the final state being empty yields

$$\begin{aligned} r_{if} = \frac{\pi e^2}{\varepsilon_0\mu(\omega)\bar{\mu}(\omega)m^2\omega}&\int\frac{\mathrm{d}^3k}{(2\pi)^3}|\hat{\boldsymbol{e}}\cdot\boldsymbol{P}_{if}(\boldsymbol{k})|^2 f(E_i(\boldsymbol{k}))(1-f(E_f(\boldsymbol{k})))\\ &\times\delta(\hbar\omega-(E_f(\boldsymbol{k})-E_i(\boldsymbol{k})))\,. \end{aligned} \tag{A16}$$

The transition rate per unit time R_{if} can be converted to the absorption coefficient per unit length by taking the difference between the rate of stimulated absorption and the rate of stimulated emission and dividing by the group velocity

$$\alpha_{if}(\omega) = \frac{r_{if} - r_{fi}}{c/\bar{\mu}(\omega)} , \tag{A17}$$

or

$$\alpha_{if}(\omega) = \frac{\pi e^2}{\varepsilon_0 \mu(\omega) m^2 c\omega} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} |\hat{\boldsymbol{e}} \cdot \boldsymbol{P}_{if}(\boldsymbol{k})|^2 (f(E_f(\boldsymbol{k})) - f(E_i(\boldsymbol{k}))) \times \delta(\hbar\omega - (E_f(\boldsymbol{k}) - E_i(\boldsymbol{k}))) . \tag{A18}$$

The above expression applies to both fundamental transitions between valence and conduction states and to intersubband absorption transitions within the conduction or valence bands. The total absorption coefficient can be obtained by summing over all state pairs where the energy of the final state, f, is greater than the energy of the initial state, i,

$$\alpha(\omega) = \sum_{f>i} \alpha_{if}(\omega) . \tag{A19}$$

This expression for the absorption coefficient requires the index of refraction, which is itself related to the absorption coefficient through the Kramers-Krönig transformations. However, the index of refraction of semiconductors near the band edge is largely due to transitions at much higher energies, so the fractional change in $n(\omega)$ near the band edge is small. Thus, it normally suffices to use a constant index of refraction in the calculation of absorption.

The absorption coefficient can be related to the imaginary part of the dielectric function

$$\varepsilon_{imag}(\omega) = \frac{\alpha(\omega)\mu(\omega)c}{\omega}\varepsilon_0 , \tag{A20}$$

or

$$\varepsilon_{imag}(\omega) = \frac{\pi e^2}{m^2\omega^2} \sum_{f>i} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} |\hat{\boldsymbol{e}} \cdot \boldsymbol{P}_{if}(\boldsymbol{k})|^2 (f(E_f(\boldsymbol{k})) - f(E_i(\boldsymbol{k}))) \times \delta(\hbar\omega - (E_f(\boldsymbol{k}) - E_i(\boldsymbol{k}))) . \tag{A21}$$

The real part of the dielectric function can then be determined from a Kramers-Krönig transformation of $\varepsilon_{imag}(\omega)$, and the absorption coefficient and index of refraction can be determined independently. However, this process requires knowledge of the electronic structure and matrix elements over a *very* wide energy range, which is often not practical.

A2.1 Differential Index of Refraction and the Linewidth Enhancement Factor

The differential index of refraction plays an important role for many laser applications, such as high-power devices and narrow-linewidth spectroscopy. Its importance is often described in terms of the linewidth enhancement factor. The starting point for the calculation of the differential index of refraction is the Kramers-Krönig relationship between the real and imaginary parts of the dielectric function

$$\varepsilon_{real}(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega'}{\omega'^2 - \omega^2}\, \varepsilon_{imag}(\omega')\, \mathrm{d}\omega', \tag{A22}$$

where P indicates the principal part of the integral. The change in $\varepsilon_{real}(\omega)$ with respect to density, n, is then

$$\frac{\mathrm{d}\varepsilon_{real}(\omega)}{\mathrm{d}n} = \frac{2}{\pi} P \int_0^\infty \frac{\omega'}{\omega'^2 - \omega^2} \frac{\mathrm{d}\varepsilon_{imag}(\omega')}{\mathrm{d}n} \mathrm{d}\omega' . \tag{A23}$$

As stated previously, the integral in Equation (A22) is very difficult to compute accurately, as $\varepsilon_{imag}(\omega')$ is required at energies significantly away from the band edge. However, the integral in Equation (A23) is much more tractable, since it only involves transition energies where $\mathrm{d}\varepsilon_{imag}(\omega')/\mathrm{d}n \neq 0$, which are primarily near the band edge. Near the band edge, where the fractional change in index is much smaller than the fractional change in the local absorption coefficient, the differential index can be written in terms of the differential absorption as

$$\frac{\mathrm{d}\mu(\omega)}{\mathrm{d}n} = \frac{c}{\pi} P \int_0^\infty \frac{\frac{\mathrm{d}\alpha(\omega')}{\mathrm{d}n}}{\omega'^2 - \omega^2} \mathrm{d}\omega' . \tag{A24}$$

The linewidth enhancement factor is defined as

$$\alpha_{lwe}(\omega) = \frac{\mathrm{d}\chi_{real}(\omega)/\mathrm{d}n}{\mathrm{d}\chi_{imag}(\omega)/\mathrm{d}n} \approx \frac{4\pi}{\lambda} \frac{\mathrm{d}\mu(\omega)/\mathrm{d}n}{\mathrm{d}\alpha(\omega)/\mathrm{d}n} . \tag{A25}$$

This can be written entirely in terms of the density-dependence of the gain spectrum

$$\alpha_{lwe}(\omega) = \frac{2\omega}{\pi} P \int_0^\infty \frac{\left[\frac{\mathrm{d}\gamma(\omega')/\mathrm{d}n}{\mathrm{d}\gamma(\omega)/\mathrm{d}n}\right]}{\omega'^2 - \omega^2}\, \mathrm{d}\omega', \tag{A26}$$

where the net gain is $\gamma(\omega) = -\alpha(\omega)$. The quantity in square braces in Equation (A26) is the differential gain spectrum *relative* to the magnitude of the differential gain at the lasing wavelength. When written in this form, the linewidth enhancement factor is a function only of the shape of the relative differential gain spectrum and not its amplitude [15].

A3 Carrier Recombination Rates

A3.1 Radiative Recombination

The expression for the radiative recombination rate can be obtained in a similar manner to Equation (A16) with the addition of a term for the optical density of states

$$\rho_{optical}(\omega) = \frac{\mu^2(\omega)\bar{\mu}(\omega)\omega^2}{\pi^2\hbar c^3} . \tag{A27}$$

The radiative recombination rate between states i and f per unit volume per unit energy is

$$R_{if}(\omega) = \frac{\pi e^2 \mu(\omega)\omega}{\varepsilon_0 m^2 \pi^2 \hbar c^3} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} |\hat{\boldsymbol{e}} \cdot \boldsymbol{P}_{if}(\boldsymbol{k})|^2 f(E_i(\boldsymbol{k}))(1 - f(E_f(\boldsymbol{k}))) \times \delta(\hbar\omega - (E_f(\boldsymbol{k}) - E_i(\boldsymbol{k}))) . \tag{A28}$$

The photoluminescence spectrum can be obtained from Equation (A28) by multiplying the rate of transitions by the energy of the emitted photon and summing over all transitions

$$I(\omega) = \frac{\pi e^2 \mu(\omega)\omega^2}{\varepsilon_0 m^2 \pi^2 c^3} \sum_{i,f} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} |\hat{\boldsymbol{e}} \cdot \boldsymbol{P}_{if}(\boldsymbol{k})|^2 f(E_i(\boldsymbol{k}))(1 - f(E_f(\boldsymbol{k}))) \times \delta(\hbar\omega - (E_f(\boldsymbol{k}) - E_i(\boldsymbol{k}))) , \tag{A29}$$

where the units of $I(\omega)\mathrm{d}(\hbar\omega)$ are watts per unit volume of emitter.

The radiative recombination rate can be obtained from Equation (A28) by integrating over all energies and summing over all transitions,

$$R = \sum_{i,f} \int \mathrm{d}(\hbar\omega) \frac{\pi e^2 \mu(\omega)\omega}{\varepsilon_0 m^2 \pi^2 \hbar c^3} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} |\hat{\boldsymbol{e}} \cdot \boldsymbol{P}_{if}(\boldsymbol{k})|^2 \times f(E_i(\boldsymbol{k}))(1 - f(E_f(\boldsymbol{k})))\delta(\hbar\omega - (E_f(\boldsymbol{k}) - E_i(\boldsymbol{k}))) . \tag{A30}$$

The units of R are carrier density per time. When the excess electron and hole densities are the same, the radiative recombination rate per carrier is

$$r = \frac{1}{\tau_{radiative}} = R/n, \tag{A31}$$

where n is the excess carrier density.

A3.2 Auger Recombination

The Auger recombination rate per unit volume can also be obtained from Fermi's golden rule:

$$R = \frac{2\pi}{\hbar V} \sum_{if} |V_{if}|^2 P(1, 1', 2, 2') \delta(E_f - E_i), \tag{A32}$$

where V is the crystal volume, i and f denote the initial (1 and 2) and final ($1'$ and $2'$) states, V_{if} is the screened Coulomb matrix element between initial and final states, and $P(1, 1', 2, 2')$ describes the occupation probability of the initial and final states. The calculations described here approximate the Coulomb matrix elements in terms of the zone-center momentum matrix elements between states using first-order $\boldsymbol{k} \cdot \boldsymbol{p}$ theory. The Auger rate per unit volume [101],

$$R \approx \frac{3 e^4 \hbar^3}{8\pi^6 \varepsilon^2 m^4} (1 - \mathrm{e}^{-(E_{F_e} - E_{F_h})/k_B T}) \int\int\int\int \mathrm{d}^3 k_1 \mathrm{d}^3 k_2 \mathrm{d}^3 k_{1'} \mathrm{d}^3 k_{2'} \times f(E_1(k_1)) f(E_2(k_2)) f(E_{1'}(k_{1'}))(1 - f(E_{2'}(k_{2'}))) \times \frac{\beta_{1,1'}(\boldsymbol{k}_1, \boldsymbol{k}_{1'})\beta_{2,2'}(\boldsymbol{k}_2, \boldsymbol{k}_{2'})}{[\lambda^2 + |\boldsymbol{k}_{2'} - \boldsymbol{k}_2|^2]^2} \delta[(\boldsymbol{k}_{1'} - \boldsymbol{k}_1) + (\boldsymbol{k}_{2'} - \boldsymbol{k}_2)] \times \delta[(E_{1'} - E_1) + (E_{2'} - E_2)], \tag{A33}$$

where E_{F_e} and E_{F_h} are the Fermi energies for the electron and hole distributions and $\beta_{if}(\boldsymbol{k}_i, \boldsymbol{k}_f)$ describes the overlap between states i and f with wavevectors $\boldsymbol{k}_i$ and $\boldsymbol{k}_f$. These terms are defined as [101]

$$\beta_{if}(\boldsymbol{k}_i, \boldsymbol{k}_f) = \left|(\boldsymbol{k}_i - \boldsymbol{k}_f) \cdot \boldsymbol{P}_{if}(\boldsymbol{k}_f)\right|^2 / \left[E_i(\boldsymbol{k_f}) - E_f(\boldsymbol{k_f})\right]^2 . \qquad \text{(A34)}$$

The integration over initial and final wavevectors constitutes a twelve-dimensional integration. The delta functions represent the conservation of energy and crystal momentum implicit in the process, and when evaluated reduce the number of integrals to eight, which are typically solved using a Monte Carlo integration method. The primary source of error in the calculation of Auger rates is in the Monte Carlo integration process. When convergence is obtained the uncertainty is on the order of a factor of three. Systems that have a large number of relevant bands (*e.g.*, quantum well structures with a large number of bands below the valence band edge of the barrier material) sometimes fail to converge. A secondary source of error is the averaging over the various spin configurations in the expression of the Coulomb matrix elements in terms of the momentum matrix elements. The error in this process is estimated to be on the order of 30% [101, 148].

The expression shown in Equation (A33) is correct for scattering events that involve only states within the first Brillouin zone. For quantum structures with growth-direction periods greater than approximately 4 nm, however, scattering events involving adjacent zones become important because of the short extent of the growth-direction zone width [27]. Scattering events involving more than the first Brillouin zone are referred to as Umklapp processes. The number of Umklapp processes that must be included in the calculation depends on the width of the growth-direction Brillouin zone; structures with a large superlattice period have narrower zones and require higher order Umklapp processes. Umklapp processes are also more of a concern in optimized structures where near-zone-center transitions have been suppressed due to the careful placement of gaps in the valence band structure.

The expression shown in Equation (A33) only describes direct Auger processes. In addition to direct Auger processes, phonon-assisted and defect-assisted Auger processes exist. The latter two involve a phonon or defect site whose presence relaxes the conservation of momentum constraint. The defect-assisted Auger recombination rate can be suppressed by reducing the density of defects, however the phonon-assisted Auger process is an intrinsic one. Fortunately, structures which have a suppressed direct Auger rate will also have a suppressed phonon-assisted rate if the final-state gap is much larger than the optical phonon energy.

A4 Carrier Mobilities

Charge carriers in a material will be accelerated by an electric field to an average drift velocity, v_d, that is proportional to the electric field, $\boldsymbol{E}$, and the mobility tensor, $\boldsymbol{\mu}$,

$$\boldsymbol{v}_d = \boldsymbol{\mu} \cdot \boldsymbol{E}. \qquad \text{(A35)}$$

The mobility tensor, $\boldsymbol{\mu}$, is related to the conductivity, $\boldsymbol{\sigma}$, by

$$\boldsymbol{\sigma} = ne\boldsymbol{\mu}\,. \tag{A36}$$

For the simple case of non-degenerate carrier densities and parabolic bands, the (isotropic) mobility,

$$\mu = \frac{e}{m^*}\langle\tau\rangle, \tag{A37}$$

where m^* is the effective mass, and $\langle\tau\rangle$ is the momentum relaxation time averaged over the carrier population. For more general cases, however, the mobility can be expressed in terms of three factors: the electronic structure, $E_i(\boldsymbol{k})$, the Fermi level, E_F, and the $\boldsymbol{k}$-dependent momentum relaxation time for band i, $\tau_i(\boldsymbol{k})$.

A particle in band i with wavevector $\boldsymbol{k}$ has a velocity

$$\boldsymbol{v}_i(\boldsymbol{k}) = \frac{1}{\hbar}\nabla_k E_i(\boldsymbol{k})\,. \tag{A38}$$

The average drift velocity can be determined by integrating $\boldsymbol{v}_i(\boldsymbol{k})$ over momentum space weighted by the probability distribution function $f(E, \boldsymbol{k})$,

$$\begin{aligned} \boldsymbol{v}_d &= \frac{\sum_i \int \frac{\mathrm{d}^3k}{(2\pi)^3} f(E_i(\boldsymbol{k}), \boldsymbol{k})\, \boldsymbol{v}_i(\boldsymbol{k})}{\sum_i \int \frac{\mathrm{d}^3k}{(2\pi)^3} f(E_i(\boldsymbol{k}), \boldsymbol{k})} \\ &= \frac{1}{n}\sum_i \int \frac{\mathrm{d}^3k}{(2\pi)^3} f(E_i(\boldsymbol{k}), \boldsymbol{k})\, \boldsymbol{v}_i(\boldsymbol{k})\,. \end{aligned} \tag{A39}$$

The probability distribution in the above expression, $f(E_i(\boldsymbol{k}), \boldsymbol{k})$, is assumed to be distorted from the equilibrium Fermi distribution, $f_0(E_i(\boldsymbol{k}))$ by the applied electric field. If the electric field is removed, the distribution is expected to relax back to the normal Fermi distribution with the band-dependent and wavevector-dependent time constant, $\tau_i(\boldsymbol{k})$. This relaxation-time approximation can be written as

$$\frac{\mathrm{d}f(E_i(\boldsymbol{k}), \boldsymbol{k})}{\mathrm{d}t} + \frac{f(E_i(\boldsymbol{k}), \boldsymbol{k}) - f_0(E_i(\boldsymbol{k}))}{\tau_i(\boldsymbol{k})} = 0\,. \tag{A40}$$

For DC fields, the solution of Equation (A40) is

$$f(E_i(\boldsymbol{k}), \boldsymbol{k}) = f_0(E_i(\boldsymbol{k})) - \tau_i(\boldsymbol{k})[\nabla_{\boldsymbol{k}} f(E_i(\boldsymbol{k}), \boldsymbol{k})] \cdot \frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t}, \tag{A41}$$

where the time derivative of the wavevector can be expressed in terms of the applied electric field:

$$\frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t} = \frac{1}{\hbar}\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \frac{e\boldsymbol{E}}{\hbar}\,. \tag{A42}$$

For small distortions from equilibrium (weak electric field strengths), we can replace $f(E_i(\boldsymbol{k}), \boldsymbol{k})$ on the right hand side of Equation (A41) with $f_0(E_i(\boldsymbol{k}))$:

$$f(E_i(\boldsymbol{k}), \boldsymbol{k}) = f_0(E_i(\boldsymbol{k})) - \tau_i(\boldsymbol{k})[\nabla_{\boldsymbol{k}} f_0(E_i(\boldsymbol{k}))] \cdot \frac{e\boldsymbol{E}}{\hbar}\,. \tag{A43}$$

We can now evaluate the drift velocity from the non-equilibrium distribution function:

$$\boldsymbol{v}_d = \frac{1}{n}\sum_i \left[\int \frac{\mathrm{d}^3k}{(2\pi)^3} f_0(\boldsymbol{k})\boldsymbol{v}_i(\boldsymbol{k}) - \int \frac{\mathrm{d}^3k}{(2\pi)^3}\boldsymbol{v}_i(\boldsymbol{k})\tau_i(\boldsymbol{k})\left(\nabla_k f_0(\boldsymbol{k})\right)\cdot\left(\frac{e\boldsymbol{E}}{\hbar}\right)\right]. \tag{A44}$$

Since $f_0(\boldsymbol{k})$ is symmetric with respect to $\boldsymbol{k}$, the first integral vanishes, giving

$$\boldsymbol{v}_d = \frac{1}{n}\sum_i \int \frac{\mathrm{d}^3k}{(2\pi)^3}\boldsymbol{v}_i(\boldsymbol{k})\tau_i(\boldsymbol{k})\left(-\nabla_k f_0(\boldsymbol{k})\right)\cdot\left(\frac{e\boldsymbol{E}}{\hbar}\right). \tag{A45}$$

The mobility is then, from Equation (A35),

$$\boldsymbol{\mu} = \frac{e}{n\hbar}\sum_i \int \frac{\mathrm{d}^3k}{(2\pi)^3}\tau_i(\boldsymbol{k})\left(\nabla_{\boldsymbol{k}} f_0(\boldsymbol{k})\right)\boldsymbol{v}_i(\boldsymbol{k}). \tag{A46}$$

The gradient of the equilibrium Fermi function can be performed analytically, yielding

$$\boldsymbol{\mu} = \frac{e}{2n\hbar k_B T}\sum_i \int \frac{\mathrm{d}^3k}{(2\pi)^3}\frac{\tau_i(\boldsymbol{k})\left[\nabla_{\boldsymbol{k}} E_i(\boldsymbol{k})\right]\left[\nabla_{\boldsymbol{k}} E_i(\boldsymbol{k})\right]}{1+\cosh\left(\frac{E_i(\boldsymbol{k})-E_F}{k_B T}\right)}. \tag{A47}$$

In order to quantify the mobility, the momentum relaxation time must be known. Several approximate expressions exist for $\tau_i(\boldsymbol{k})$ based on the relaxation mechanism (*e.g.*, acoustic or optical phonon scattering, impurity scattering, *etc.*). For room temperature calculations in type-II structures, a typical value of 100 fs can be used to obtain a mobility estimate. It is, however, best to measure the mobility in order to estimate $\tau_i(\boldsymbol{k})$ for a system. For example, if an estimate of $\tau_i(\boldsymbol{k})$ can be made by measuring the in-plane mobility of a system, then the vertical mobility (which is much more difficult to obtain experimentally) can be calculated. These calculations can also estimate the relative effect of electronic structure perturbations on the mobility.

References

1. S. F. Chichibu, S. Nakamura, eds.: *Nitride Semiconductor Blue Lasers and Light Emitting Diodes* (CRC Press, 2000)
2. E. Bründermann: *Long-Wavelength Infrared Semiconductor Lasers*, chapter 6, pp. 279–350 (Wiley, New York, 2004)
3. A. R. Beattie, P. T. Landsberg: Auger effect in semiconductors. Proc. R. Soc. London A **249**, 16–29 (1959)
4. E. O. Kane: Band structure of indium antimonide. J. Phys. Chem. Solids **1**, 249–261 (1957)
5. A. R. Adams: Band-structure engineering for low-threshold high-efficency semiconductor lasers. Electron. Lett. **22**, 249–250 (1986)
6. E. Yablonovitch, E. Kane: Reduction of lasing threshold current density by the lowering of valence band effective mass. J. Lightwave Technol. **4**, 504–506 (1986)
7. G. Chen, C. L. Tien, X. Wu, J. S. Smith: Thermal diffusivity measurement of GaAs/AlGaAs thin-film structures. J. Heat Transf. **116**, 325–331 (2001)

8. T. Borca-Tasciuc, D. Achimov, W. L. Liu, G. Chen, H.-W. Ren, C.-H. Lin, S. S. Pei: Thermal conductivity of InAs/AlSb superlattices. Microscale Therm. Eng. **5**, 225–231 (2001)
9. R. N. Hall, G. E. Fenner, J. D. Kingsley, T. J. Soltys, R. O. Carlson: Coherent light emission from GaAs junctions. Phys. Rev. Lett. **9**, 366–368 (1962)
10. M. I. Nathan, W. P. Dumke, G. Burns, J. Frederick H. Dill, G. Lasher: Stimulated emission of radiation from GaAs p-n junctions. Appl. Phys. Lett. **1**, 62–64 (1962)
11. N. Holonyak, J. S. F. Bevacqua: Coherent (visible) light emission from Ga($As_{1-x}P_x$) junctions. Appl. Phys. Lett. **1**, 82–83 (1962)
12. R. N. Hall: Injection lasers. Trans. Electron Devices **23**, 700–704 (1976)
13. I. Hayashi, M. B. Panish, P. W. Foy: A low-threshold room-temperature injection laser. J. Quant. Electron. **5**, 211–212 (1969)
14. H. Kressel, H. Nelson: Close-confinement gallium arsenide pn junction lasers with reduced optical loss at room temperature. RCA Rev. **30**, 106–113 (1969)
15. J. T. Olesberg, M. E. Flatté, T. F. Boggess: Comparison of linewidth enhancement factors in midinfrared active regions. J. Appl. Phys. **87**, 7164 (2000)
16. G. H. B. Thompson, P. A. Kirkby: (GaAl)As lasers with a heterostructure for optical confinement and additional heterojunctions for extreme carrier confinement. J. Quant. Electron. **9**, 311–318 (1973)
17. G. H. B. Thompson: *Physics of Semiconductor Laser Devices* (Wiley, New York, 1980)
18. P. W. A. McIlroy, A. Kurobe, Y. Uematsu: Analysis and application of theoretical gain curves to the design of multi-quantum-well lasers. J. Quant. Electron. **21**, 1958–1963 (1985)
19. L. A. Coldren, S. W. Corzine: *Diode Lasers and Photonic Integrated Circuits* (Wiley, New York, 1995)
20. M. E. Flatté, J. T. Olesberg, C. H. Grein: Ideal performance of cascade and noncascade intersubband and interband long-wavelength semiconductor lasers. Appl. Phys. Lett. **75**, 2020–2022 (1999)
21. J. T. Olesberg, M. E. Flatté, B. J. Brown, C. H. Grein, T. C. Hasenberg, S. A. Anson, T. F. Boggess: Optimization of active regions in midinfrared lasers. Appl. Phys. Lett. **74**, 188–190 (1999)
22. J. T. Olesberg, M. E. Flatté, B. J. Brown, T. C. Hasenberg, S. A. Anson, T. F. Boggess, C. H. Grein: Comparison of mid-infrared laser diode active regions. In *In-Plane Semiconductor Lasers III*, volume 3628 of *Proc. SPIE*, pp. 148–155 (1999)
23. C. H. Grein, M. E. Flatté, J. T. Olesberg, S. A. Anson, L. Zhang, T. F. Boggess: Auger recombination in narrow-gap semiconductor superlattices incorporating antimony. J. Appl. Phys. **92**, 7311–7316 (2002)
24. W. H. Lau, J. T. Olesberg, M. E. Flatté: Electron-spin decoherence in bulk and quantum-well zinc-blende semiconductors. Phys. Rev. B **64**, 161301(R) (2001)
25. J. T. Olesberg, W. H. Lau, M. E. Flatté, C. Yu, E. Altunkaya, E. M. Shaw, T. C. Hasenberg, T. F. Boggess: Interface contributions to spin relaxation in a short-period InAs/GaSb superlattice. Phys. Rev. B **64**, 201301(R) (2001)
26. S. A. Anson, J. T. Olesberg, M. E. Flatté, T. C. Hasenberg, T. F. Boggess: Differential gain, differential index, and linewidth enhancement factor for a 4 μm superlattice laser active layer. J. Appl. Phys. **86**, 713–718 (1999)
27. M. E. Flatté, C. H. Grein, T. C. Hasenberg, S. A. Anson, D.-J. Jang, J. T. Olesberg, T. F. Boggess: Carrier recombination rates in narrow-gap InAs/GaInSb-based superlattices. Phys. Rev. B **59**, 5745–5750 (1999)
28. D.-J. Jang, M. E. Flatté, C. H. Grein, J. T. Olesberg, T. C. Hasenberg, T. F. Boggess: Temperature dependence of Auger recombination in a multilayer narrow-band-gap superlattice. Phys. Rev. B **58**, 13047–13054 (1998)

29. J. T. Olesberg, S. A. Anson, S. W. McCahon, M. E. Flatté, T. F. Boggess, D. H. Chow, T. C. Hasenberg: Experimental and theoretical density-dependent absorption spectra in (GaInSb/InAs)/AlGaSb superlattice multiple quantum wells. Appl. Phys. Lett. **72**, 229–231 (1998)
30. M. Panish, H. Casey, Jr., S. Sumski, P. Foy: Reduction of threshold current density in GaAs-Al$_x$Ga$_{1-x}$As heterostructure lasers by separate optical and carrier confinement. Appl. Phys. Lett. **22**, 590–591 (1973)
31. W. T. Tsang: A graded-index waveguide separate-confinement laser with very low threshold and a narrow Gaussian beam. Appl. Phys. Lett. **39**, 134–137 (1981)
32. S. Hersee, M. Baldy, P. Assenat, B. de Cremoux, J. P. Duchemin: Low-threshold GRIN-SCH GaAs/GaAlAs laser structure grown by OM VPE. Electron. Lett. **18**, 618–620 (1982)
33. J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, A. Y. Cho: Quantum cascade laser. Science **264**, 553–556 (1994)
34. J. C. Garcia, E. Rosencher, P. Collot, N. Laurent, J. L. Guyaux, B. Vinter, J. Nagle: Epitaxially stacked lasers with Esaki junctions: A bipolar cascade laser. Appl. Phys. Lett. **71**, 3752–3754 (1997)
35. J. K. Kim, E. Hall, O. Sjolund, L. A. Coldren: Epitaxially-stacked multiple-active-region 1.55 μm lasers for increased differential efficiency. Appl. Phys. Lett. **74**, 3251–3253 (1999)
36. A. A. Allerman, R. M. Biefeld, S. R. Kurtz: InAsSb-based mid-infrared lasers (3.8–3.9 μm) and light-emitting diodes with AlAsSb claddings and semimetal electron injection, grown by metalorganic chemical vapor deposition. Appl. Phys. Lett. **69**, 465–467 (1996)
37. J. R. Meyer, I. Vurgaftman, R. Q. Yang, L. R. Ram-Mohan: Type-II and type-I interband cascade lasers. Electron. Lett. **32**, 45–46 (1996)
38. C. Sirtori, J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, S. N. G. Chu, A. Y. Cho: Continuous wave operation of midinfrared (7.4–8.6 μm) quantum cascade lasers up to 110 K temperature. Appl. Phys. Lett. **68**, 1745–1747 (1996)
39. J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, C. Sirtori, S. N. G. Chu, A. Y. Cho: Quantum cascade laser: temperature dependence of the performance characteristics and high T_0 operation. Appl. Phys. Lett. **65**, 2901–2903 (1994)
40. J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, A. Y. Cho: Vertical transition quantum cascade laser with Bragg confined excited state. Appl. Phys. Lett. **66**, 538–540 (1995)
41. J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, J. N. Baillargeon, A. L. Hutchinson, S.-N. G. Chu, A. Y. Cho: High power mid-infrared ($\lambda \sim 5$ μm) quantum cascade lasers operating above room temperature. Appl. Phys. Lett. **68**, 3680–3682 (1996)
42. R. Teissier, D. Barate, A. Vicet, C. Alibert, A. N. Baranov, X. Marcadet, C. Renard, M. Garcia, C. Sirtori, D. Revin, J. Cockburn: Room temperature operation of InAs/AlSb quantum cascade lasers. Appl. Phys. Lett. **85**, 167–169 (2004)
43. J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, S.-N. G. Chu, A. Y. Cho: Short wavelength ($\lambda \sim 3.4$ μm) quantum cascade laser based on strained compensated InGaAs/AlInAs. Appl. Phys. Lett. **72**, 680–682 (1998)
44. C. Sirtori, J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, A. Y. Cho: Quantum cascade laser with plasmon-enhanced waveguide operating at 8.4 μm wavelength. Appl. Phys. Lett. **66**, 3242–3244 (1995)
45. C. Sirtori, J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, A. Y. Cho: Long wavelength infrared ($\lambda \approx 11$ μm) quantum cascade lasers. Appl. Phys. Lett. **69**, 2810–2812 (1996)

46. A. Tredicucci, F. Capasso, C. Gmachl, D. L. Sivco, A. L. Hutchinson, A. Y. Cho, J. Faist, G. Scamarcio: High-power inter-miniband lasing in intrinsic superlattices. Appl. Phys. Lett. **72**, 2388–2390 (1998)
47. C. Gmachl, A. Tredicucci, F. Capasso, A. L. Hutchinson, D. L. Sivco, J. N. Baillargeon, A. Y. Cho: High-power $\lambda \sim 8$ μm quantum cascade lasers with near optimum performance. Appl. Phys. Lett. **72**, 3130–3132 (1998)
48. C. Gmachl, F. Capasso, A. Tredicucci, D. L. Sivco, A. L. Hutchinson, S. N. G. Chu, A. Y. Cho: Noncascaded intersubband injection lasers at $\lambda \sim 7.7$ μm. Appl. Phys. Lett. **73**, 3830–3832 (1998)
49. C. Sirtori, P. Kruck, S. Barbieri, P. Collot, J. Nagle, M. Beck, J. Faist, U. Oesterle: GaAs/Al_xGa_{1-x}As quantum cascade lasers. Appl. Phys. Lett. **73**, 3486–3488 (1998)
50. G. Scamarcio, C. Gmachl, F. Capasso, A. Tredicucci, A. L. Hutchinson, D. L. Sivco, A. Y. Cho: Long-wavelength ($\lambda \approx 11$ μm) interminiband Fabry-Pérot and distributed feedback quantum cascade lasers. Semicond. Sci. Technol. **13**, 1333–1339 (1998)
51. G. Strasser, S. Gianordoli, L. Hvozdara, W. Schrenk, K. Unterrainer, E. Gornik: GaAs/AlGaAs superlattice quantum cascade lasers at $\lambda \approx 13$ μm. Appl. Phys. Lett. **75**, 1345–1347 (1999)
52. A. Tredicucci, C. Gmachl, F. Capasso, D. L. Sivco, A. L. Hutchinson, A. Y. Cho: Long wavelength superlattice quantum cascade lasers at $\lambda \approx 17\mu$m. Appl. Phys. Lett. **74**, 638–640 (1999)
53. I. Melngailis: Maser action in InAs diodes. Appl. Phys. Lett. **2**, 176–178 (1963)
54. Y. W. Tung, M. L. Cohen: Relativistic band structure and electronic properties of SnTe, GeTe, and PbTe. Phys. Rev. **180**, 823–826 (1969)
55. S. E. Kohn, P. Y. Yu, Y. Petroff, Y. R. Shen, Y. Tsang, M. L. Cohen: Electronic band structure and optical properties of PbTe, PbSe, and PbS. Phys. Rev. B **8**, 1477–1488 (1973)
56. D. L. Mitchell, R. F. Wallis: Theoretical energy-band parameters for the lead salts. Phys. Rev. **151**, 581–595 (1966)
57. P. C. Findlay, C. R. Pidgeon, R. Kotitschke, A. Hollingworth, B. N. Murdin, C. J. G. M. Langerak, A. F. G. van der Meer, C. M. Ciesla, J. Oswald, A. Homer, G. Springholz, G. Bauer: Auger recombination dynamics of lead salts under picosecond free-electron-laser excitation. Phys. Rev. B **58**, 12908–12915 (1998)
58. J. S. Blakemore: Approximations for Fermi-Dirac integrals, especially the function used to describe electron density in a semiconductor. Solid-State Electron. **25**, 1067–1076 (1982)
59. K. L. Vodopyanov, H. Graener, C. C. Phillips, T. J. Tate: Picosecond carrier dynamics and studies of Auger recombination processes in indium arsenide at room temperature. Phys. Rev. B **46**, 13194–13200 (1992)
60. A. E. Bochkarev, L. M. Dolginov, A. E. Drakin, L. V. Druzhinina, P. G. Eliseev, B. N. Sverdlov: Room-temperature in GaSbAs injection-lasers at the wavelength of 1.9–2.3 μm. Sov. J. Quantum Electron. **15**, 869–871 (1985)
61. C. Caneau, A. K. Srivastava, A. G. Dentai, J. L. Zyskind, M. A. Pollack: Room-temperature GaInAsSb/AlGaAsSb DH injection lasers at 2. 2 μm. Electron. Lett. **21**, 815–817 (1985)
62. T. H. Chiu, W. T. Tsang, J. A. Ditzenberger, J. P. van der Ziel: Room-temperature operation of InGaAsSb/AlGaSb double heterostructure lasers near 2.2 μm prepared by molecular beam epitaxy. Appl. Phys. Lett. **49**, 1051–1052 (1986)
63. C. Caneau, J. L. Zyskind, J. W. Sulhoff, T. E. Glover, J. Centanni, C. A. Burrus, A. G. Dentai, M. A. Pollack: 2.2 μm GaInAsSb/AlGaAsSb injection lasers with low threshold current density. Appl. Phys. Lett. **51**, 764–766 (1987)

64. J. L. Zyskind, J. C. Dewinter, C. A. Burrus, J. C. Centanni, A. G. Dentai, M. A. Pollack: Highly uniform, high quantum efficiency GaInAsSb/AlGaAsSb double heterostructure lasers emitting at 2.2 μm. Electron. Lett. **25**, 568–570 (1989)
65. H. K. Choi, S. J. Eglash: High-efficiency high-power GaInAsSb-AlGaAsSb double-heterostructure lasers emitting at 2.3 μm. J. Quant. Electron. **27**, 1555–1559 (1991)
66. H. K. Choi, S. J. Eglash: Room-temperature CW operation at 2.2 μm of GaInAsSb/AlGaAsSb diode lasers grown by molecular beam epitaxy. Appl. Phys. Lett. **59**, 1165–1166 (1991)
67. H. K. Choi, S. J. Eglash: High-power multiple-quantum-well GaInAsSb/AlGaAsSb diode lasers emitting at 2.1 μm with low threshold current density. Appl. Phys. Lett. **61**, 1154–1156 (1992)
68. A. N. Baranov, Y. Cuminal, G. Boissier, C. Alibert, A. Joullie: Low-threshold laser diodes based on type-II GaInAsSb/GaSb quantum-wells operating at 2.36 μm at room temperature. Electron. Lett. **32**, 2279–2280 (1996)
69. D. Z. Garbuzov, H. Lee, V. Khalfin, R. Martinelli, J. C. Connolly, G. L. Belenky: 2.3–2.7 μm room temperature CW operation of InGaAsSb-AlGaAsSb broad waveguide SCH-QW diode lasers. Photon. Technol. Lett. **11**, 794–796 (1999)
70. C. Mermelstein, S. Simanowski, M. Mayer, R. Kiefer, J. Schmitz, M. Walther, J. Wagner: Room-temperature low-threshold low-loss continuous-wave operation of 2.26 μm GaInAsSb/AlGaAsSb quantum-well laser diodes. Appl. Phys. Lett. **77**, 1581–1583 (2000)
71. J. G. Kim, L. Shterengas, R. U. Martinelli, G. L. Belenky, D. Z. Garbuzov, W. K. Chan: Room-temperature 2.5 μm InGaAsSb/AlGaAsSb diode lasers emitting 1 W continuous waves. Appl. Phys. Lett. **81**, 3146–3148 (2002)
72. A. Salhi, Y. Rouillard, A. Perona, P. Grech, M. Garcia, C. Sirtori: Low-threshold GaInAsSb/AlGaAsSb quantum well laser diodes emitting near 2.3 μm. Semicond. Sci. Technol. **19**, 260–262 (2004)
73. A. Salhi, Y. Rouillard, J. Angellier, M. Garcia: Very-low-threshold 2.4 μm GaInAsSb-AlGaAsSb laser diodes operating at room temperature in the continuous-wave regime. Photon. Technol. Lett. **16**, 2424–2426 (2004)
74. P. Brosson, J. Benoit, A. Joullie, B. Sermage: Analysis of threshold current density in 2.2 μm GaInAsSb/GaAlAsSb/GaSb DH lasers. Electron. Lett. **23**, 417–419 (1987)
75. A. N. Baranov, C. Fouillant, P. Grunberg, J. L. Lazzari, S. Gaillard, A. Joullie: High temperature operation of GaInAsSb/AlGaAsSb double-heterostructure lasers emitting near 2.1 μm. Appl. Phys. Lett. **65**, 616–617 (1994)
76. G. W. Turner, H. K. Choi, M. J. Manfra: Ultralow-threshold (50 A/cm^{-1}) strained single-quantum-well GaInAsSb/AlGaAsSb lasers emitting at 2.05 μm. Appl. Phys. Lett. **72**, 876–878 (1998)
77. I. Riech, M. L. Gomez-Herrera, P. Diaz, J. G. Mendoza-Alvarez, J. L. Herrera-Perez, E. Marin: Measurement of the Auger lifetime in GaInAsSb/GaSb heterostructures using the photoacoustic technique. Appl. Phys. Lett. **79**, 964–966 (2001)
78. S. Anikeev, D. Donetsky, G. Belenky, S. Luryi, C. A. Wang, J. M. Borrego, G. Nichols: Measurement of the Auger recombination rate in p-type 0.54 eV GaInAsSb by time-resolved photoluminescence. Appl. Phys. Lett. **83**, 3317–3319 (2003)
79. H. K. Choi, S. J. Eglash, G. W. Turner: Double-heterostructure diode lasers emitting at 3 μm with a metastable GaInAsSb active layer and AlGaAsSb cladding layers. Appl. Phys. Lett. **64**, 2474–2476 (1994)
80. H. Lee, P. K. York, R. J. Menna, R. U. Martinelli, D. Z. Garbuzov, S. Y. Narayan, J. C. Connolly: Room-temperature 2.78 μm AlGaAsSb/InGaAsSb quantum-well lasers. Appl. Phys. Lett. **66**, 1942–1944 (1995)

81. H. Lee, P. K. York, R. J. Menna, R. U. Martinelli, D. Garbuzov, S. Y. Narayan: 2.78 μm InGaAsSb/AlGaAsSb multiple quantum-well lasers with metastable InGaAsSb wells grown by molecular beam epitaxy. J. Cryst. Growth **150**, 1354–1357 (1995)
82. M. Grau, C. Lin, M.-C. Amann: Low threshold 2.72 μm GaInAsSb/ AlGaAsSb multiple-quantum-well laser. Electron. Lett. **38**, 1678–1679 (2002)
83. M. Grau, C. Lin, O. Dier, M.-C. Amann: Continuous-wave GaInAsSb/AlGaAsSb type-I double quantum well lasers for 2.6 μm wavelength. Electron. Lett. **39**, 1816–1817 (2003)
84. M. Grau, C. Lin, M.-C. Amann: Room-temperature 2.81-μm continuous-wave operation of GaInAsSb-AlGaAsSb laser. Photon. Technol. Lett. **16**, 383–385 (2004)
85. A. Salhi, Y. Rouillard, J. Angellier, P. Grech, A. Vicet: 2.61 μm GaInAsSb/AlGaAsSb type I quantum well laser diodes with low threshold. Electron. Lett. **40**, 424–425 (2004)
86. C. Lin, M. Grau, O. Dier, M.-C. Amann: Low threshold room-temperature continuous-wave operation of 2.24-3.04 μm GaInAsSb/AlGaAsSb quantum-well lasers. Appl. Phys. Lett. **84**, 5088–5090 (2004)
87. D. Garbuzov, M. Maiorov, H. Lee, V. Khalfin, R. Martinelli, J. Connolly: Temperature dependence of continuous wave threshold current for 2.3–2.6 μm InGaAsSb/AlGaAsSb separate confinement heterostructure quantum well semiconductor diode lasers. Appl. Phys. Lett. **74**, 2990–2992 (1999)
88. G. A. Sai-Halasz, R. Tsu, L. Esaki: A new semiconductor superlattice. Appl. Phys. Lett. **30**, 651–653 (1977)
89. D. L. Smith, C. Mailhiot: Proposal for strained type II superlattice infrared detectors. J. Appl. Phys. **62**, 2545–2548 (1987)
90. C. H. Grein, P. M. Young, H. Ehrenreich: Theoretical performance of InAs/$In_xGa_{1-x}Sb$ superlattice-based midwave infrared lasers. J. Appl. Phys. **76**, 1940–1942 (1994)
91. M. E. Flatté, C. H. Grein, H. Ehrenreich, R. H. Miles, H. Cruz: Theoretical performance limits of 2.1–4.1 μm InAs/InGaSb, HgCdTe, and InGaAsSb lasers. J. Appl. Phys. **78**, 4552–4559 (1995)
92. S. J. Eglash, H. K. Choi: InAsSb/AlAsSb double-heterostructure diode lasers emitting at 4 μm. Appl. Phys. Lett. **64**, 833–835 (1994)
93. D. H. Chow, R. H. Miles, T. C. Hasenberg, A. R. Kost, Y. H. Zhang, H. L. Dunlap, L. West: Mid-wave infrared diode lasers based on GaInSb/InAs and InAs/AlSb superlattices. Appl. Phys. Lett. **67**, 3700–3702 (1995)
94. T. C. Hasenberg, D. H. Chow, A. R. Kost, R. H. Miles, L. West: Demonstration of 3.5 μm $Ga_{1-x}In_xSb$/InAs superlattice diode-laser. Electron. Lett. **31**, 275–276 (1995)
95. T. C. Hasenberg, R. H. Miles, A. R. Kost, L. West: Recent advances in Sb-based midwave-infrared lasers. J. Quant. Electron. **33**, 1403–1406 (1997)
96. M. E. Flatté, J. T. Olesberg, S. A. Anson, T. F. Boggess, T. C. Hasenberg, R. H. Miles, C. H. Grein: Theoretical performance of mid-infrared broken-gap multilayer superlattice lasers. Appl. Phys. Lett. **70**, 3212–3214 (1997)
97. J. R. Meyer, C. A. Hoffman, F. J. Bartoli, L. R. Ram-Mohan: Type-II quantum-well lasers for the mid-wavelength infrared. Appl. Phys. Lett. **67**, 757–759 (1995)
98. H. K. Choi, G. W. Turner: InAsSb/InAlAsSb strained quantum-well diode lasers emitting at 3.9 μm. Appl. Phys. Lett. **67**, 332–334 (1995)
99. J. R. Lindle, J. R. Meyer, C. A. Hoffman, F. J. Bartoli, G. W. Turner, H. K. Choi: Auger lifetime in InAs, InAsSb, and InAsSb-InAlAsSb quantum wells. Appl. Phys. Lett. **67**, 3153–3155 (1995)
100. C. H. Grein, P. M. Young, H. Ehrenreich: Minority carrier lifetimes in ideal InGaSb/InAs superlattices. Appl. Phys. Lett. **61**, 2905–2907 (1992)

101. C. H. Grein, P. M. Young, M. E. Flatté, H. Ehrenreich: Long wavelength InAs/InGaSb infrared detectors: Optimization of carrier lifetimes. J. Appl. Phys. **78**, 7143–7152 (1995)
102. M. E. Flatté, C. H. Grein, H. Ehrenreich: Sensitivity of optimization of mid-infrared InAs/InGaSb laser active regions to temperature and composition variations. Appl. Phys. Lett. **72**, 1424–1426 (1998)
103. M. E. Flatté, T. C. Hasenberg, J. T. Olesberg, S. A. Anson, T. F. Boggess, C. Yan, D. L. J. McDaniel: III-V interband 5.2 μm laser operating at 185 K. Appl. Phys. Lett. **71**, 3764–3766 (1997)
104. M. E. Flatté, C. H. Grein: Auger optimization in mid-infrared lasers: the importance of final-state optimization. Optics Express **2**, 131–136 (1998)
105. J. T. Olesberg, M. E. Flatté, T. C. Hasenberg, C. H. Grein: Mid-infrared InAs/GaInSb separate confinement heterostructure laser diode structures. J. Appl. Phys. **89**, 3283–3289 (2001)
106. M. E. Flatté, J. T. Olesberg, C. H. Grein: Theoretical performance of mid-infrared broken-gap multilayer superlattice lasers. In *Proceedings of the 1997 MRS Fall Symposium*, volume 484 of *Mat. Res. Soc. Proc.*, pp. 71–81 (1997)
107. D. Z. Garbuzov, R. U. Martinelli, H. Lee, P. K. York, R. J. Menna, J. C. Connolly, S. Y. Narayan: Ultralow-loss broadened-waveguide high-power 2 μm AlGaAsSb/InGaAsSb/GaSb separate-confinement quantum-well lasers. Appl. Phys. Lett. **69**, 2006–2008 (1999)
108. D. Z. Garbuzov, R. U. Martinelli, H. Lee, R. J. Menna, P. K. York, L. A. DiMarco, M. G. Harvey, R. J. Matarese, S. Y. Narayan, J. C. Connolly: 4 W quasi-continuous-wave output power from 2 μm AlGaAsSb/InGaAsSb single-quantum-well broadened waveguide laser diodes. Appl. Phys. Lett. **70**, 2931–2933 (1997)
109. J. N. Schulman, T. C. McGill: Electronic properties of the AlAs-GaAs (001) interface and superlattice. Phys. Rev. B **19**, 6341–6349 (1979)
110. J. N. Schulman, T. C. McGill: The CdTe/HgTe superlattice: proposal for a new infrared material. Appl. Phys. Lett. **34**, 663–665 (1979)
111. J. N. Schulman, T. C. McGill: Complex band structure and superlattice electronic states. Phys. Rev. B **23**, 4149–4155 (1981)
112. A. Madhukar, R. N. Nucho: The electronic structure of InAs/GaSb (001) superlattices – two dimensional effects. Solid State Commun. **32**, 331–336 (1979)
113. M. Fornari, H. H. Chen, L. Fu, R. D. Graft, D. J. Lohrmann, S. Moroni, G. P. Parravicini, L. Resca, M. A. Stroscio: Electronic structure and wave functions of interface states in HgTe-CdTe quantum wells and superlattices. Phys. Rev. B **55**, 16339–16348 (1997)
114. M. Jaros, K. B. Wong, M. A. Gell: Electronic structure of GaAs-$Ga_{1-x}Al_xAs$ quantum well and sawtooth superlattices. Phys. Rev. B **31**, 1205–1207 (1985)
115. I. Morrison, M. Jaros: Electronic and optical properties of ultrathin Si/Ge(001) superlattices. Phys. Rev. B **37**, 916–921 (1988)
116. B. M. Adderley, R. J. Turton, M. Jaros: Absorption spectra of perfect and imperfect Si/Ge superlattices. Phys. Rev. B **49**, 16622–16631 (1994)
117. H. Fu, L.-W. Wang, A. Zunger: Comparison of the $\mathbf{k} \cdot \mathbf{p}$ and the direct diagonalization approaches for describing the electronic structure of quantum dots. Appl. Phys. Lett. **71**, 3433–3435 (1997)
118. L.-W. Wang, A. Zunger: Pseudopotential-based multiband $\mathbf{k} \cdot \mathbf{p}$ method for $\sim$250,000-atom nanostructure systems. Phys. Rev. B **54**, 11417–11435 (1996)
119. C. Jenner, E. Corbin, B. M. Adderley, M. Jaros: InAs/$Ga_{1-x}In_xSb$ and InAs/$Al_{1-x}Ga_xSb$ superlattices for infrared applications. Semicond. Sci. Technol. **13**, 359–375 (1998)

120. G. C. Dente, M. L. Tilton: Pseudopotential methods for superlattices: applications to mid-infrared semiconductor lasers. J. Appl. Phys. **86**, 1420–1429 (1999)
121. G. C. Dente, M. L. Tilton: Comparing pseudopotential predictions for InAs/GaSb superlattices. Phys. Rev. B **66**, 165307 (2002)
122. R. Magri, A. Zunger: Segregation effects on the optical properties of (InAs)/(GaSb) superlattices. Physica E **13**, 325–328 (2002)
123. R. Magri, A. Zunger: Effects of interfacial atomic segregation and intermixing on the electronic properties of InAs/GaSb superlattices. Phys. Rev. B **65**, 165302 (2002)
124. R. Magri, A. Zunger: Effects of interfacial atomic segregation on optical properties of InAs/GaSb superlattices. Phys. Rev. B **64**, 081305 (2001)
125. M. E. Flatté, P. M. Young, L.-H. Peng, H. Ehrenreich: $\mathbf{k} \cdot \mathbf{p}$ superlattice theory and intersubband optical transitions. Phys. Rev. B **53**, 1963–1978 (1996)
126. M. F. H. Schuurmans, G. W. 't Hooft: Simple calculations of confinement states in a quantum well. Phys. Rev. B **31**, 8041–8048 (1985)
127. N. F. Johnson, H. Ehrenreich, P. M. Hui, P. M. Young: Electronic and optical properties of III-V and II-VI semiconductor superlattices. Phys. Rev. B **41**, 3655–3669 (1990)
128. D. L. Smith, C. Mailhiot: Theory of semiconductor superlattice electronic structure. Rev. Mod. Phys. **62**, 173–234 (1990)
129. L. R. Ram-Mohan, J. R. Meyer: Multiband finite element modeling of wavefunction-engineered electro-optical devices. Journal of Nonlinear Optical Physics and Materials **4**, 191–243 (1995)
130. M. S. Hybertsen, M. Schlüter: Theory of optical transitions in Si/Ge(001) strained-layer superlattices. Phys. Rev. B **36**, 9683–9693 (1987)
131. A. Ongstad, R. Kaspi, C. E. Moeller, M. L. Tilton, D. M. Gianardi, J. R. Chavez, G. C. Dente: Spectral blueshift and improved luminescent properties with increasing GaSb layer thickness in InAs-GaSb type-II superlattices. J. Appl. Phys. **89**, 2185–2188 (2001)
132. W. H. Lau, M. E. Flatté: Effect of interface structure on the optical properties of InAs/GaSb laser active regions. Appl. Phys. Lett. **80**, 1683–1685 (2002)
133. P. Y. Yu, M. Cardona: *Fundamentals of Semiconductors* (Springer, New York, 1995)
134. J. R. Chelikowsky, M. L. Cohen: Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors. Phys. Rev. B **14**, 556–582 (1976)
135. M. Cardona, F. H. Pollak: Energy-band structure of germanium and silicon. Phys. Rev. **142**, 530–543 (1966)
136. C. Mailhiot, T. C. McGill, D. L. Smith: New approach to the $\mathbf{k} \cdot \mathbf{p}$ theory of semiconductor superlattices. J. Vac. Sci. Technol. B **2**, 371–375 (1984)
137. I. Prevot, B. Vinter, F. H. Julien, F. Fossard, X. Marcadet: Experimental and theoretical investigation of interband and intersubband transitions in type-II InAs/AlSb superlattices. Phys. Rev. B **64**, 195318 (2001)
138. M. G. Burt: Fundamentals of envelope function theory for electronic states and photonic modes in nanostructures. J. Phys. Condens. Matter **11**, R53–R83 (1999)
139. R. Kaspi, C. Moeller, A. Ongstad, M. L. Tilton, D. Gianardi, G. Dente, P. Gopaladasu: Absorbance spectroscopy and identification of valence subband transitions in type-II InAs/GaSb superlattices. Appl. Phys. Lett. **76**, 409–411 (2000)
140. O. Krebs, P. Voisin: Giant optical anisotropy of semiconductor heterostructures with no common atom and the quantum-confined pockels effect. Phys. Rev. Lett. **77**, 1829–1832 (1996)
141. S. W. McCahon, S. A. Anson, D.-J. Jang, T. F. Boggess: Generation of 3-4 μm femtosecond pulses from a synchronously pumped, critically phase-matched $KTiOPO_4$ optical parametric oscillator. Opt. Lett. **20**, 2309–2311 (1995)

142. S. W. McCahon, S. A. Anson, D.-J. Jang, M. E. Flatté, T. F. Boggess, D. H. Chow, T. C. Hasenberg, C. H. Grein: Carrier recombination dynamics in a (GaInSb/InAs)/AlGaSb superlattice multiple quantum well. Appl. Phys. Lett. **68**, 2135–2137 (1996)
143. P. M. Young, P. M. Hui, H. Ehrenreich: Excitons and interband transitions in iii-v semiconductor superlattices. Phys. Rev. B **44**, 12969–12976 (1991)
144. W. W. Bewley, C. L. Felix, E. H. Aifer, I. Vurgaftman, L. J. Olafsen, J. R. Meyer, H. Lee, U. Martinelli, J. C. Connolly, A. R. Sugg, G. H. Olsen, M. J. Yang, B. R. Bennett, B. V. Shanabrook: Above-room-temperature optically pumped midinfrared W lasers. Appl. Phys. Lett. **73**, 3833–3835 (1998)
145. D.-J. Jang, J. T. Olesberg, M. E. Flatté, T. F. Boggess, T. C. Hasenberg: Hot carrier dynamics in a (GaInSb/InAs)/GaInAlAsSb superlattice multiple quantum well measured with mid-wave infrared, subpicosecond photoluminescence upconversion. Appl. Phys. Lett. **70**, 1125–1127 (1997)
146. W. W. Bewley, C. L. Felix, I. Vurgaftman, D. W. Stokes, E. H. Aifer, L. J. Olafsen, J. R. Meyer, M. J. Yang, B. V. Shanabrook, H. Lee, U. Martinelli, A. R. Sugg: High-temperature continuous-wave 3–6.1 μm "W" lasers with diamond-pressure-bond heat sinking. Appl. Phys. Lett. **74**, 1075–1077 (1999)
147. W. W. Bewley, I. Vurgaftman, C. L. Felix, J. R. Meyer, C.-H. Lin, D. Zhang, S. J. Murry, S. S. Pei, L. R. Ram-Mohan: Role of internal loss in limiting type-II mid-IR laser performance. J. Appl. Phys. **83**, 2384–2391 (1998)
148. A. Sugimura: Band-to-band Auger effect in GaSb and InAs lasers. J. Appl. Phys. **51**, 4405–4411 (1980)

Band Structure and High-pressure Measurements

B.N. Murdin, A.R. Adams and S.J. Sweeney
Advanced Technology Institute, University of Surrey,
Guildford, GU2 7XH, UK

1 Band Structure Calculation by the k.p Method

The electronic transitions involved in mid–infrared semiconductor optoelectronic devices are either between the conduction and valence bands, as in standard visible and near-IR technology but with different materials having appropriately narrow bandgaps, or quantum confined states within one band as in Quantum Cascade lasers [1] (QCLs) and Quantum Well Infrared Photodetectors (QWIPs).

QCLs have the obvious advantage for some applications that they can be constructed from large gap, better controlled materials (mainly InP- and GaAs-based to date) and have a peak operating frequency which is determined simply by the design of the particular quantum structure. Whilst interband devices offer an obvious way of increasing the gain by increasing the number of quantum wells in parallel, there is no such simple way of electrically pumping multiple well intersubband devices, and it was only with the advent of cascade structures that subband lasing was first achieved. The cascade offers the ability to pump multiple wells in series, and through the use of resonant tunnelling enhances the carrier injection and extraction from the lasing states. Some interband devices use a cascade series for pumping, for the reasons that firstly one can then use high barriers for electron confinement without ruining transport, and secondly that it trades current for voltage and reduces I^2R losses in contacts.

The interband alloys fall, broadly speaking, into two categories. The zinc-blende structure crystals include both the III-Vs such as (In,Ga)(Sb,As) *etc.* and the II-VIs such as (Hg,Cd)Te. The IV-VI lead salts, which have rock-salt crystal structure, include (Pb,Sn)(Te,Se). To predict the properties of these materials and the transitions it is of course necessary to understand the electronic bandstructure. There are naturally a number of sources that describe in detail the methods for calculating allowed energy states and bands in semiconductors. In this chapter we describe the most often used techniques for bulk and heterostructure band structures.

There are three external parameters that can be easily used to change the bandstructure, namely temperature, pressure and magnetic field. These are very useful for modification of the bandstructure in a controlled way, either because the bandstructure itself needs to be measured, or because an effect that depends on the bandstructure in a known way should be extracted from other competing effects. High hydrostatic pressure is particularly useful in this regard, as its effect on the bandstructure is primarily to change the fundamental energy gap while temperature also changes the electronic occupation function and magnetic field changes the symmetry. Magnetic field has the advantage that it sharpens and splits the smooth

density of states so that optical transitions are more numerous and easier to see, and this can give data that is easier to fit to.

1.1 Bulk Band Structure

One of the purposes of this chapter is to explain the important changes required when dealing with mid-infrared materials to the usual isotropic, parabolic band effective mass approximation, defined by

$$\begin{aligned} E_c &= E_g + \frac{\hbar^2 |\mathbf{k}|^2}{2m_c^*}. \\ E_v &= -\frac{\hbar^2 |\mathbf{k}|^2}{2m_v^*} \end{aligned} \tag{1}$$

Here c,v refer to the conduction and valence bands, E_g is the energy band gap, $\hbar$ is Planck's constant and $m^*_{c,v}$ are the effective masses. These are the most important bandstructure parameters, and can be used to describe a wealth of phenomena, but it is often the case that more detail is required to obtain accurate enough agreement with experiment. In particular, the type of materials used for mid-infrared applications often require an even more detailed or complex analysis than wider bandgap semiconductors, because non-parabolicity is stronger. For thin-layer quantum structures made from materials of different bandgaps, it is important to have a good model for the nonparabolic dispersion away from the band extrema. This is because states which may be near to the extremum in the higher bandgap material on one side of the interface will be in resonance with, and therefore mix with, states at high energy in the narrower bandgap material on the other side.

The main trends from one material to another, and also the best overall picture of the crystal bands, are produced with first-principles techniques such as pseudo-potential and tight-binding methods. However, these techniques are rather complicated and difficult to implement well, and often fail in detailed comparison with experiment. Other methods such as quasi-particles [2] and linear combination of atomic orbitals are also in use. It is therefore desirable to look for functions that can be used to describe the shape of the bands in terms of parameters that can be determined from experiment, as the effective mass approximation does. The most widely accessible method for calculation of the energy bandstructure in semiconductors is the **k.p** method, due to Kane [3] and others. It is a semi-empirical extension of the effective mass approximation, and as such maintains the ideas of simplicity, good agreement with the true band structure, and the smallest possible set of parameters that need to be determined by experiment (or by comparison with first principles calculations). The set of parameters is minimised by group theoretical considerations when the energy momentum relationship is expanded about a high symmetry point of the Brillouin zone (BZ). The method works well near band extrema, the point of interest when discussing optoelectronic properties.

The Schroedinger Equation for one electron is given by;

$$H\psi_{n,\mathbf{k}}(\mathbf{r})=\left[\frac{\hbar}{4m_0^2c^2}(\boldsymbol{\sigma}\times\nabla V)\bullet\mathbf{p}+V(\mathbf{r})+\frac{p^2}{2m_0}\right]\psi_{n,\mathbf{k}}(\mathbf{r})=E_{n,\mathbf{k}}\psi_{n,\mathbf{k}}(\mathbf{r}) \tag{2}$$

where the first term in the bracket is the spin-orbit coupling due to relativistic effects, $V(\mathbf{r})$ is the crystal potential, and $\boldsymbol{p}=-i\hbar\nabla$ is the momentum operator. The electronic wavefunctions in a periodic potential satisfy Bloch's theorem:

$$\psi_{n,\mathbf{k}}(\mathbf{r})=e^{i\mathbf{k}\bullet\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r}) \tag{3}$$

i.e. that they may be separated into a slowly varying envelope function, with wave–vector **k**, and a cell-periodic Bloch function $u_{n\mathbf{k}}(\mathbf{r})$. The index n runs over all bands. Substitution of Equation (3) into Equation (2) and elimination of $\exp(i\mathbf{k}.\mathbf{r})$ yields

$$\left[H_{SO}+H_0+\frac{\hbar}{m_0}\mathbf{k}\bullet\mathbf{p}+\frac{\hbar^2k^2}{2m_0}\right]u_{n,\mathbf{k}}(\mathbf{r})=E_{n,\mathbf{k}}u_{n,\mathbf{k}}(\mathbf{r}) \tag{4}$$

where H_{so} is the spin-orbit term mentioned above and $H_0 = V(\mathbf{r}) + p^2/2m_0$. In discussion of opto-electronic materials it is natural to consider the states near to the band extrema, and for zinc-blende structure materials this is at $\mathbf{k} = 0$, *i.e.* the Γ-symmetry point at the centre of the BZ. For a single band we can recover the effective mass approximation of Equation (1) using second order perturbation theory:

$$\begin{aligned}E_{n,\mathbf{k}}&=E_{n,0}+\frac{\hbar}{m_0}\mathbf{k}\bullet\mathbf{p}_{n,n}+\frac{\hbar^2k^2}{2m_0}+\frac{\hbar^2}{m_0^2}\sum_{n'\neq n}\frac{\left|\mathbf{k}\bullet\mathbf{p}_{n,n'}\right|^2}{E_{n,0}-E_{n',0}}\\&=E_{n,0}+\frac{\hbar^2k^2}{2m_0}\left[1+\frac{2}{m_0}\sum_{n'\neq n}\frac{\frac{1}{k^2}\left|\mathbf{k}\bullet\mathbf{p}_{n,n'}\right|^2}{E_{n,0}-E_{n',0}}\right]\\&=E_{n,0}+\frac{\hbar^2k^2}{2m_n^*}\end{aligned} \tag{5}$$

where we have used the fact that the second term in the first equality, which is linear in **k**, must be zero near a band extremum. In the last equality the effect of all the $\mathbf{p}_{n,n'}$ parameters has been lumped into a single parameter, m_n*. In the narrow bandgap materials of interest here $E_{n,0} - E_{n',0}$ is small and their interaction should not be treated by perturbation theory, but treated exactly. This will naturally have the effect of adding more unknown parameters, but the shape of the non-parabolicity of the bands will be reproduced more accurately. We express the u_{nk} in terms of the complete set $u_{n,0}$ at the band-edge.

$$u_{n,\mathbf{k}}(\mathbf{r})=\sum_m f_{m,\mathbf{k}}^n(\mathbf{r})u_{m,0}(\mathbf{r}) \tag{6}$$

The matrix eigen-equation for the envelope functions is therefore

$$\sum_m\left[H_{n,m}^{SO}+H_{n,m}^0+\frac{\hbar}{m_0}\mathbf{k}\bullet\mathbf{p}_{n,m}+\frac{\hbar^2k^2}{2m_0}\delta_{n,m}-E_{n,\mathbf{k}}\delta_{n,m}\right]f_{m,\mathbf{k}}^n(\mathbf{r})=0 \tag{7}$$

where, $$p_{n,m}=\langle n|p|m\rangle=\int_{\text{unit cell}}u_{n,0}^*pu_{m,0}d^3r \tag{8}$$

We choose n,m to run over the bands which interact strongly on either side of the gap. As it stands this matrix equation can be solved only if a large number (n^2)

of matrix elements for each of the three first terms are known. Furthermore, having performed the diagonalisation, we should then include the effect of far bands, again by second order perturbation theory, which introduces yet more parameters. Fortunately group theoretical arguments can reduce the number considerably.

1.1.1 Zinc-blende Crystals

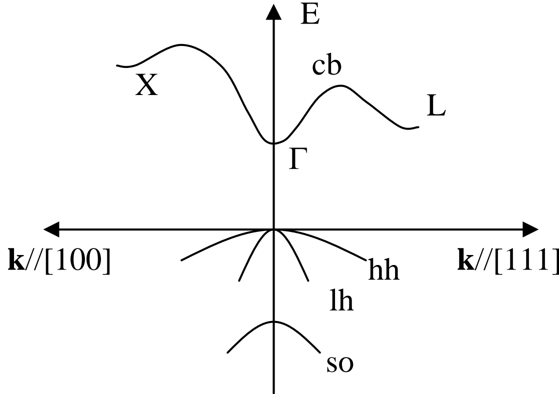


Figure 1. Schematic band structure of zinc-blende semiconductors, with symmetry points

The band structures of narrow bandgap semiconductors (*i.e.* those with mid-infrared absorption edge) with the zinc blende crystal structure, *i.e.* the III-V and II-VI show many common features. A schematic is shown in Figure 1. In these materials the conduction band arises from (and has the symmetry of) an antibonding hybrid of an S-like atomic orbital, whereas the valence band is triply degenerate and arises from bonding P-like symmetry states; X, Y or Z. One can immediately see from symmetry that the matrix element of ∇ and hence **p** between any pair of X, Y or Z is zero, and that between S and Z (or X or Y) it is a vector that points along z (or x or y respectively). In other words the components of $\mathbf{p}_{n.m}$ between these four bands are only non zero for

$$\langle S|p_x|X\rangle = \langle S|p_y|Y\rangle = \langle S|p_z|Z\rangle = i\frac{m_0}{\hbar}P \tag{9}$$

The consequence is that, for **k** along z, the **k**.**p** interaction in the matrix Equation (7) between bands originating from S and Z is large so they repel each other strongly and have light mass, whereas those originating from X and Y have heavy mass. Similarly for **k** along x or y, X or Y have light mass respectively. The parameter P is usually replaced by a parameter that has dimensions of energy,

$$E_p = \frac{2m_0P^2}{\hbar^2} \tag{10}$$

In III-V and II-VI materials E_p is typically about 20 eV.

The spin-orbit interaction term may also be reduced to a single parameter, Δ, by symmetry arguments. Generally, spin-orbit coupling increases with increasing atomic number. In zinc-blende materials its effect is to split the three-fold degeneracy at the zone center into two-fold degenerate heavy-hole and light-hole bands, and a split-off band, as shown in Figure 2. For III-V semiconductors the

magnitude of the energy splitting, equal to Δ, is determined mainly by the group V atom, and for phosphides it is of order 0.1 eV, for arsenides, 0.3 eV, whereas for antimonides it is of order 0.8 eV.

The S,X,Y,Z basis is usually changed to a linear combination which diagonalises the first two terms in Equation (6) exactly so they can be replaced by the band-edge energies $E_{n,0}$. The resulting basis set is a conduction band (cb) labelled by its group theoretical symmetry Γ_6, with total angular momentum J=1/2, doubly degenerate with spin (m_J=±1/2); a valence band labelled Γ_8 with total angular momentum J=3/2, four-fold degenerate (m_J=±3/2 or ±1/2); and a split-off (so) band labelled Γ_7 with total angular momentum J=1/2, doubly degenerate with spin (m_J=±1/2). The degeneracy of the Γ_8 is lifted at non-zero **k** by the **k.p** interaction giving a heavy hole (hh, m_J=±3/2) and light hole (lh, m_J=±1/2) band.

As mentioned above, Kane included the effects of higher bands by second order perturbation theory after diagonalisation, introducing more bandstructure parameters. Luttinger [4] showed how to include the higher bands before the diagonalisation for the lh and hh bands, with parameters that are linear combinations of Kane's, which made it easier to include a magnetic field. This approach was extended by Pidgeon and Brown [5] (PB) to include the cb and so bands, *i.e.* a total of eight bands including spin. Luttinger parameters γ_1^L and γ_2^L describe coupling of the valence band states to the higher bands, γ_3^L describes anisotropy, and a parameter F describes the interaction of the conduction band with the higher bands. Terms in the Hamiltonian involving the parameter κ^L, which gives rise to spin splitting, are zero in the absence of magnetic field.

While the band edge energies E_g and Δ, which feed directly into the Hamiltonian, are relatively easy to measure directly, the others are much more difficult to determine accurately since experiments tend to measure the effect of combinations of many parameters. The effective masses m^* and g-factors for example are sensitive to several parameters as shown below. The best results have usually come from magneto-optical data due to the fact that the smooth density of states is broken up into a series of sharp peaks that are more easily resolved (and furthermore it is sensitive to anisotropy). This is the reason for the importance of the PB model. For recent tables of the band parameters for the III-V semiconductors, see Vurgaftman and Meyer [6], and for 2-4μm quaternary alloys in particular see Adachi [7].

In the PB model with external magnetic field, the basis set becomes combinations of the Bloch functions u_n (n running over cb, hh, lh, so) with harmonic oscillator functions ϕ_l (l = 0,1,2,3…), *i.e.* potentially an infinite problem, and the components of k are replaced by creation and annihilation operators. It was shown that if inversion asymmetry terms are neglected (which are small for compound semiconductors and zero for elemental semiconductors) then the basis states only couple in groups with two from each of the four bands, so that again an 8×8 eigen problem is recovered. The model is therefore sometimes referred to as the quasi-germanium model. Originally PB considered only bulk materials, so it was natural to set k_z=0 where the density of states is peaked, and this restriction enabled the decoupling of the 8×8 matrix eigenproblem into two 4×4 problems. The full 8×8 matrix is given, *e.g.* by Weiler [8].

For the simplified case where $k_y=k_z=0$, with no external magnetic field, the PB Hamiltonian matrix is

$E_g+(2F+1)E_{k0}$			
$\sqrt{E_pE_{k0}/2}$	$-(\gamma_1+\gamma_2)E_{k0}$		
$-\sqrt{E_pE_{k0}/6}$	$-\sqrt{3}\gamma_2E_{k0}/2$	$-(\gamma_1+\gamma_2)E_{k0}$	
$\sqrt{E_pE_{k0}/3}$	$-\sqrt{3}\gamma_2E_{k0}/2$	$-\sqrt{2}\gamma_2E_{k0}$	$-\Delta-\gamma_1E_{k0}$
S	(3/2,±3/2)	(3/2,±1/2)	(1/2,±1/2)

(11)

where $\gamma_1=\gamma_1^L-E_p/3E_g$, $\gamma_{2,3}=\gamma_{2,3}^L-E_p/6E_g$ and $E_{k0}=\hbar^2k_x^2/2m_0$ We show only the lower triangle as the matrix is real and symmetric.

In general one chooses a value of **k** and finds the four eigen-values of Equation (11) which correspond to the energies of the four bands at that point in k-space. In the case when $\Delta>>E_g$ as *e.g.* for the mid-infrared material InSb, and neglecting the contribution of the far bands, the eigen-values of this matrix equation are

$$\begin{aligned}
E_c &= E_g+\frac{\hbar^2k^2}{2m_0}\left[\frac{E_g}{2}\frac{2m_0}{\hbar^2k^2}\left(\sqrt{1+2\frac{2}{E_g}\frac{2E_p}{3E_g}\frac{\hbar^2k^2}{2m_0}}-1\right)+1\right]\\
E_{hh} &= \frac{\hbar^2k^2}{2m_0}\\
E_{lh} &= -\frac{\hbar^2k^2}{2m_0}\left[\frac{E_g}{2}\frac{2m_0}{\hbar^2k^2}\left(\sqrt{1+2\frac{2}{E_g}\frac{2E_p}{3E_g}\frac{\hbar^2k^2}{2m_0}}-1\right)-1\right]\\
E_{so} &= -\Delta-\frac{\hbar^2k^2}{2m_0}\left[\frac{E_p}{3(E_g+\Delta)}-1\right]
\end{aligned} \qquad (12)$$

The terms in square brackets are the reciprocal effective masses, m_0/m^*_n. The heavy hole of Equation (12) is non-interacting, and has an effective mass equal to the free electron mass and with the wrong sign. It is only when the far-bands are taken into account ($\gamma_{1,2,3}^L \neq 0$) that the heavy hole dispersion becomes of the form $E_{hh} = -\hbar^2k^2/2m^*_{hh}$, *i.e.* the mass is rectified. The split-off band is also parabolic with a large mass. It can be seen from Equation (12) that the conduction and light hole bands repel strongly for large E_p, and have nearly equal band-edge effective mass, $m^*_{c,lh}/m_0 \sim 3E_g/2E_p$. Furthermore, the dispersion relations for the electrons and the light holes are now non-parabolic, *i.e.* the masses are k-dependent unless the momentum is small enough that the kinetic energy for the states of interest $<< E_g/2$. In mid-infrared applications E_g is a few hundred meV, and so non-parabolicity can be very important, especially at room temperature when k_BT = 25meV. Furthermore, one is often interested in Auger recombination (see below), which is a process that

involves states E_g away from the band edge where non-parabolicity is certain to play a role. For large k the dispersion becomes linear in k.
Approximate relationships for the effective masses at the band edges now including the Luttinger parameters *etc.* are

$$\frac{m_0}{m^*_e} = (1+2F) + \frac{E_P(2/3 + E_g/\Delta)}{E_g(1+E_g/\Delta)}$$

$$\frac{m_0}{m^*_\pm} = \frac{E_P}{3E_g} + \gamma_1 \pm \left[\left(\frac{E_P}{6E_g} + \gamma_2 \right)^2 + 3\left(\frac{E_P}{6E_g} + \frac{\gamma_2 + \gamma_3}{2} \right)^2 \right]^{\frac{1}{2}} \tag{13}$$

where the upper (lower) sign is for the light (heavy) hole.

The 8-band PB model compares well with more rigorous calculations up to about a quarter of the way to the BZ boundary. It is usually not necessary, and often introduces more questions than it solves, but more parameters can be introduced by including more bands in the exact diagonalisation, as in 14-band models [9,10]. The 14 band model has the advantages that it reproduces certain features of *a-priori* calculations better than the 8-band model: a) it includes spin-orbit interaction properly, b) the light hole band is very highly non-parabolic and quickly becomes heavy in mass, c) cubic anisotropy is included, d) the inversion asymmetry which gives rise to a small spin-splitting even in zero magnetic field is included, e) although no local minimum is reproduced the conduction band flattens off towards the L-point of the BZ.

A particularly difficult material to find accurate band structures for is the InAs/GaSb crossed gap system. Here the conduction band of the InAs is below the valence band of the GaSb, and electrons can be promoted from the valence band of the GaSb to the conduction band of the InAs without any thermal energy. The system is thus semi-metallic and the resulting spatial separation of the electrons and holes gives rise to strong electric fields that bend the bands in real space. Only when very thin layer superlattices are used do quantum confinement energies produce a positive bandgap and a semiconducting state. The gaps are small and it is therefore commonly employed in mid-IR applications. For a comparison of different bandstructure methods see Tawar *et al.*[11].

1.1.2 Lead-salts

The lead-salt compounds (PbTe, PbSe, PbS) and the other IV-VI alloys cover the whole wavelength region between 3 and 30 μm. They have a rock-salt crystal structure, and direct bandgaps at the four equivalent L points of the BZ. It is quite unusual to find materials with the valence band maximum away from k=0, and this is a consequence of the very large spin orbit coupling. In PbTe for example, the upper valence band and lower conduction band arise mainly from the Te (5p) and Pb (6p) states respectively, but the Pb (6s) sits just below and hybridises with the Te (5p). The band edges are singly degenerate due to the s-o coupling, with odd L_6^- symmetry in the conduction band (P_Z-like Bloch function where Z denotes the $\langle 111 \rangle$ family of crystal directions of the cubic lattice), while the valence band edge has

even L_6^+ (S-like) symmetry. The symmetry is the same for PbSe and PbS. The lead salts exhibit nearly mirror-image conduction and valence bands, *i.e.* they have almost equal masses, though they are highly anisotropic; the Fermi surfaces are ellipsoids of revolution with their principal axes along the ⟨111⟩ directions. The ellipsoids are prolate, *i.e.* the mass is heavy longitudinally along ⟨111⟩ and light transversely.

The **k.p** model for description of the band structure expanded about an L-point is due to Mitchell and Wallis [12].

$$H = \begin{bmatrix} H_{cc} & H_{cv} \\ H_{vc} & -H_{vv} \end{bmatrix} \tag{14}$$

where

$$H_{ii} = \left(E_i + \frac{\hbar^2 k_t^2}{2m_t^i} + \frac{\hbar^2 k_3^2}{2m_l^i} \right) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad H_{i \neq j} = \begin{bmatrix} \frac{\hbar}{m_0} P_l k_3 & \frac{\hbar}{m_0} P_t k_- \\ \frac{\hbar}{m_0} P_t k_+ & -\frac{\hbar}{m_0} P_l k_3 \end{bmatrix}$$

and $k_t^2 = k_1^2 + k_2^2$, $k_\pm = k_1 \pm ik_2$, $P_{l,t}$ represent the transverse and longitudinal momentum matrix elements between the band edge states, and $m_{t,l}^{c,v}$ are the far band contributions to the transverse and longitudinal band edge masses. The Cartesian axes 1,2,3 are aligned with 3 along the axis of the L valley under consideration. The eigen-values of Equation (14) are

$$E_\pm = \frac{E_g}{2} + \frac{\hbar^2}{4}\left(\frac{k_t^2}{m_t^c} + \frac{k_3^2}{m_t^c} + \frac{k_t^2}{m_t^v} + \frac{k_3^2}{m_t^v} \right) \pm \sqrt{\frac{1}{4}\left[E_g + \frac{\hbar^2}{2}\left(\frac{k_t^2}{m_t^c} + \frac{k_3^2}{m_t^c} - \frac{k_t^2}{m_t^v} - \frac{k_3^2}{m_t^v} \right) \right]^2 + E_P^t \frac{\hbar^2 k_t^2}{2m_0} + E_P^l \frac{\hbar^2 k_3^2}{2m_0}} \tag{15}$$

The dispersion relation is highly non-parabolic owing to the interaction of the conduction and valence bands, and is also highly anisotropic owing to the difference between the longitudinal and transverse parameters. The effective mass at the band edge is

$$\frac{m_0}{m_{l,t}^{\pm *}} = \frac{m_0}{m_{l,t}^{c,v}} + \frac{E_p^{l,t}}{E_g} \tag{16}$$

where the upper (lower) sign corresponds to the superscript c (v). For example in PbTe at T = 5K, E_g = 189.7 meV, and the interband momentum matrix elements are E_P^l = 1.76eV and E_P^t =6.02eV. The far band contributions to the effective masses for the conduction band are $m_l^+ = 0.505m_0$, $m_t^+ = 0.060m_0$, for the valence band $m_l^- = 0.920m_0$, $m_t^- = 0.102m_0$, and the corresponding effective masses from Equation (16) are 0.089, 0.020, 0.097, 0.024 respectively [13].

1.2 Quantum Well Structures

Just as for bulk materials, there are various methods commonly employed for the calculation of the electronic states of heterostructures where size quantisation is important. The tight binding approach gives results that reproduce the bands well across the whole BZ, whereas the envelope function approximation (EFA) is an extension of the effective mass approximation, *i.e.* it is simpler to implement and works well near the symmetry point of interest. Continuing the above discussion we describe the EFA. The method for parabolic bands is dealt with in many text books, and involves replacing k_z with d/dz and separating the Schroedinger Equation into longitudinal and transverse parts. The motion in the growth direction, z, is quantised, whereas in the plane of the well the motion is free, and these directions are independent and separable. The transverse part gives rise to a parabolic term as before, and the longitudinal part must be solved by choosing trial values for the energy and integrating the left hand side of the Schrodinger equation to see if that energy gives a valid solution. An arbitrary point along z may be chosen for initialising the wavefunction and starting the integration, usually the left-most interface, and normalisation is done after solutions are found. The integration may be done analytically if the bands are flat in real space, and gives rise to purely sinusoidal or purely evanescent waves, but the solution for the energy is transcendental and requires numerical solution.

For coupled bands, one can use a very similar principle for finding the allowed energy states, though the procedure is slightly more complicated. The transfer matrix method described here is due to Ram-Mohan *et al.*[14]. Replacing k_z with d/dz in Equation (7) one finds a second order differential matrix equation of the form;

$$\sum_m \left[\begin{array}{l} H^{SO}_{n,m} + H^{0}_{n,m} \\ -\dfrac{i\hbar}{m_0}\left(k_x p^x_{n,m} + k_y p^y_{n,m} + p^z_{n,m}\dfrac{d}{dz} \right) \\ +\dfrac{\hbar^2}{2m_0}\left(k_x^2 + k_y^2 + \dfrac{d^2}{dz^2} \right)\delta_{n,m} \end{array} \right] f^n_{m,\mathbf{k}//,nz}(\mathbf{r}) = E_{n,\mathbf{k}//,nz} f^n_{n,\mathbf{k}//,nz}(\mathbf{r})$$

$$= \underline{\underline{\mathbf{A}}}\underline{\mathbf{f}}'' + \underline{\underline{\mathbf{B}}}\underline{\mathbf{f}}' + \underline{\underline{\mathbf{C}}}\underline{\mathbf{f}} = E\underline{\mathbf{f}} \tag{17}$$

where for a given material the coefficients are fixed for particular values of $\mathbf{k}_{\|} = (k_x, k_y)$ and a trial value for the energy is chosen. In Equation (7) the eigenvalues and eigenvectors have a subscript **k** to indicate that k_x, k_y and k_z are the quantum numbers. In Equation (17) $\mathbf{k}_{\|}$ and the subband index, n_z, are the quantum numbers. The first boundary condition for the wavefunction $\underline{\mathbf{f}}$ at the heterojunction interface is that it is continuous (or there would be infinite current there), and integrating Equation (17) across the junction **Af'+Bf** is also continuous. One can therefore notice that

$$\begin{bmatrix} -\underline{\underline{\mathbf{A}}}^{-1}\underline{\underline{\mathbf{B}}} & \underline{\underline{\mathbf{A}}}^{-1} \\ E\underline{\underline{\mathbf{I}}}-\underline{\underline{\mathbf{C}}} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \underline{\underline{\mathbf{A}}}\mathbf{f}'+\underline{\underline{\mathbf{B}}}\mathbf{f} \end{bmatrix} = \begin{bmatrix} \mathbf{f}' \\ -\underline{\underline{\mathbf{C}}}\mathbf{f} \end{bmatrix} = \begin{bmatrix} \mathbf{f}' \\ \underline{\underline{\mathbf{A}}}\mathbf{f}''+\underline{\underline{\mathbf{B}}}\mathbf{f}' \end{bmatrix} \equiv \underline{\underline{\mathbf{\Lambda}}}\underline{\mathbf{\Phi}}(z) = \underline{\mathbf{\Phi}}'(z) \tag{18}$$

where $\underline{\mathbf{\Phi}}$ is a vector that is continuous (and which has the envelope functions as its first half). Equation (18) has solutions

$$\underline{\mathbf{\Phi}}(z) = \exp\left(\underline{\underline{\mathbf{\Lambda}}}z\right)\underline{\mathbf{\Phi}}(0) \tag{19}$$

As $\underline{\mathbf{\Phi}}$ is continuous across boundaries one can easily propagate the transfer matrix $\underline{\underline{\mathbf{T}}} = \exp(\underline{\underline{\mathbf{\Lambda}}}z)$. For the example of a single well with z = 0 at the left-hand interface;

$$\underline{\mathbf{\Phi}}(z) = \exp\left(\underline{\underline{\mathbf{\Lambda}}}_w d_w\right)\exp\left(\underline{\underline{\mathbf{\Lambda}}}_b z\right)\underline{\mathbf{\Phi}}(0) \tag{20}$$

when z is in the right-hand barrier, where the subscripts indicate the material. To calculate the transfer matrix one notes that

$$\begin{aligned} \exp\left(\underline{\underline{\mathbf{\Lambda}}}z\right) &= \underline{\underline{1}} + \underline{\underline{\mathbf{\Lambda}}}z + \tfrac{1}{2}\underline{\underline{\mathbf{\Lambda}}}^2 z^2 + \ldots \\ &= \underline{\underline{\mathbf{U}}}\left(\underline{\underline{1}} + \underline{\underline{\mathbf{U}}}^{-1}\underline{\underline{\mathbf{\Lambda}}}\underline{\underline{\mathbf{U}}}z + \tfrac{1}{2}\underline{\underline{\mathbf{U}}}^{-1}\underline{\underline{\mathbf{\Lambda}}}\underline{\underline{\mathbf{U}}}\underline{\underline{\mathbf{U}}}^{-1}\underline{\underline{\mathbf{\Lambda}}}\underline{\underline{\mathbf{U}}}z^2 + \ldots\right)\underline{\underline{\mathbf{U}}}^{-1} \\ &= \underline{\underline{\mathbf{U}}}\exp\left(\underline{\underline{\mathbf{U}}}^{-1}\underline{\underline{\mathbf{\Lambda}}}\underline{\underline{\mathbf{U}}}z\right)\underline{\underline{\mathbf{U}}}^{-1} = \underline{\underline{\mathbf{U}}}\exp\left(\underline{\underline{\mathbf{\lambda}}}z\right)\underline{\underline{\mathbf{U}}}^{-1} \end{aligned} \tag{21}$$

where $\underline{\underline{\mathbf{U}}}$ may be chosen to be the matrix of eigenvectors of $\underline{\underline{\mathbf{\Lambda}}}$ so that $\underline{\underline{\mathbf{\lambda}}}$ is diagonal (the matrix of eigenvalues), in which case the exponential of $\underline{\underline{\mathbf{\lambda}}}$ is found simply by replacing each eigenvalue by $\exp(\lambda_i z)$ on the diagonal. Some care must be taken to eliminate large values of λ_I (which represent unphysically quickly varying envelopes and will cause numerical overflows when exponentiating), e.g., by rounding them to a cut-off value if their magnitude is larger than the cut-off, typically about 0.1Å^{-1}.

Proceeding as for uncoupled bands, one fixes $\mathbf{k}_{||}$ and chooses trial values of energy. For a quantum well valid solutions are obtained if $\underline{\mathbf{\Phi}}$(z), and hence also the determinant of the transfer matrix, tends to zero as $|z|$ tends to infinity. In the case of periodic boundary conditions, *i.e.* a superlattice, valid solutions are obtained if

$$\underline{\mathbf{\Phi}}(d_{tot}) = \underline{\underline{\mathbf{T}}}_{tot}\underline{\mathbf{\Phi}}(0) = \exp\left(iqd_{tot}\right)\underline{\mathbf{\Phi}}(0) \tag{22}$$

(where q is the superlattice wave–vector). We can see that solutions are obtained when one of the eigenvalues of the total transfer matrix has a magnitude of unity. Of course the superlattice and quantum well cases converge for large barrier thickness.

In lead-salt materials, because of the anisotropy, each L valley must be considered separately when calculating the level stucture of quantum wells. For example, when growing along the [111] direction, one of the valleys, the longitudinal one, has an effective mass simply given by the longitudinal mass, and the other three oblique valleys all at 70° to the growth direction, have a lighter mass.

1.3 Temperature Dependence of the Band Structure

Both E_P, F and the γ's are usually taken to be independent of temperature, which means that the temperature variation of the effective mass arises only through the temperature dependences of the energy gaps. In a bulk semiconductor, both direct and indirect energy gaps in semiconductor materials are temperature-dependent quantities, with the functional form often fitted to the empirical Varshni form [15]:

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta}, \tag{23}$$

where α and β are adjustable parameters. See Vurgaftmann and Meyer [6] for recom -mended values. In III-V materials α is positive and the bandgap decreases with increasing temperature, whereas in II-VI and IV-VI narrow gap semiconductors α is negative. For lead-salts the temperature coefficients of the bandgap are nearly equal for all lead-salt compounds. For PbTe, α = -0.51 meV/K, β=56 K. Note that α is negative, *i.e.* the bandgap increases with increasing temperature dEg/dT ~ +0.50 meV/K at 300 K [13].

1.4 Hydrostatic Pressure and Strained Layer Quantum Wells

The growth of semiconductor nanostructures usually involves incorporation of one material in the form of a film, wire or box into a matrix of another semiconductor. If the lattice constants of the two materials are different there will be a resultant strain, that will affect the bandstructure. The effects can very beneficial, as in the example of the strained layer quantum well laser. Furthermore, the application of external stress can be used as an externally controllable variable, so that effects that depend sensitively on the bandstructure in known ways can be extracted. This has provided an extremely powerful diagnostic tool, as has controlling the temperature or external fields.

The application of a hydrostatic, *i.e.* isotropic, pressure to a crystal produces changes in the lattice parameter, which in turn produce changes in the electronic bandstructure. By simple consideration of the fact that bonding-antibonding splittings increase with reduction in bond length, the conduction band (S antibonding) should rise and the valence band (P bonding) should decrease in energy with increasing pressure. This trend is demonstrated by the tendency of the bandgap to widen as materials are alloyed with smaller atoms, reducing the lattice constant, *e.g.* as In is replaced by Ga and then Al. The shift in energy of a particular state per unit strain is defined as the deformation potential, and these are typically in the range 1-10 eV. Additionally the **k.p** parameters are altered though these shifts are normally neglected, and in some cases degeneracies are removed. Effective masses therefore also change, due to their dependence on the gap and other parameters as described above.

Stress is a force (per unit area) acing on a surface, and hence requires specification of the surface normal and direction of the force, which need not be in the same direction, and therefore it is a second rank tensor field. Similarly for the response, *i.e.* the strain. The relationships in the different crystal symmetries between the stress and strain, which are given by the fourth rank stiffness tensor,

has been described e.g. by Nye [16]. We use here the Voigt notation that reduces the number of indices to two by symmetry.

Because hydrostatic pressure is usually compressive, it is normally treated as a negative scalar stress, $\sigma_{ii} = -P$. For a cubic crystal under hydrostatic pressure,

$$\varepsilon^h = \frac{1}{3}\frac{dV}{V} = -\frac{1}{3B}P = -\frac{1}{\left(C_{11} + 2C_{12}\right)}P \tag{24}$$

where C_{ij} are the components of the elastic stiffness tensor, B is the bulk modulus and dV/V is the fractional change in volume. For tensile uniaxial stress P_z along the z cubic crystal axis, the strains along the three crystal axes are

$$\begin{aligned} \varepsilon_z^u &= \frac{P_z}{E} = \frac{C_{11} + C_{12}}{\left(C_{11} + 2C_{12}\right)\left(C_{11} - C_{12}\right)}P_z \\ \varepsilon_{x,y}^u &= -\nu\varepsilon_z^u = -\frac{C_{12}}{C_{11} + C_{12}}\varepsilon_z^u \end{aligned} \tag{25}$$

where E is the Young's modulus, and ν is Poisson's ratio. For pseudomorphic strained layers (with or without external stress) the in-plane lattice constants of the film and substrate are equal, so the in-plane strains are related to the natural lattice constants a_{lc} by

$$\varepsilon_{x,y}^f = \left(\varepsilon_{x,y}^s + 1\right)\frac{a_{lc}^s}{a_{lc}^f} - 1 \tag{26}$$

where the superscripts are f for the film and s for the substrate. The substrate strain is zero without external stress, or otherwise is given by one of the previous two Equations (24) and (25). Including hydrostatic stress only the out of plane strain in the film is

$$\varepsilon_z^f = \frac{1}{C_{11}^f}\left(-P - 2C_{12}^f\varepsilon_{x,y}^f\right) \tag{27}$$

It is normally assumed that the values of the stiffness tensor, *etc*,. are pressure independent, but, for example, in InAs it has been shown [17] that although B is constant, C_{11} and C_{12} change by 5-10% per GPa.

The procedure for generating additional terms in the Hamiltonian due to the presence of the strain has been described in detail by Bir and Pikus [18]. This is usually resolved into isotropic, uniaxial and shear components. The isotropic component changes the volume, is proportional to $\varepsilon_V = dV/V = \varepsilon_z+\varepsilon_y+\varepsilon_z$, and it shifts the energy of the band i with deformation potential a_i. The conduction bands move up, and the valence bands move down, as mentioned above. The uniaxial component is associated with strains of tetrahedral symmetry, and is proportional to $\varepsilon_u = \varepsilon_z-(\varepsilon_y+\varepsilon_z)/2$. It splits the J=3/2 multiplet at **k**=0 with deformation potential b, so that the valence band originating from Z (which has light mass along z) moves up under uniaxial tension and the heavy band moves down as shown in Figure 2. The shear (of rhombohedral symmetry) has deformation potential d and is not considered here. Neglecting small strain dependent spin-orbit terms, the Hamiltonian for zinc-blende materials under hydrostatic stress P plus uniaxial stress P_z along [001], which should be added on to the **k.p** Hamiltonian (Equation 11) is

$a_c\varepsilon_V$			
	$-a_v\varepsilon_V - b\varepsilon_u$		
		$-a_v\varepsilon_V + b\varepsilon_u$	$\sqrt{2}b\varepsilon_u$
		$\sqrt{2}b\varepsilon_u$	$-a_v\varepsilon_V$
S	(3/2,±3/2)	(3/2,±1/2)	(1/2,±1/2)

(28)

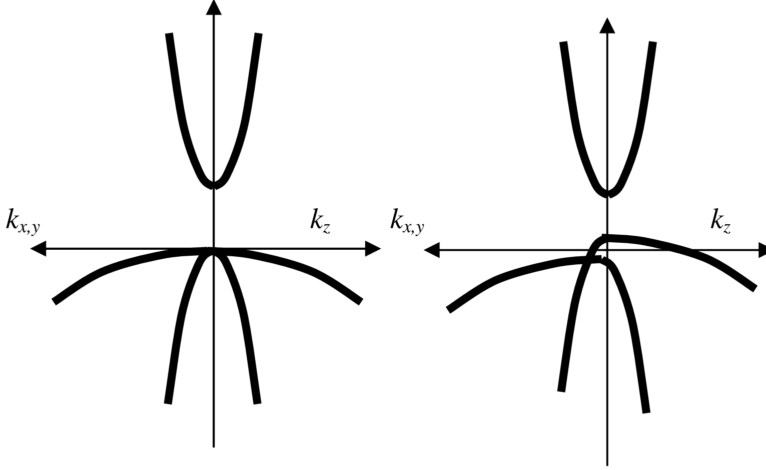


Figure 2. Schematic energy *vs* momentum in Kane bandstructure (split-off band not shown) **a**) in equilibrium, **b**) with bisotropic compression in the *x-y* plane (and/or uniaxial tension along z), corresponding to pseudomorphic growth of a large lattice constant material on a small lattice constant substrate.

It is difficult to measure the terms of Equation (28) separately, and usually experiments reveal only the change in bandgap, $dE_g/dP = (a_c+a_v)\varepsilon_V$. For most III-V materials the value of this is approximately 100 meV/GPa, *i.e.* the bandgap increases with hydrostatic compression. For the mid-IR materials InAs and InSb the hydrostatic deformation potentials have been calculated from pseudopotentials [19,20], and the valence band splittings have been determined from experiment [21]. For design of strained layer structures see [22].

For all lead salts the pressure coefficient of the bandgap is roughly the same, $dE_g/dP \sim -90$ meV/GPa [23]. The strain component of the Hamiltonian that should be added to Equationn (14) is [24]

$$H_{ii} = \left(D_d^c Tr\varepsilon + D_u^c \varepsilon_{33}\right)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \quad H_{i\neq j} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} \tag{29}$$

2 Transition Rates

Having calculated the band structure, for optoelectronic applications we are naturally interested in the transition rates between bands. The important transitions can be either radiative or non-radiative, and in mid-infrared materials the most

important non-radiative process is Auger recombination. We describe the principles of the calculation of these two processes.

2.1 Interband Optical Transition Rates

The average optical transition rate is proportional to the coupling strength K between the two states with the light field, and which may be found from Fermi's Golden Rule,

$$\begin{aligned} K_{\hbar\omega}(i,f) &= \frac{2\pi}{\hbar}\left|\left\langle \Psi_f(\mathbf{r})\right|\hat{H}\left|\Psi_i(\mathbf{r})\right\rangle\right|^2 \\ &= \frac{\pi q^2}{m_0^2\omega\varepsilon_0 V}\delta_{ki,kf}\left|\left\langle \Psi_f(\mathbf{r})\right|\mathbf{e}\bullet\mathbf{p}\left|\Psi_i(\mathbf{r})\right\rangle\right|^2 \end{aligned} \tag{30}$$

where **e** is the polarisation, q is the charge, ω is the photon frequency and ε is the dielectric constant. The wavefunctions at the band edges in bulk material are just given by the basis set S, $|{}^3/_2, {}^3/_2\rangle$, $|{}^3/_2, {}^1/_2\rangle$, $|{}^1/_2, {}^1/_2\rangle$, and away from k=0 they are linear combinations of the basis that may be found as described above. It remains only to find the matrix elements between the basis set, and for example using x-polarised light $|\langle S|e_x{\cdot}p_x|{}^3/_2,{}^3/_2\rangle|^2 = m_0^2P^2/2\hbar^2$ for the heavy hole, and $|\langle S|e_x{\cdot}p_x|{}^3/_2,{}^1/_2\rangle|^2 = m_0^2P^2/6\hbar^2$ for the light hole, where P is the Kane momentum matrix element. The transition rates are therefore in the ratio 3:1. See Voon [25].

To find the absorption strength it is also necessary to take into account the joint density of states ρ_j for the optical transitions, and in the case of non-parabolic materials this is not a trivial expression to write down.

$$\alpha_{\hbar\omega} = \frac{\eta}{c}\frac{V}{8\pi^3}K\rho_j(\hbar\omega)(f_i - f_f) \tag{31}$$

where η is the refractive index and f is the occupation function. We can write

$$\rho_j^{-1} = \rho_i^{-1} + \rho_f^{-1} \tag{32}$$

and for isotropic bands this may be reduced to

$$\rho_j^{-1} = 2\pi^2\left(\frac{\hbar^2}{2m_0}\right)^{\frac{3}{2}}\left(\frac{dE_f}{dE_{k0}} - \frac{dE_i}{dE_{k0}}\right)\frac{1}{\sqrt{E_{k0}}} \tag{33}$$

where $E_{k0} = \hbar^2k^2/2m_0$ as before. We can now insert analytic expressions for $E_i(k)$ if known (*e.g.* Equation 12) or numerically differentiate the calculated band structure.

2.2 Auger Recombination

The dominant carrier loss process in MIR interband devices is Auger recombination, and the problem is illustrated in Figure 3 and as follows. In wide band gap materials the excitation energy is large, so that in an Auger process there is a large momentum transfer. For zinc blende materials m_c* is relatively large for wide gaps, which increases it even further. Holes involved must therefore be in states far from **k**=0, *i.e.* at high energy, and the process is unlikely. Narrow gap materials, conversely,

have small electron mass, small **k**-transfer, and the Auger process involves hole states close to **k**=0 with low energy, which are much more likely to be occupied.

The Auger rate per unit volume is [26]

$$R = \frac{2\pi}{\hbar}\left(\frac{1}{(2\pi)^3}\right)\int\int\int\int |M_{if}|^2 P(1,1',2,2') \\ \times \delta(k_1 + k_2 - k_{1'} - k_{2'})\delta(E_i - E_f)d^3k_1 d^3k_2 d^3k_{1'} d^3k_{2'} \quad (34)$$

Energy and momentum conservation is included through the delta functions, and M_{if} is the matrix element for the screened Coulomb interaction potential between the two colliding carriers; 1 and 1' refer to the electron and hole recombination pair respectively; 2 refers to the second electron which is excited to 2'. Thus the un-primes are the initial, filled with electron (or empty hole) states and primes are the final, empty electron (or filled hole) states. P(1,1',2,2') denotes the net probability for the Auger-1 process. The temperature dependence arises from this probability factor which is governed by Fermi-Dirac statistics. For the n-type CHCC process

$$P(1,1',2,2') = f_{e1} f_{h1} f_{e2} (1 - f_{e2'}) \quad (35)$$

where the f_i are the Fermi-Dirac distribution functions for the electrons (f_e) and holes (f_h) which are referenced to the conduction and valence band quasi-Fermi levels, respectively. The simplest approximation states that the first three terms for the highly occupied states are proportional to n (assuming $n=p$) and that the last term for the empty state is replaced by a Boltzmann factor, which gives the temperature dependence;

$$P \propto n^3 \, exp\left(-\frac{E_{2'} - E_g}{k_B T}\right) \quad (36)$$

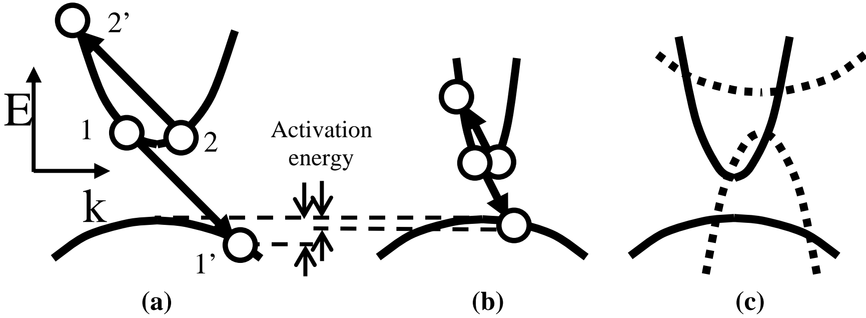


Figure 3. Schematic energy *vs* momentum in Kane band structure (light hole and split-off band not shown): **a**) wide gap and **b**) narrow gap. The CHCC Auger process is depicted for two electrons at the peak of their distribution. Conservation of energy and momentum for these electrons requires that the hole is in a state excited away from the edge. The mid–infrared material has a narrower gap, smaller conduction effective mass and a much smaller activation energy for Auger recombination. **c**) A graphical solution to the conservation laws is obtained by drawing a bandstructure inverted about any arbitrary point; choice of any pair of intersections marks four states that can take part in a collision (one intersection marks the initial states and one the final states, *e.g.* for a CHCC process one cb-cb and one cb-hh intersection are required). According to Equation (36) the threshold process is the one that minimizes E_2.

This approximation shows the importance of the state furthest from the band edge: P is largest when $E_{2'}$ is smallest. Since there exists a minimum energy for $E_{2'}$ below which the δ-functions cannot be satisfied, the Auger rate is dominated via the exponential by contributions at this "threshold energy", E_T. The threshold process does not necessarily involve electrons and holes at the band edge, but while processes which do involve such carriers can always be found, their probability is always reduced by the need to involve other states with much lower occupation. In the case of an emitter device, in which there are many electrons and holes near the band edge, the activation energy is simply the excess kinetic energy required to perform the threshold process. For zinc-blende materials, where $m^*_{hh}/m^*_c \sim (\gamma_1-2\gamma_2)2E_P/3E_g >>1$, we may use the approximation that the valence band is approximately flat when looking for the threshold process. In this case the electrons are right at the band edge, *i.e.* $E_{2'} \sim 2E_g$ (measured from the top of the valence band) and hence, for parabolic bands, the total initial excess energy, equal to the hole energy is

$$E_T = -E_{1'} \approx \frac{m_c^*}{m_v^*} E_g \tag{37}$$

N.B. for hole processes such as CHHH a similar expression is derived with m^*_v and m^*_c interchanged. By inverting Equation (12) for the conduction band momentum, one can show that non-parabolicity in very narrow gap materials such as InSb approximately doubles the threshold which would otherwise be expected for a parabolic band with the same band edge mass, because it pushes up the momentum transfer required, and hence also pushes up the hole energy at threshold. Inserting the results for the band edge masses from Equation (13) for large Δ and small valence band anisotropy, Equation (31) becomes

$$E_T \approx (\gamma_1 - 2\gamma_2)\frac{3E_g}{2E_P} E_g \tag{38}$$

i.e. proportional to the square of the bandgap (since the other parameters are similar for all III-Vs).

It is a commonly used approximation to include only the threshold process in integration of the equation for R, which gives the well-known approximation for the carrier density dependence:

$$R = \text{const}\,.\, n^3 \left|M\right|^2 \exp\left[-\frac{E_T}{kT}\right] = Cn^3 \tag{39}$$

where C is called the Auger coefficient and the inverse Auger lifetime $1/\tau_{\text{Auger}} = R/n = Cn^2$. The transition rate can therefore be influenced by manipulation of:

a) The matrix element M
b) The threshold energy
c) The carrier density
d) The temperature

Of course it is the rise of the Auger rate with temperature which inhibits room temperature operation, and it is this which needs to be offset by reductions in other ways, such as using Type II indirect transitions to reduce M.

The δ-functions of Equation (34) provide constraints on the integral of the Auger rate and determine the size of E_T. However if the momentum conservation constraint is relaxed due to occupied levels existing with a broad range of momenta, such as at some band extremum with a large effective mass, then the threshold will be reduced and the Auger rate will be strongly enhanced. Or to put it the other way round, the Auger rate can be reduced by engineering the band structures such that the heavy hole effective mass is reduced (*e.g.* by stress - see below) to approach that of the conduction band. Furthermore, since we are concerned with nonradiative recombinations across the bandgap with energy E_g, any energy levels (again band extrema in particular) separated from the conduction or valence band edges by E_g are particularly susceptible to the excitation part of the Auger process. The Auger rate may therefore be reduced by arranging that there are no band extrema at energies E_g from the conduction and valence band edges. This is important in InAs and its alloys, since the spin-orbit split-off valence bandgap is in close proximity to the conduction-valence bandgap, which enhances CHSH processes. It is therefore important to choose alloy combinations which remove this resonance. Hydrostatic pressure has found real success as a diagnostic tool for Auger recombination due to the dependence of the threshold on E_g and m^*, both of which can be tuned with pressure.

For more sophisticated calculations of Auger rates, both the large number of dimensions of the integrals and satisfaction of the δ functions when non-parabolicity is important make it very costly in terms of computation time. Several approaches have been employed to reduce this, and these include Monte Carlo methods to reduce the number of points in $\mathbf{k}_{1,1',2,2'}$ space included. The so-called Flat Valence Band Approximation [27] is one way to simplify satisfaction of the δ-functions (illustrated in Figure 3c). Because the valence band mass is much heavier than the conduction band (or light hole for the CHLH process), one can remove the energy conservation easily (as done in the discussion of Equations 37 and 38). It is then possible to find a very useful analytic simplification of the Auger integral for the CHCC and CHLH Auger rates, $R(n,T)$. The results of this approximation agree with more detailed calculations over a large range of alloys, temperatures and equilibrium and non-equilibrium excitations.

By far the most widespread implementation of Auger suppression has been by reduction of the carrier density required for the application. It has been used at telecommunications wavelengths in strained layer quantum well lasers, which reduces the hole mass, m^*_h, with compressively strained layer heterostructures [28]. This has the effect of more closely matching the conduction and valence band masses and, via the Bernard-Duraffourg condition, reducing the excited carrier density, n, required for emission, and this increases τ_{Auger}. The other effect of the lighter valence band mass is to concentrate the holes closer to the zone centre, which counteracts the detrimental effect of a small electron mass described above. In effect E_T is increased, reducing C via Equation (37).

2.3 Auger Rates in Bulk Zinc Blende Materials

The main thrust of band structure engineering techniques for Auger suppression has been in III-V systems, and it is important to gain more quantitative information on

Auger rates in those systems than is given in the above approximations. In fact an important device which has been demonstrated to operate between 3-5 μm (though only up to ~170 K) is based simply on homostructures of bulk InSb or dilute $In_{1-x}Ga_xSb$ [29].

Of course it is also important to test the models against experiment. The alloy $In_{1-x}Ga_xSb$ has been incorporated into homojunction lasers and identified as potentially an important component in band structure engineering of quantum structures based on the narrow gap III-V material InSb. A time-resolved pump-probe measurement, which gives the most direct way of lifetime determination, has been made of carrier recombination in $In_{1-x}Ga_xSb$, for a range of alloy compositions $x = 0.0$ to 0.2 (corresponding to a room temperature energy gap variation from 0.175 to 0.215 eV), excited carrier densities (5×10^{16} to 3×10^{17}cm^{-3}), and temperatures (30–300 K)[30]. The room temperature Auger coefficient, C, dropped from the value 1.22×10^{-26} cm^{6}s^{-1} at $x = 0$ (*i.e.* InSb) to 1.03×10^{-26} cm^{6}s^{-1} at $x = 0.2$. The Flat Valence Band model gives 1.17×10^{-26} cm^{6}s^{-1} and 0.98×0^{-26} cm^{6}s^{-1} respectively, and agrees very well. Assuming that only E_g changes with composition, Equation (37) predicts a 12% drop, which also agrees well considering the simplicity of the approximation.

The Flat Valence Band model has made similarly accurate predictions for bulk $Hg_xCd_{1-x}Te$ [31], and since $Hg_xCd_{1-x}Te$ has wide bandgap tunability the results of the model make a convenient benchmark for comparisons of Auger rates.

2.4 Auger Suppression in Bulk Materials: Pb-salts

As an initial step towards quantum well devices it is interesting to consider the possibilities for natural Auger suppression by choice of bulk material. As mentioned above, zinc blende materials have an effective mass ratio that diverges from unity as the bandgap is reduced, and this has the effect of reducing E_T. The lead salts, conveniently, have mirror bands, thus maximising the Auger threshold; $E_T = E_g/2$ in the parabolic approximation, and these materials are widely used in mid-IR optoelectronic emitter and detector devices.

Pump-probe transmission experiments have been performed on PbSe in the temperature range 30 to 300 K, using a ps free electron laser [32]. For temperatures below 200K and carrier densities above the threshold for stimulated emission, stimulated recombination represents the most efficient recombination mechanism with relatively fast kinetics in the 50-100 ps regime, in good agreement with earlier reports of amplified spontaneous emission. Above this temperature Auger recombination dominates. The Auger coefficient is approximately constant (with a value of about 2×10^{-28}cm^{6}s^{-1}) between 300 and 70 K. In this range E_g drops with reducing T so E_T also drops, balancing the effect of the temperature change in the exponential term of Equation (37). Below 70 K, the change in T wins and C drops to a value of about 5×10^{-29}cm^{6}s^{-1} at 30 K. The fact that Auger recombination is so much lower in IV-VI semiconductors than Kane bandstructure materials in the wavelength range of 3-5μm is one of the main reasons why lead-salts are able to provide quantum structures achieving room temperature laser operation [33], at least for optical pumping.

2.5 Auger Suppression in Bulk Materials: Dilute III-V Nitrides

It is also possible to find alloys of III-V semiconductors with more symmetric effective masses than equivalent bandgap HgCdTe. It has been recently observed that dilute alloys of III-V semiconductors with nitrogen exhibit a very strong bandgap bowing and this has opened up the possibility of long wavelength emission. We observed strong bandgap bowing to long wavelengths in InSbN [34], with associated strong increase in non-parabolicity. Negative bandgaps in InSbN have been shown even at very low N concentrations [35]. Furthermore the conduction band non-parabolicity is increased due to the mixing of the localised N-level [36]. The resulting increase of the conduction band mass towards the valence band mass increases E_T via Equation (37) and suppresses Auger recombination.

We have performed time-resolved pump-probe transmission measurements on a sample of InN_xSb_{1-x} with a bandgap near 11μm [37]. The time-resolved bleaching experiment at 10μm mimics the threshold condition for lasing at that wavelength. Pump-probe results are shown in Figure 4 for a pump wavelength of 10μm, plotted on a log-linear scale so that 1/τ is the slope. There is a rapid decay (steep slope) corresponding to an initial lifetime per excited carrier of 300 ps. The recombination is dominated by Auger processes and marked by a lifetime which slows with time. Eventually defect related processes take over and the lifetime is constant.

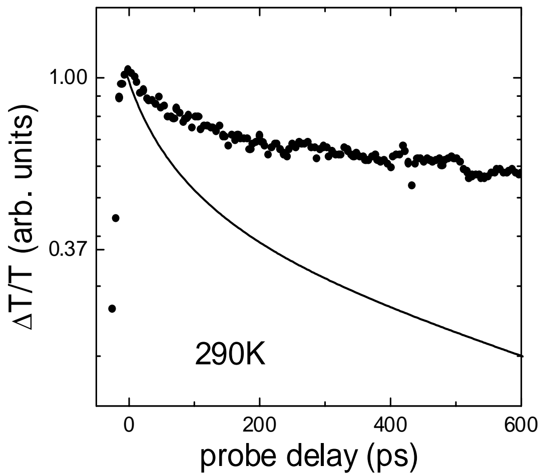


Figure 4. Differential transmission change of an interband probe pulse as a function of time delay after a degenerate interband pump pulse. *Symbols*: experimental results for an InSbN sample with $E_g = 0.113$ eV ≡11 μm at 290 K, with $\hbar\omega_{pump} = 0.124$ eV ≡ 10 μm. *Curve*: theoretical transmission change as a function of time for an initial density $n = 4.6\times10^{17}$ cm^{-3} in $Hg_{1-y}Cd_yTe$ with $y = 0.170$ with the same bandgap and same pump energy. Reprinted with permission from [37] (Copyright 2001, American Institute of Physics).

In order to examine the benefit of the N-alloy over Kane band-structures we have compared them with equivalent bandgap material such as $Hg_{1-y}Cd_yTe$. A bandgap of 0.113 eV at 300K is achieved at *y*= 0.170, and we calculated the actual Auger transition rates using the flat valence band approximation for this alloy at 300K. The Auger decay rate for $Hg_{1-y}Cd_yTe$ with the same bandgap (11μm) as our N-alloy InSb sample, and the same pump wavelength as used in the experiment (10μm) is shown, demonstrating a much shorter initial lifetime of 97 ps. The decay curve is also shown in Figure 4. This result is critically important for emitter devices, where it is the Auger lifetime at a given quasi-Fermi level separation which determines the threshold current. We have shown that the Auger lifetime at this excitation is substantially longer than would be obtained with a Kane band-structure of similar gap, and indeed is only marginally shorter than that for InSb itself (bandgap ~ 7 μm). This experimental conclusion does not require any knowledge of the InN_xSb_{1-x} band-structure (though as we have explained it is a consequence of it), its Auger coefficient or the carrier density. There is great scope for low-dimensional structures based on this material to suppress the Auger rate even further.

2.6 Auger Suppression in Quantum Wells

Detailed calculations of the Auger rates in quantum wells are not so simple, and many authors have shown that simple design rules are difficult to give, especially in strained and Type II systems where the bandstructure is so complicated. However, from a theoretical point of view, it turns out to be very important to include transitions which involve higher subbands and/or unbound states. It is quite easy to quench substantially the main $C_1H_1C_1C_1$ Auger process (where the subscript indicates the subband index), as described below, but may be considerably harder to quench others. In particular $C_1H_1H_1H_n$ (*i.e.* hole processes involving higher subbands) will be present unless the confinement energies and/or strain energies are higher than the bandgap. Here we use the notation H to refer to any valence band state, since Auger recombination normally involves states away from **k**=0 where the mixing makes attempts to classify states as light or heavy irrelevant. $C_1H_1C_1C_n$ will occur to a lesser extent due to the greater complexity of the valence band. $C_1H_1C_1C_{unbound}$ and $C_1H_1H_{unbound}H_1$ (*i.e.* electron and hole processes involving states outside the wells) will also be present unless the band offsets are greater than the bandgap [38]. Although these higher order processes are only weakly allowed due to the change in z-momentum, they may easily dominate over radiative recombination if the carrier density is high. The former type, involving higher bound states, is resonant when there is a subband edge at an energy E_g away from the band edge. The latter type involving unbound states is resonant when the band offset is equal to the bandgap, and decays slowly for smaller offsets due to the mismatch of the quickly varying unbound state with the slowly varying bound state. A design rule which follows immediately is that ideally systems should be found with large band offsets and large confinement energies relative to the bandgap, so that there is only one subband from each band involved in the recombination processes. This is, of course, quite hard to do in practice.

As mentioned above, Auger recombination may be suppressed in quantum structures and this principle has been very successfully applied to NIR devices using

compressive and tensile strain to reduce the threshold carrier concentration. Compressive strain also acts to increase E_T by reducing the valence band edge mass, and this principle has also proved useful in MIR devices.

Considering an unstrained quantum well heterostructure, the carriers are quantised along the growth direction, but remain free to move in the well plane, with wave–vector $k_{\parallel}$. The confinement energies, however, vary inversely with the mass along the growth direction, corresponding to wave–vector $k_{\perp}$ and this determines the labelling of the states. Since the heavy in-plane states arise from the light perpendicular states and vice versa, in an unstrained well the highest valence subband, HH1, has light in-plane dispersion. For a layer under biaxial compression (e.g. AlSb on GaSb), or equivalently uniaxial tension, the HH subband edge shifts upwards with respect to the LH subband edge, which can increase further the energy splitting EHH_1 - ELH_1. This has considerable benefits for the inhibition of Auger recombination as it helps symmetrise the band edge masses, thus increasing E_T.

In a superlattice it may also be expected that minigaps in the conduction band may have the effect of inhibition of the Auger rate, just as does subband splitting in quantum wells. A simple design rule is therefore that there should be no minibands at an energy Eg away from the band edge which would give a resonant Auger path. In quantum structures such as InAsSb/InAs superlattices, reduction of the density of states Eg away from the band edge gives rise to reduction in the recombination time [39].

Probably the most versatile system is the III-V broken gap system InAs/GaSb/AlSb. These binaries are nearly lattice matched, and there exists a large range of possibilities for strained layer quantum wells. Most interestingly, the conduction band of InAs lies below the valence band of GaSb and without quantum confinement electrons and holes can coexist even at zero temperature at the interface. Quantum wells of this system, where the confinement energy raises the degeneracy, may therefore have any bandgap from zero up to the near IR. Furthermore, the band-edge envelope functions are spatially indirect, and this enables modification of the Auger rate via the matrix element. Because the Auger matrix element is quite different from the optical matrix element, cleverly designed coupled wells and superlattices with extended states may have quite large optical matrix elements in spite of a suppressed Auger matrix element. In any case, for laser devices it is even desirable to suppress the spontaneous emission rate.

Many authors have considered in detail the Auger scattering rates in these rather complex SLS structures where the two-dimensional integral and sum over the subband indices becomes a three-dimensional integral to incorporate the full dispersion in the z-direction. The formula for the rate looks similar to the bulk case, except that the wavefunctions are now those of a superlattice bandstructure. Large strain induced splittings of the valence band are possible and give rise to an inhibition of the p-type process, CHLH. The splitting may be larger than the bandgap which means at least one of the carriers involved must come from a region of k-space with low occupation. In practice for experimental reasons it has not been possible to grow the thick SLSs as originally envisaged and early results for long wavelength lasers on what have been referred to as superlattice multiple quantum well structures, whilst very encouraging at low temperatures, have not been as successful as hoped at room temperature. Similar calculations have been made of

the reduction in Auger rate expected for the InAsSb-based type II SLS systems [40]. These authors have also shown that the Coulomb interaction responsible for Auger recombination may also have a long enough range to cause inter-well scattering in multi-quantum well structures (*i.e.* with charges 1 and 2 in different wells). This particular process may be suppressed by simply ensuring the wells are separated by at least the screening length.

3 High Pressure

The number of studies that have been made of mid-infrared optoelectronic devices under pressure is very small, which is in some sense surprising considering the importance of the information that has been obtained for near-IR materials and devices. Possibly the two main reasons are the requirement for cooling of mid-IR devices and the need for a pressure medium that transmits at these wavelengths. The most convenient pressure cells that allow electrical and optical access to the device under test use a liquid such as an oil or solvent to transmit the pressure. Cooling the cell results in freezing of the medium which has a significant problem for the magnitude and uniformity of the pressure. Furthermore, these molecules have significant numbers of vibrational and rotational absorption lines in this spectral region, and these cause trouble especially because they shift with pressure. The best medium from these points of view is helium, because it remains fluid at lower temperatures and higher pressures, and does not have spectral features in the mid-IR. On the other hand, it is much more difficult to seal, and being gaseous, it is highly compressible meaning a dangerous amount of energy is stored in the system.

3.1 Near-infrared Applications

Because of their commercial importance in optical fibre communication systems, there have been detailed studies of the radiative and non-radiative recombination processes in 1.3μm and 1.5μm wavelength lasers based on InP. The trends that have been established provide useful information for the prediction of the characteristics of Type-I Mid-IR quantum-well lasers based on the antimonides. It is also of interest to compare and contrast the two materials systems which have similar band structures except for the relative sizes of the band gap, E_g, to spin-orbit splitting, Δ, energies.

3.1.1 Radiative Recombination

In order to distinguish between the radiative and non-radiative processes that contribute to the threshold current, it is convenient to use the experimental arrangement illustrated in Figure 5. In addition to observing the stimulated emission from the end facet, the spontaneous emission from a window *etc.*hed in the substrate electrode is monitored. Since this light travels perpendicular to the plane of the well, it suffers negligible re-absorption or gain and the integrated spontaneous emission at threshold is directly proportional to J_{rad}, the radiative threshold current density. We first consider the temperature dependence of the radiative recombination current at

threshold. Details of the measurements and devices can be found in the paper of Phillips *et al.*[41]. Figure 6(a) shows [42] the observed variation of the integrated spontaneous emission, L, as a function of current at 78 K and 94 K. The characteristics are linear up to the threshold current, I_{th}, above which point L becomes constant since the carrier concentration becomes pinned by the stimulated emission process.

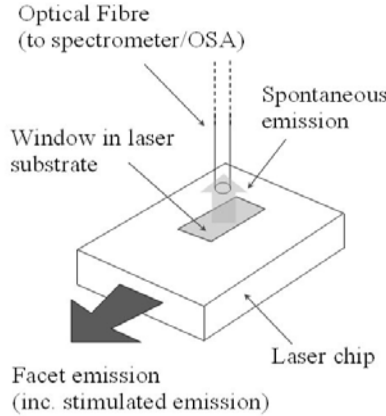


Figure 5. Schematic of the set-up used to measure the spontaneous emission from semiconductor lasers. From these measurements, the radiative and non-radiative currents may be quantified.

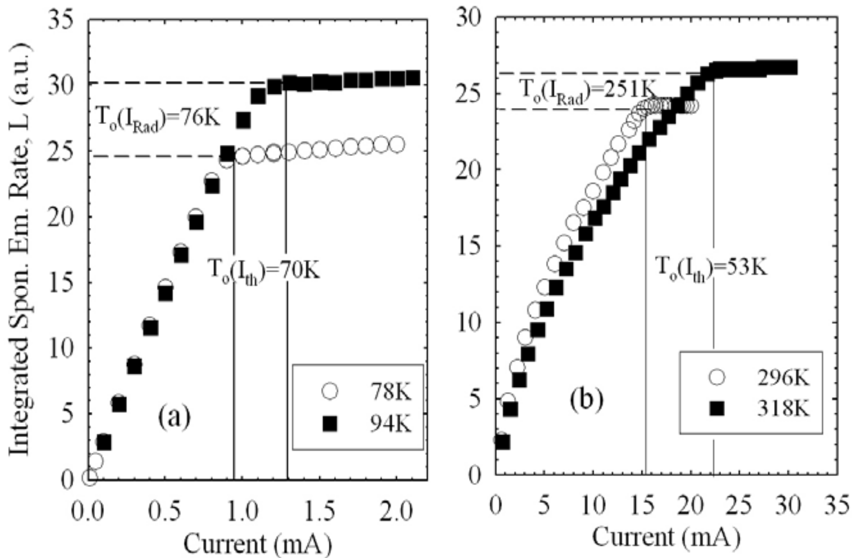


Figure 6. Integrated Spontaneous emission *vs* current for a 1.5 μm laser at (**a**) 78 K and 94 K and (**b**) 296 K and 318 K. Reprinted with permission from [42] (Copyright 1998 IEEE).

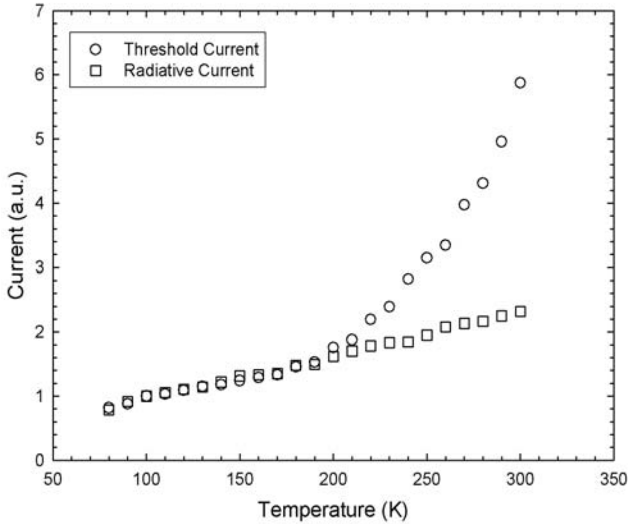


Figure 7. Threshold current (*circles*) and radiative current (*squares*) *vs* temperature for a 1.3μm InGaAs(P) laser. The radiative current is determined from the pinning level of the spontaneous emission which is normalised to threshold at low temperature.

The T_0 for both I_{th} and I_{rad} are shown in the diagram. As can be seen, they are both close to T_0=T, where T is the absolute temperature as one would expect for an ideal quantum well dominated by radiative recombination [43]. Extending this, Figure 7 shows the measured variation of I_{th} from 80 K up to 300 K together with the variation of I_{rad} determined from the pinning level of the spontaneous emission. Up to ~180K, I_{th} and I_{rad} have the same linear ideal temperature dependence hence at low temperature we may normalise I_{rad} to I_{th}. Clearly, at higher temperatures, $I_{th} \neq I_{rad}$ due to non-radiative recombination as will be discussed later.

Second, we consider the pressure dependence of the radiative recombination. Figure 8 shows the measured photon energy of the lasing wavelength in 1.3μm and 1.5μm low temperature and near room temperature. It varies as 8-9 meV per kbar which is close to the measured pressure dependence of the band gap [44].

Figure 9 shows the measured variation of the threshold current of a 1.3 μm InGaAsP laser at 125 K where we know that it is dominated by radiative recombination (see Figure 7). Also shown by the dashed line is the expected [44] variation of $I_{rad} \propto E_g^2$. The difference in the two can be explained by the fact that the optical confinement factor, Γ, increases with decreasing wavelength. Taking this into account we obtain the solid curve in Figure 9 which is in good agreement with the experiment [45].

In conclusion, measurements on near IR InP based quantum well lasers show that the radiative component of the threshold current increases linearly with temperature T and increases as E_g^2 in lasers with the same optical confinement factor. Thus we would expect that longer wavelength Sb based lasers would exhibit a lower threshold current density than their near IR InP based counterparts if no non-radiative recombination processes are occurring.

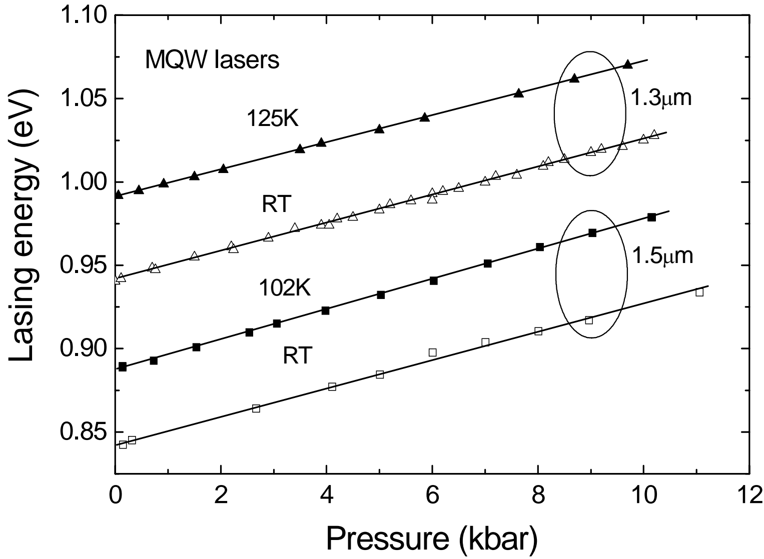


Figure 8. Pressure dependence of the lasing energy for a 1.3 μm (*triangles*) and 1.5 μm (*squares*) InGaAs(P) laser at low temperature and at room temperature. Reprinted with permission from [45].

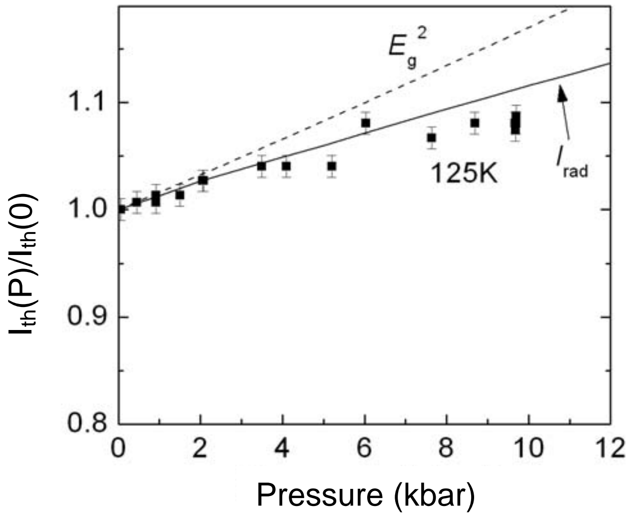


Figure 9. Pressure dependence of the threshold current for a 1.3 μm InGaAsP laser at low temperature. The *dashed line* is the ideal expected variation of the radiative current whilst the *solid line* is the calculated radiative current taking account the pressure-induced increase in optical confinement factor. Reprinted with permission from [45].

3.1.2 Non-radiative Recombination

Close to room temperature there is a sub-linear increase in spontaneous emission rate L with current, as shown in Figure 6(b), indicating the presence of a nonradiative current path. This results in a decrease in T_0 for the threshold current, T_0 (I_{th})=53K although T_0 (I_{rad}) remains similar to T. The presence of a strongly temperature sensitive nonradiative current path at higher temperature is also clear from Figure 4.3 where it can be seen that above a certain break-point temperature [42], in this case ~180 K, I_{th} starts to increase much more rapidly than I_{rad}.

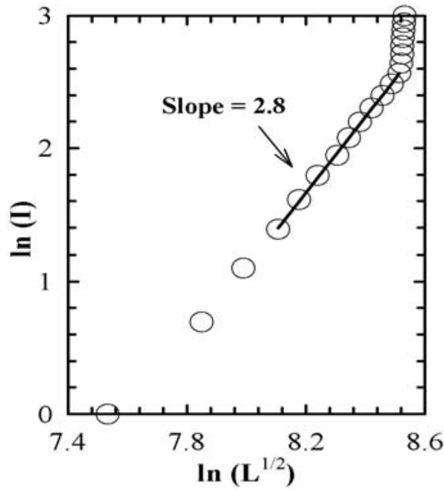


Figure 10. Plot of ln(I) *vs* ln($L^{1/2}$) for a 1.5 μm InGaAs laser at room temperature. The slope of the plot, Z, is the power dependence of the current on carrier density. At this temperature, Z=2.8≈3, hence Auger recombination ($\propto n^3$) dominates.

In order to identify the nature of the non-radiative recombination at room temperature we plot ln(I) against ln($L^{1/2}$) in Figure 10. Over a limited current range we may write that $I \propto n^Z$, where Z will have a value of 2 if the current is dominated by radiative recombination ($\propto n^2$) and 3 if Auger recombination ($\propto n^3$) plays a significant role. However, the total integrated spontaneous emission rate $L \propto Bn^2$ and therefore $n \propto L^{1/2}$, from which it follows that $I \propto (L^{1/2})^Z$. thus plotting a graph of ln(I) against ln($L^{1/2}$) yields directly the value of Z from the gradient. The value of Z=2.8 obtained in Figure 10 indicates that at the threshold current is dominated by Auger recombination. Figure 11 shows how Z increases from about 2 at low temperatures up to about 3 at higher temperatures as Auger recombination becomes increasingly important.

When Auger recombination is dominant then we would expect that in an ideal quantum well laser we may write [43] that $T_0=T/[3+(E_a/k_BT)]$, where E_a is the activation energy for the Auger process and k_B is the Boltzmann constant. Thus the presence of Auger recombination explains well the low value of T_0 (I_{th}) shown in Figure 6(b).

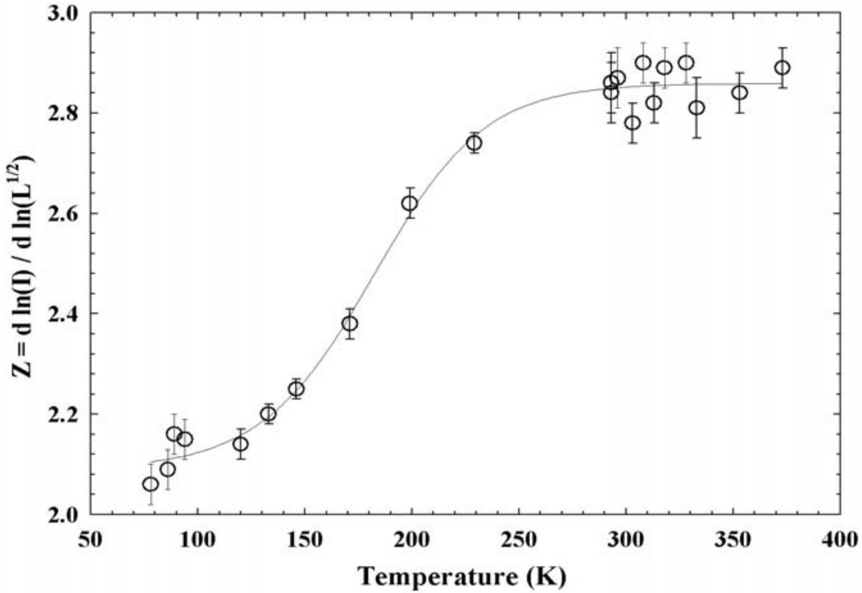


Figure 11. Z(T) for a 1.5 μm InGaAs laser. At low temperature, Z≈2 due to dominant radiative recombination ($\propto n^2$). However as the temperature increases, Auger recombination becomes more important such that by room temperature, Z≈3 due to dominant Auger recombination ($\propto n^3$). Reprinted with permission from [42] (Copyright 1998 IEEE).

As discussed earlier in this chapter, the two primary Auger recombination processes are; the CHCC process which leads to the generation of a hot electron in the conduction band and the CHSH process which gives rise to a hot hole in the spin-split-off band. These hot carriers have energies greater than the heterojunction barrier heights in InP based lasers and can give rise to further current losses if the hot electron diffuses out of the waveguide region into the p-doped cladding layer or the hot-hole diffuses into the n-type cladding layer. Sweeney *et al.*[46] performed an interesting experiment in which they studied lasers with quantum wells offset from the centre of the waveguide either towards the n-side or p-side. From this they concluded that in 1.5 μm InP based lasers the CHSH process is stronger than the CHCC process. This result is of considerable importance for mid-IR antimonide based lasers where Δ is larger than E_g and the CHSH process can no longer occur. Finally, having established the importance of Auger recombination in determining the threshold current in InP based near-IR lasers, it is interesting to study the variation of the total I_{th} with band gap. This is shown in Figure 12 where measurements of the pressure dependence of I_{th} for several near-IR lasers have been normalised and plotted *vs* lasing energy. As can be seen, I_{th} decreases swiftly with increasing band gap. Since it has been established that the radiative component of I_{th} increases with increasing band gap, the results of Figure 12 can only be explained if I_{th} is dominated by a non-radiative process which decreases quickly with increasing band gap. Detailed theoretical analysis shows that the change is consistent with the direct CHSH Auger processes [44,46].

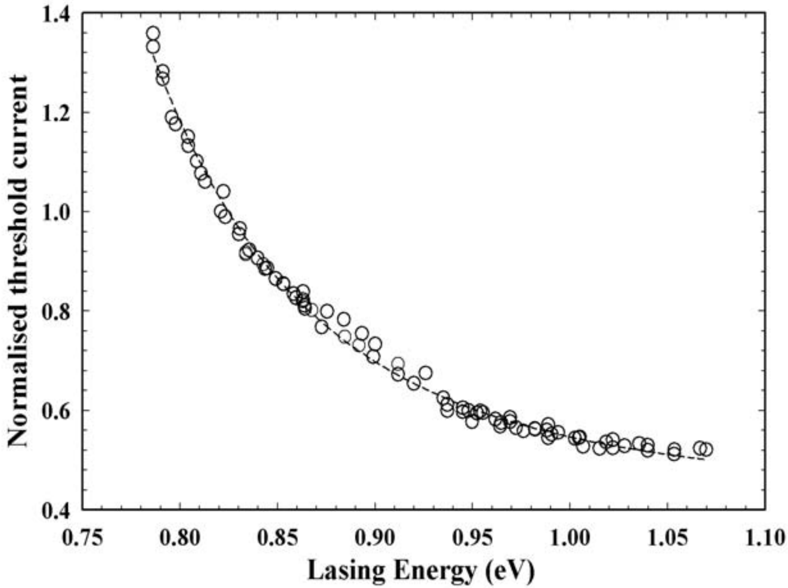


Figure 12. Normalised variation of the measured threshold current with lasing energy determined by performing high pressure measurements on several lasers over the nominal wavelength range 1.3-1.6 μm. With increasing lasing energy (band gap), the threshold current decreases strongly due to the decreased contribution of Auger recombination. The *line* is a guide to the eye.

3.2 Determination of Mid-IR Material Pressure Coefficients

The values of the pressure coefficients of many narrow bandgap materials have been determined. Because the g-factor (which gives the magnitude of the spin splitting in magnetic field) depends on the band structure parameters as does the effective mass, spin resonance may be used to determine E_p *etc.* E_p has been shown to increase in $In_{0.53}Ga_{0.47}As$ by about 3% per GPa [47]. In quantum structures it is also important to know the way that band offsets vary with pressure, and this has been investigated in narrow gap materials InAs/InGaSb by measuring the bandgap (from photo-luminescence) in a clamp cell, and comparing with **k**.**p** calculations (giving ~10% per GPa increase)[48]. Other corrections have been observed in the elastic constants with pressure, and have been shown to be important in calculating the pressure coefficients of strained layer quantum wells in GaInAs [49].

One of the most striking changes that occurs in III-V materials with pressure apart from the fundamental bandgap, is the movement of the valleys at other points in the Brillouin Zone relative to each other. In particular the X-valley energy decreases at ~10meV/GPa (c.f. the Γ-minimum which rises at ~100meV/GPa). This has been observed in InAs quantum dots in the pressure dependence of the luminescence in a diamond anvil cell [50].

As described above dilute concentrations of nitrogen have been shown to be very useful for reducing the bandgap of near-infrared and mid-infrared III-V materials. The bandstructure is sensitive to the position of the nitrogen level, and of course this also moves with pressure, and this has been shown to move at 25 meV/GPa for near-IR InGaAsN/GaAs quantum wells [51].

It is also important to know the phase transition pressures when interpreting high pressure data, and one way to examine this is using Raman spectroscopy to observe the change in vibrational frequencies, as for example in InAs/InAsSb [52].

The band structure in GaSb/InAs presents special problems because of the overlap of the InAs conduction band and GaSb valence band. The band structure, and hence optical overlaps and non-radiative transition rates, are quite sensitive to the structure and growth conditions, and in particular to the type of bond at the interface which can be InSb-like or GaAs-like. Studies of the magneto-transport at hydrostatic pressures up to 1.4 GPa showed that the band overlap for InSb-like interfaces have an overlap 30 meV larger than GaAs-like samples, and thus also greater intrinsic charge [53]. High pressure up to 1.1 GPa has been used on resonant tunnelling structures of InAs/GaSb/InAs [54]. The current characteristic of such devices shows negative differential resistance (NDR), arising from resonant conduction through quantum confined hole states in the well. The NDR is lost as well width is reduced below 6 nm, and it had previously been proposed that this is because the dominant conduction is through the LH1, whose overlap at this point with the barrier conduction band is removed by increasing quantum confinement. However the hydrostatic pressure results showed that NDR remains in wider well devices even when the LH1 level is pushed out of overlap by pressure. The 6nm threshold was found to be due to the change in electron wavefunction penetration due to its mixing with the light-hole state in the well.

3.3 Interband Mid–IR Devices

The first report of a laser being tuned by hydrostatic pressure was in the mid-infrared using a PbSe device [55]. The devices were mounted in a cell with He pressure medium at 77K, with poly-Si windows. The devices were tuned from 8 μm at ambient pressure to 22 μm at 1.4 GPa, and the output allowed spectroscopy of the phonon bands of the window. Since those first measurements, there have been virtually no other reports of mid-infrared optoelectronic devices under pressure. We present here some of our recent studies on mid-IR LEDs and lasers.

3.3.1 InAs LEDs

We have measured the characteristics of InAs-based LED devices under He gas pressures up to 1 GPa [56]. In an InAs/InAsSbP device where the electrons were confined in the InAs and the holes in the alloy (*i.e.* Type-II), a very significant increase in radiative efficiency was seen, of a factor of nearly four, indicating that it accounted for less than a quarter of the current initially, and that Auger recombination was very strong. Calculations predicted 14% at ambient conditions increasing by a factor of nearly three, in good agreement. For an InAs/InAsSb Type-I device operating at 3.8 μm under similar conditions the increase was less than a

factor of two. This indicates although the Type-II nature of the interface should suppress Auger recombination by reduction of the Auger coefficient C (*via* the matrix element M), and the wider bandgap (3.3 µm) should reduce E_T, the carrier concentration which enters the current as n^3, has risen. This is likely to be due to the reduction in active volume by confinement at the interface, so for a given current the density is higher. Furthermore, although in InAs the spin-orbit split-off bandgap is close to E_g and brought into resonance by the pressure, no resonance was seen in the output characteristic. This indicates that the CHCC rather than CHSH was the dominant process.

3.3.2 Comparison of Near-IR and 2.37 µm InGaAsSb Lasers

Excellent quantum well lasers operating at 2.37 µm have been produced by the MBE growth of InGaAsSb [57] and these probably afford the best examples for comparison with the devices described above. From our studies of the recombination processes in near-IR InP based devices we would expect that the threshold current density in mid-IR GaSb based devices with similar structures should be much lower. Most importantly the room temperature direct band gap E_g of 0.52 eV is less than Δ=0.72 eV and thus the CHSH Auger process, the major component of the threshold current in the 1.5 µm lasers cannot occur. Second, the radiative component of I_{th} will have decreased as E_g^2, to about 40% of that at 1.5 µm. In Figure 13 we plot J_{th} as a function of temperature for a 1.5 µm laser with four wells (the optimum number for a low J_{th}) [58] and a 2.37 µm laser with three wells. As can be seen, the threshold current density is indeed lower for the longer wavelength device as predicted [59]. However, although J_{th} is lower, we see that

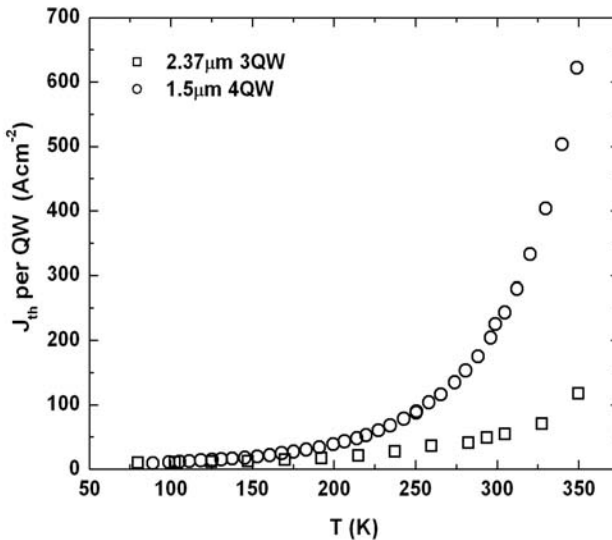


Figure 13. Threshold current density (J_{th}) per quantum well *vs* temperature for a 2.37 µm InGaAsSb/AlGaAsSb laser (*squares*) compared with a standard 1.5 µm InGaAs/InP laser (*circles*). The threshold current density per QW is lower for the 2.37 µm device as expected.

it remains temperature sensitive. Indeed the value of T_0 is approximately 100 K at room temperature. That is T_0 is approximately equal to T/3 which is again indicative that the threshold current is dominated by Auger recombination.

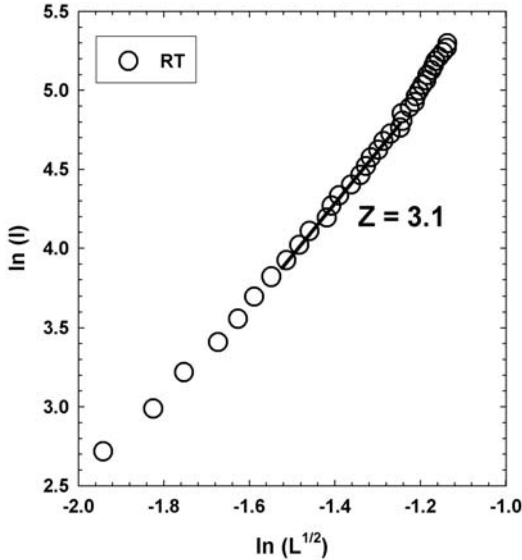


Figure 14. Plot of ln(I) *vs* ln($L^{1/2}$) for a 2.37 μm InGaAsSb laser at room temperature. In a similar way to the 1.5 μm device, we measure Z=3.1≈3 at room temperature suggesting that Auger recombination ($\propto n^3$) dominates at this temperature.

To test whether this hypothesis is correct, measurements were again made [59] of the spontaneous emission collected through a window milled in the substrate electrode and the observed variation of ln(I) against ln($L^{1/2}$) is shown in Figure 14. At room temperature, the value of Z=3.1 supports the view that Auger recombination is present. In Figure 15 we plot the temperature dependence of Z for the 2.37 μm device. As for the 1.5 μm device, we find that whilst Z≈2 at low temperature, indicating that radiative recombination dominates, as the temperature is increased, Auger recombination becomes steadily more significant.

Finally, if Auger recombination is dominant, we would expect that the threshold current density would decrease with increasing pressure. Such measurements were carried out by Adamiec *et al.*[57]. Figure 16 shows the observed rate of variation in the lasing photon energy with pressure. The linear pressure coefficient of dE_g/dP ≈12 meV/kbar is similar to that for the 1.5 μm devices.

The resulting variation of threshold current is extremely interesting and is shown in Figure 17. At first J_{th} does indeed decrease with pressure. However, the rate of change soon decreases and when the photon energy exceeds 0.6 eV it starts to increase strongly. This is consistent with the calculations which predict that the band gap becomes equal to the spin-orbit splitting (E_g=Δ=0.72 eV) at a pressure of ~19 kbar. This suggests that the increase in J_{th} is due to the onset of the strong CHSH Auger process already identified as being important in the 1.5μm devices.

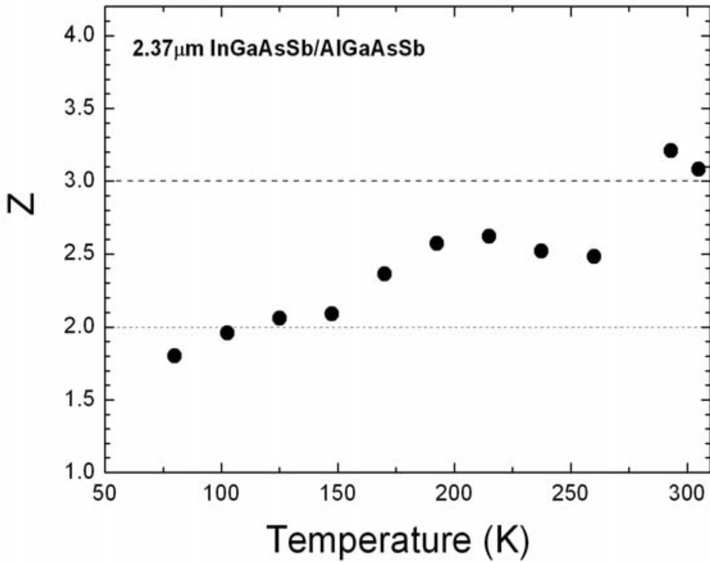


Figure 15. Z(T) for the 2.37 μm InGaAsSb laser. At low temperature, Z≈2 where $I_{th}=I_{rad}$. With increasing temperature Z rises steadily such that by room temperature, Z≈3 corresponding to a transition from radiative to Auger dominated behaviour.

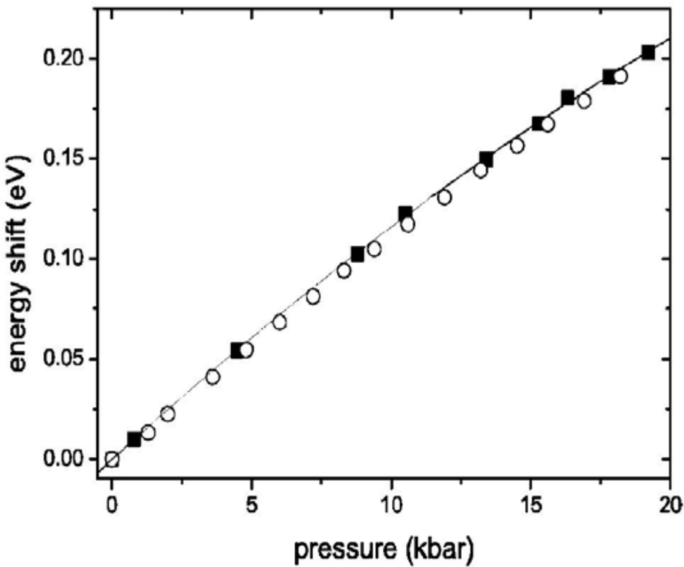


Figure 16. Lasing energy shift as a function of pressure for two 2.37 μm devices (*circles and squares*). For these devices, dE_{lase}/dP ≈ 12 meV/kbar, similar to that for the 1.5 μm devices. Reprinted with permission from [57] (Copyright 2001, American Institute of Physics).

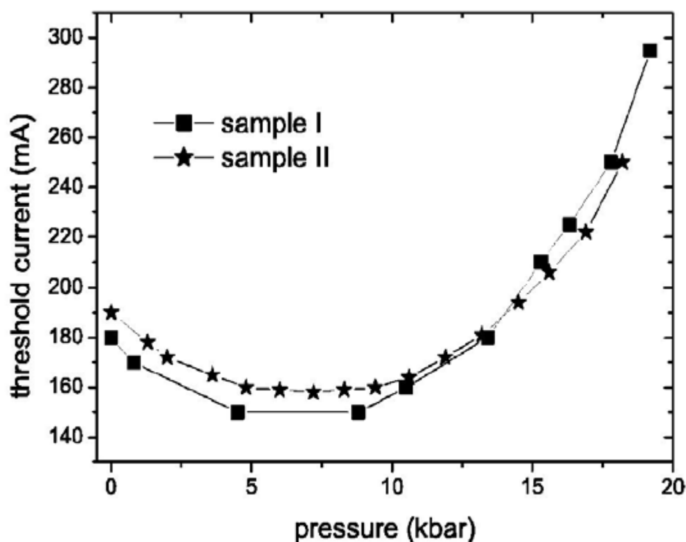


Figure 17. Pressure dependence of I_{th} for the 2.37 μm devices. At first I_{th} does indeed decrease with pressure. However, the rate of change soon decreases and above ~10kbar it starts to increase strongly. This is consistent with calculations which predict that the band gap becomes equal to the spin-orbit splitting at a pressure of ~19 kbar. Reprinted with permission from [57] (Copyright 2001, American Institute of Physics).

4 Conclusion

Determination of the electronic energy *vs* momentum relationship in semiconductors is essential for the prediction of almost all of their properties. In materials useful for mid-infrared applications, the simplest parabolic band approximations are usually insufficient. However relatively straight-forward numerical techniques based on the **k.p** method can yield good predictions for the bandstructure. The theoretical bandstructures can be compared with experiment using magneto-optics and magneto–transport, but one of the most useful tools for controllably tuning the system is hydrostatic stress. The strain modifies the bandstructure in a rather simple way, principally by a linear increase in the fundamental gap, and thus it can be used to separate out effects that depend on bandgap. A large literature has built up on the study of near-infrared optoelectronic devices under pressure. These have been used to establish the variations of the radiative and non-radiative Auger recombination processes with band structure. The results predict that III-V mid-IR lasers with direct band gaps less than the spin-orbit gap should have threshold current densities less than the near-IR lasers based on InP or GaAs. These predictions are found to be consistent with the threshold current density and its variation with pressure observed in InGaAsSb/AlGaAsSb operating at 2.37 μm at atmospheric pressure. Clearly high-pressure techniques provide exciting opportunities for the study of mid-infrared devices.

References

1. J. Faist *et al.*, *Semiconductors and Semimetals*, Vol 62, Academic Press (2000); M. Helm, ibid.
2. Zhu, XJ; Louie, SG, Physical Review B 43 14142 (1991)
3. E. O. Kane, in *Semiconductors and semimetals*, Vol 1 p75, Ed RK Willardson and AC Beer (Academic, New York, 1966)
4. J.M. Luttinger and W. Kohn, Phys Rev 97, 869 (1955)
5. C.R. Pidgeon and R.N. Brown, Phys Rev 147 575 (1966)
6. Vurgaftman, I; Meyer, JR; Ram-Mohan, LR, J Appl Phys 89 5815 (2001)
7. Adachi, S, J Appl Phys 61 4869 (1987)
8. M.H. Weiler in *Semiconductors and semimetals*, Vol 16, p119, Ed RK Willardson and AC Beer (Academic, New York, 1981)
9. U. Rössler, Solid State Commun. 49, 943 (1984).
10. P. Pfeffer and W. Zawadzki, Phys. Rev. B 41, 1561 (1990).
11. Talwar, DN; Loehr, JP; Jogai, B Phys Rev B 49 10345 (1994)
12. DL Mitchell and RF Wallis, Phys Rev 151 581-595 (1966).
13. S Yuan, G Springholtz, G Bauer and M Kreichbaum, Phys Rev B 49, 5476 (1994)
14. LR Ram-Mohan, KH Yoo and RL Aggarwal, Phys Rev B 38 6151 (1988)
15. Y. P. Varshni, Physica 34, 149 (1967).
16. J.F. Nye, *Physical properties of crystals*, (Clarendon Press, Oxford, 1964)
17. SW Ellaway and DA Faux, Phys. Stat Sol B 235, 437 (2003)
18. G. L. Bir and G. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, NY, 1974).
19. Bouarissa, N; Aourag, H, Mat Sci and Eng B 33 122 (1995)
20. Gorczyca, I; Pfeffer, P; Zawadzki, W Semicond Sci and Technol 6 963 (1991)
21. Daunov, MI Kamilov, IK Gabibov, SF Magomedov, AB Phys Stat Sol (b) 235, 297 (2003)
22. Adams, AR Dunstan, DJ Semicond Sci and Technol 5, 1194 (1990)
23. Dalven R (1973) Solid State Phys 28, 179
24. M Kriechbaum P Koecevar, H Pasdiev and G. Bauer, IEEE J. Quant. Electron. 24, 1727 (1988)
25. Voon, LCLY; Rammohan, LR Phys Rev B 47 15500 (1993)
26. Beattie, A.R., and Landsberg, P.T., Proc. Roy. London 249, 16 (1959).
27. Beattie, A.R., and White, A.M., J. Appl. Phys. 79, 802 (1996)
28. Adams, A.R., Electron. Lett. 22, 249 (1986)
29. Ashley T, Beswick J, Elliott CT, Jeffries R, Johnson A, Pryce G, IEE Proc. Optoelectronics 145, 265 (1998)
30. P. C. Findlay, C. R. Pidgeon, H. Pellemans, R. Kotitschke, B.N. Murdin, T. Ashley, A. D. Johnson, A. M. White, C. T. Elliott Semicond. Sci. Technol. 14, 1026-1030 (1999)
31. CM Ciesla, BN Murdin, TJ Phillips, AM White, AR Beattie, CJGM Langerak, CT Elliott, CR Pidgeon and S.Sivananthan. Appl Phys Lett 71, 491 (1997)
32. P Findlay, C R Pidgeon, B N Murdin, C J G M Langerak, C M Ciesla, J Oswald, A Homer, G Springholz and G Bauer, Phys. Rev. B58 12908 (1998)
33. Heiss W, Boberl M, Schwarzl T, Springholz G, Furst J, Pascher H, IEE Proceedings-Optoelectronics 150, 332 (2003)
34. Murdin BN, P Murzyn, C R Pidgeon, I V Bradley, J-P R Wells, T Burke, Y H Matsuda and N Miura, Appl. Phys. Lett. 81, 256 (2002)
35. Veal, TD; Mahboob, I; McConville, CF Phys Rev Lett 92 136801 (2004)
36. Lindsay A, O'Reilly EP, Solid State Commun. 112, 443 (1999)

37. Murdin, BN; Kamal-Saadi, M; Lindsay, A; O'Reilly, EP; Adams, AR; Nott, GJ; Crowder, JG; Pidgeon, CR; Bradley, IV; Wells, JPR; Burke, T; Johnson, AD; Ashley, T, Appl Phys Lett 78 1568 (2001)
38. Andreev, unpublished
39. Ciesla, CM; Murdin, BN; Pidgeon, CR; Stradling, RA; Phillips, CC; Bain, DJ; Galbraith, I; Jaroszynski, DA; Langerak, CJGM; Tang, PJP; Pullin, MJ, IEE Proceedings-Optoelectronics 144 331 (1997)
40. Hjalmarson, H.P., and Kurtz, S.R., Appl. Phys. Lett. 69, 949 (1996).
41. F. Phillips, S. J. Sweeney, A. R. Adams and P. J. A. Thijs, IEEE Jour. Sel. Top. Quant. Electr., 5, 401 (1999).
42. S. J. Sweeney, A. F. Phillips, A. R. Adams, E. P. O'Reilly and P. J. A. Thijs, IEEE Phot. Tech. Lett., 10, 1076 (1998).
43. E. P. O'Reilly and M. Silver, Appl. Phys. Lett., 63, 3318 (1993).
44. R. Adams, M. Silver and J. Allam, in High Pressure in Semiconductor Physics II, vol. 55, T. Suski and W. Paul, Eds., New York: Academic, 1998, pp301-352.
45. S. J. Sweeney, S. R. Jin, C. N. Ahmad, A. R. Adams and B. N. Murdin, Phys. Stat. Sol. (b), 241, 3399 (2004).
46. S. J. Sweeney, A. R. Adams, M. Silver, E. P. O'Reilly, J. R. Watling, A. B. Walker and P. J. A. Thijs, Phys. Stat. Sol. (b), 211, 525 (1999).
47. Charlebois, S Beerens, J Miner, CJ Puetz, N, Phys Rev B 54, 13456 (1996)
48. Cheong, HM Paul, W Flatte, ME Miles, RH, Phys Rev B 55, 4477 (1997)
49. Frogley, MD Downes, JR Dunstan, DJ, Phys Rev B 62, 13612 (2000)
50. Li, GH Goni, AR Syassen, K Brandt, O Ploog, K, Phys Rev B 50, 18420 (1994)
51. Choulis, SA Hosea, TJC Tomic, S Kamal-Saadi, M Adams, AR O'Reilly, EP Weinstein, BA Klar, PJ, Phys Rev B 66, 165321 (2002)
52. Webb, SJ Stradling, RA Nagata, K, Compound Semiconductors 1997, 156, 315 (1998)
53. Daly, MS Symons, DM Lakrimi, M Nicholas, RJ Mason, NJ Walker, PJ, Surf Sci 362, 205 (1996)
54. Roberts, M Chung, YC Lyapin, S Mason, NJ Nicholas, RJ Klipstein, PC, Phys Rev B 65, 235326 (2002)
55. J. M. Besson, J. F. Butler, A. R. Calawa, W. Paul, and R. H. Rediker, Appl. Phys. Lett., vol. 7, pp. 206–208, Oct. 1965; J. M. Besson, W. Paul, and A. R. Calawa, Phys. Rev., vol. 173, pp. 699–713, Sept. 1968.
56. Choulis, SA Andreev, A Merrick, M Adams, AR Murdin, BN Krier, A Sherstnev, VV, Appl Phys Lett 82, 1149 (2003)
57. P. Adamiec, A. Salhi, R. Bohdan, A. Bercha, F. Dybala, W. Trzeciakowski, Y. Rouillard and A. Joullie, Appl. Phys. Lett., 85, 4292 (2004).
58. M. Silver and E. P. O'Reilly, IEEE Jour. Quant. Electr., 31, 1193 (1995).
59. K. O'Brien *et al.*, to be published.

Part II

Lasers

III-Sb-based Type-I QW Diode Lasers

M. Rattunde, J. Schmitz, C. Mermelstein, R. Kiefer and J. Wagner
Fraunhofer-Institut für Angewandte Festkörperphysik
Tullastrasse 72, 79108 Freiburg, Germany

1 Introduction

There is an increasing number of applications, such as trace gas spectroscopy, medical diagnostics, laser surgery and material processing, which require compact and robust laser sources emitting in the 2-3 μm wavelength range. Because of the strong absorption lines of a variety of relevant technical gases in this wavelength range [1], high resolution spectroscopic gas detection for industrial process monitoring or environmental control can be realized using such diode lasers. Continuous wave (CW) output powers at room temperature of a few tens of mW are required for these applications, together with a good spectral purity. Furthermore, LIDAR-measurements and optical free-space telecommunications can be achieved through the 2-2.5 μm atmospheric window. Because of the distinct absorption spectra of human tissue [2], medical diagnostic, such as noninvasive optical blood glucose monitoring [3] and laser surgery [2] are other promising applications. For the latter, output powers exceeding 1 W are required. Given the wide range of applications of lasers in material processing, not all can be dealt with using state-of-the-art GaAs-based high-power diode lasers, emitting around 1 μm. For example the welding of transparent plastic demands high power lasers at 2 μm, an application which is currently addressed by Ho-YAG solid state laser systems [4]. High-power diode lasers emitting at around 2 μm, with their potential lower cost and compactness can replace those systems. Finally, optical pumping of solid state lasers systems [5,6] is another application of high power 2 μm diode lasers.

The III-V compound semiconductor material system (AlGaIn)(AsSb) constitutes an ideal basis for the realization of diode lasers in this wavelength regime. GaInAsSb, either lattice matched to GaSb or deliberately strained, can be used for the active layer with a direct band gap between 1.7 μm and well above 3 μm. For the barrier and cladding layers, AlGaAsSb is well suited because of its larger band gap energy and lower refractive index than GaInAsSb.

Starting with the first III-V mid-IR laser in 1963, based on InAs homojunctions and a low-temperature emission at 3.1 μm [7], double-heterostructure (DH) lasers were developed with GaInAsSb active region and AlGaAsSb barriers, grown on GaSb-substrates. These lasers showed excellent performance at room temperature in the 2.0 to 2.5 μm range [8,9]. For the 3-4 μm range, DH-lasers employing InAsSb / AlGaAsSb layer sequences on GaSb or InAsSb / InAsPSb layers on InAs-substrates were fabricated with a maximum operating temperature of around 170 K in pulsed operation [10-12]. In 1992, the concept of strained quantum-well (QW) lasers was for the first time implemented in GaSb-based lasers [13] yielding significant improvements in laser output power and maximum operating temperature [14]. Nowadays GaInAsSb /AlGaAsSb Type-I QW lasers can be operated at room

temperature up to wavelength of 3.04 μm [15]. Above 3 μm, Type-II laser concepts, such as the W-Laser [16] seem to have a better performance than Type-I QW-lasers with an InAsSb active layer, mostly because the potential suppression of the CHHS-Auger recombination in these Type-II laser structures [17]. In this chapter, we will focus on GaInAsSb / AlGaAsSb Type-I diode lasers grown on GaSb-substrates.

2 III-Sb-based Material System

2.1 AlGaAsSb

The quaternary semiconductors $Al_xGa_{1-x}As_ySb_{1-y}$ are the ideal materials for barrier-, waveguide- and cladding layers of III-Sb based Type-I diode lasers. Because of the large width of the latter layers, they are almost exclusively grown lattice matched to the GaSb-substrate, which is achieved by adding a small amount of As to the AlGaSb to form $Al_xGa_{1-x}As_ySb_{1-y}$ with $y = 0.08\,x$ [18]. The direct band gap for AlGaAsSb at 300 K, lattice matched to GaSb, is given by $E_g(\Gamma) = 2.297\,x + 0.727\,(1-x) - 0.48\,x\,(1-x)$ eV [18].

The ternary alloy $Al_xGa_{1-x}Sb$ has the specific property of changing the character of the fundamental band gap twice upon increasing Al-content: from direct, with the Γ-conduction band minimum being lowest in energy, to indirect, with the L-point minima being lowest in energy for Al-contents above 25 % and to the X-point minima being lowest for an Al-content beyond approx. 45 % [18-20]. For AlGaAsSb lattice matched to GaSb, a similar behavior is expected as only a small amount of As is added, although some calculations suggest a direct to indirect crossover at an Al-concentration of only 14 % [18].

2.2 Strained GaInAsSb Layers

For the active layers, $Ga_{1-x}In_xAs_ySb_{1-y}$ is used which has a direct bandgap for all alloy composition and is lattice matched to GaSb if the condition $y = 0.913\,x$ is satisfied [18,21]. Using a quaternary material for the active layer adds an additional degree of freedom for the design of QWs compared to a ternary material (such as the well known GaInAs): by changing the composition, two of the three relevant parameters bandgap E_g, strain ε_{zz}, and band offsets ΔE can be adjusted individually within certain limitations. To illustrate this property of the GaInAsSb material system, Figure 1 shows the band-edge profile for $Ga_{1-x}In_xAs_ySb_{1-y}$ grown on GaSb as a function of the In-content x for three cases with different As-contents y and thus different strain ε_{zz}. Starting with GaSb on the left side of each plot, the bandgap is decreased with increasing In-content, in all cases. Also shown are the band edges of AlGaAsSb lattice matched to GaSb with an Al content of 30 % as a reference. The latter represents a typical barrier material, used by many different groups in their GaSb-based QW lasers [22-24]. All material data for this calculation are taken from [18]. Without the addition of As in Figure 1a, that means for $Ga_{1-x}In_xSb$, the layers are compressively strained with $\varepsilon_{zz} = +0.063\,x$ and due to strain effects, a Type-I band alignment is formed with regard to the chosen barrier material in the whole composition range with the heavy hole (hh) band as the

topmost valence band. For material lattice matched to GaSb with y= 0.913 x (Figure 1b), the band offset in the valence band ΔE_V is reduced with increasing In-content x and leads to a Type-II band alignment with respect to this barrier material for x>0.35. In Figure 1c, tensile strained material is shown, with higher As- than In-concentration (y= 1.7 x) and a strain given by ε_{zz}= -0.063 x. In this case of tensile strain, the light hole band (lh) forms the topmost valence band and the band alignment is Type-I is again only for low In-concentrations (x<0.3). Therefore compressively strained GaInAsSb is favorable in order to allow sufficient confinement of the holes in the active QWs. This is especially important for GaSb-based lasers with longer wavelength towards 3 µm (see Section 6).

The use of strained material in general has the advantage that the hh-lh-degeneracy is removed, reducing the density of states at the valence band-edge.

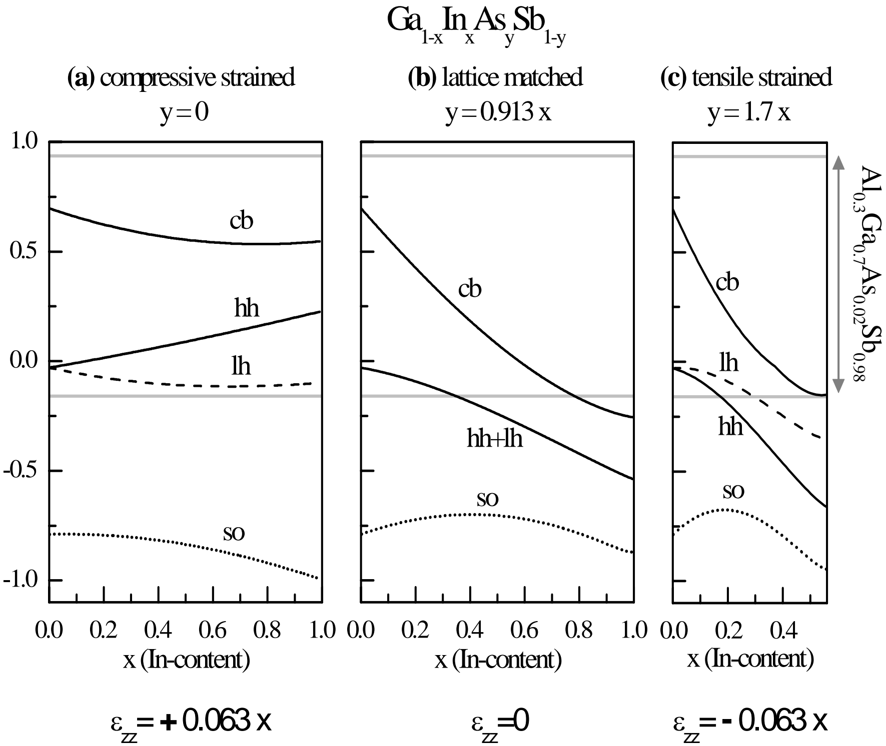


Figure 1. Relative positions of the conduction band (cb), heavy-hole (hh), light-hole (lh) valence band, and split-off (so) valence band vs In-content for the active region QW material $Ga_{1-x}In_xAs_ySb_{1-y}$ grown on GaSb; for comparison conduction and valence band edges of the $Al_{0.3}Ga_{0.7}As_{0.02}Sb_{0.98}$ barriers are also indicated (*gray lines*). Three representative cases are shown: (**a**) compressively strained QW material with y = 0 (**b**) lattice matched material with y = 0.913 x and (**c**) tensile strained material with y = 1.7 x. For the strained GaInAsSb in (a) and (c), the strain ε_{zz} is given by + 0.063 x and - 0.063 x, respectively.

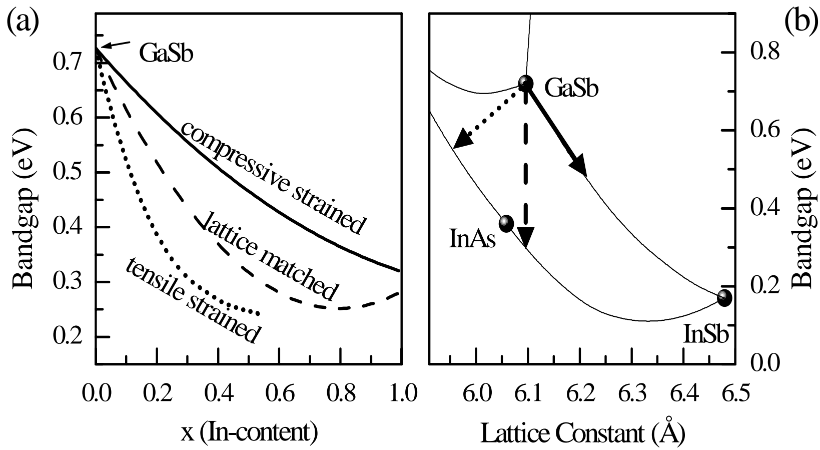


Figure 2. Bandgap E_g dependence of the In-content x for the three different cases of compressively strained (*solid line*), lattice matched (*dashed line*), and tensile strained (*dotted line*) $Ga_{1-x}In_xAs_ySb_{1-y}$ on GaSb as displayed in Figure 1. In (b), the bandgap vs. lattice constant relation is illustrated for these three different cases.

Compressively strained material has the additional advantage, that the in-plane effective mass of the heavy hole is very small [25,26], leading to a further reduction of the 2D-density of states in the valence band and therefore a further equalization ofthe density of states in the conduction and valence band. This reduces the threshold carrier density N_{th} and increases the differential gain dg/dN of the active layer [27]. In [28], a detailed theoretical examination of lattice matched, compressive- and tensile strained GaInAsSb QWs is carried out.

For the three different cases of compressive strained, lattice matched and tensile strained material discussed above, the bandgap E_g of GaInAsSb is plotted vs the In-concentration in Figure 2a and the bandgap vs lattice constant relation is illustrated in Figure 2b. In the case of compressively strained GaInSb, the bandgap decreases at a smaller rate with increasing In-content than for the other two cases (Figure 2a), *i.e.* the addition of As leads to an additional reduction of the bandgap. This is due to bowing and strain effects, as compressive strain increases E_g. However, the addition of As decreases the band offset in the valence band ΔE_V (see Figure 1). Note that for all calculations shown in Figures 1 and 2 pseudomorphic growth of the GaInAsSb-layer is assumed despite the fact that for strained material, the critical layer thickness approaches zero for increasing In-content.

2.3 Laser Structure

A typical III-Sb based Type-I laser structure for an emission wavelength of 2.0 µm is shown in Figure 3 [29]. For most of the laser structures reported on here, one to three compressively strained QWs are used as the active region with a width of 10-20 nm [15,22,30]. They are separated by 20-30 nm thick AlGaAsSb barrier layers with an Al-content of 20 % - 40 % [31] which are either lattice matched to the GaSb substrate or tensile strained in order to reduce the average strain in the active region

[32,33]. The AlGaAsSb spacer or separate confinement layers (SCL), which are lattice matched to GaSb and each between 20 and 600 nm thick [34,35] have usually

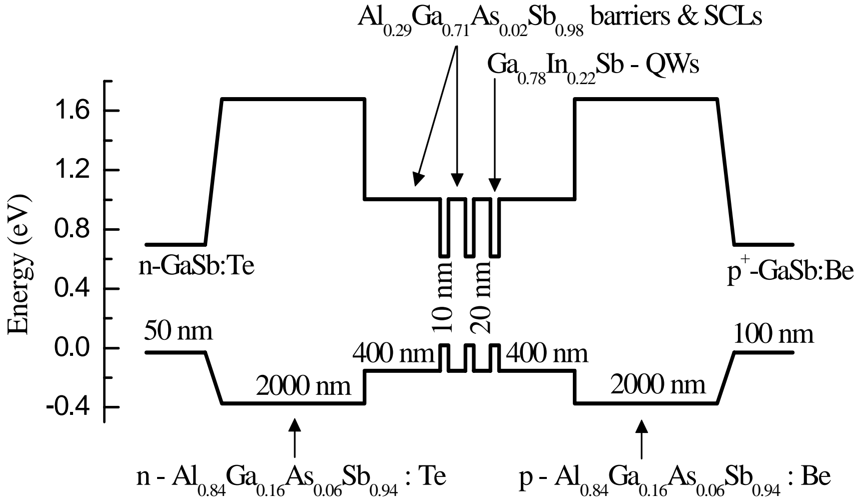


Figure 3. Direct gap band-edge profile of a GaSb-based Type-I laser structure with an emission wavelength of 2 µm [29]

the same Al-content as the barrier layer material. This waveguide core is nominally undoped and therefore weakly p-type in the range of 1×10^{16} cm^{-3} because of the native acceptors in GaSb-based material [36]. The AlGaAsSb cladding layers exhibit a higher Al-content than the barrier layers and SCLs in order to confine the optical mode to the waveguide core as well as to act as an additional electronic barrier for the carriers. An Al-content of 100 % in these layers is not advisable due to the high oxidation rate of AlSb, therefore typical values are in the range of 85-90 % [37,38]. The AlGaAsSb cladding layers are n-doped with Te (bottom) and p-doped with Be (top). For the p-doped layer, it is advantageous to lower the Be-concentration towards the waveguide layer in order to reduce the waveguide loss due to the free carrier absorption (see Section 4).

For the 2.0 µm laser structure presented in Figure 3, the resulting refractive index profile is shown in Figure 4a together with the calculated optical mode intensity. The confinement factor Γ_{QW} of the optical mode with the QWs, which determines the modal gain $\Gamma_{QW}\cdot g$ of the laser, yields a value of $\Gamma_{QW} = 4.5$ or 1.5 % per QW. The confinement factor with the highly doped cladding layers is $\Gamma_{Clad} = 10$ %. The overlap of the optical mode with the doped regions of the laser structure determines predominantly the optical losses in the laser structure (see Section 4). In Figure 4b, the calculated (line) and measured (symbols) far-field characteristic in the vertical direction (fast axis) are shown for the laser structures, mentioned above. In agreement with the simulated characteristic, the measurement shows a large beam divergence with a full width at $1/e^2$ of 120°. This large

divergence is the drawback of this waveguide structure, which is designed in order to supply a high confinement factor Γ_{QW} per QW and low optical losses in the doped cladding layers.

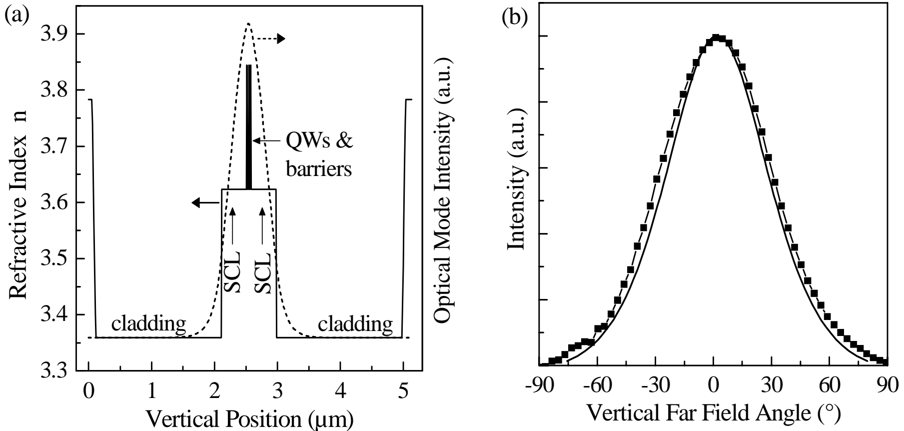


Figure 4. (**a**) Refractive index profile (*solid line*) for the 2.0 µm laser structure together with the optical mode intensity (*dashed line*), (**b**) Calculated (*solid line*) as well as measured (*symbols*) fast-axis far-field characteristics for this laser

3 Fabrication of (AlGaIn)(AsSb)-based Diode Lasers

3.1 Growth

Although there are some reports of GaSb-based lasers grown by metal-organic vapor phase epitaxy (MOVPE) [39], molecular beam epitaxy (MBE) is most widely used [14,40]. The laser structures are grown on (100)-oriented n-doped GaSb substrates. The MBE system is typically equipped with valved cracker effusion cells for the group-V elements As and Sb, providing precisely controlled fluxes of As_2 and predominantly Sb_2 as required for the growth of mixed group-V ternaries and quaternaries. The group-III elements Al, Ga, and In as well as the p- and n-type dopants Be and Te are supplied from conventional effusion cells, where GaTe or Sb_2Te_3 is used as the Te doping source [41]. The lower cladding layer is grown at higher temperatures (around 550°C) than the laser core and the top-cladding layer (around 475°C). Growing the top cladding layer at a lower temperature than the bottom cladding ensures that the GaInAsSb-QWs do not degrade due to In-inter-diffusion during growth of the former [40].

The barrier and cladding material AlGaAsSb lattice matched to GaSb is in the stable growth regime for all Al-concentrations for growth temperature above 400 °C [14]. However, the active layer material GaInAsSb is not miscible in its entire composition range. In Figure 5 the calculated spinodal isotherms for GaInAsSb are shown, indicating the growth temperature dependent boundaries for unstable alloys

[42]. For material lattice matched to GaSb, there is a miscibility gap between approximately 20 and 85 % In-content for the growth temperature mentioned above. MBE growth is still possible within this range, but the material quality is usually inferior. Compressively strained GaInAsSb with high In and low As concentration has the advantage of avoiding this miscibility gap to some extent.

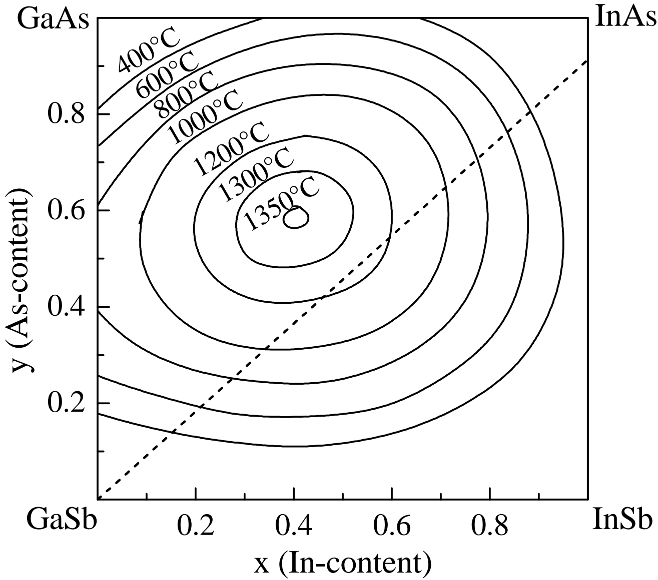


Figure 5. Temperature-dependent spinodal isotherms for $Ga_{1-x}In_xAs_ySb_{1-y}$ (*solid lines*) after Onabe [42]. The *dashed line* indicates the composition for lattice-matching to GaSb.

3.2 Epitaxial Layer Characterization

The epitaxial layer structures are routinely analysed by high-resolution X-ray diffraction (HRXRD), photoluminescence-spectroscopy (PL) and secondary ion mass spectroscopy (SIMS). In Figure 6a, a typical HRXRD profile for a triple-QW diode laser is shown, covering the (004) reflection range of GaSb. The diagram reveals a well-resolved superlattice (SL) diffraction peak signature of the triple-QW active region along with narrow diffraction peaks associated with the AlGaAsSb cladding and separate confinement layers. From a detailed analysis of the measured HRXRD profile with the help of simulated diffraction profiles (see lower curve in Figure 6a), the individual thicknesses of QW and barrier layers can be determined as well as the individual lattice mismatch perpendicular to the growth plane $(\Delta a/a_0)_\perp$. From the latter value, the strain ε_{zz} can be determined and thus the strain effects on the band structure can be calculated [43]. The residual lattice mismatch of both cladding layers is below $(\Delta a/a_0)_\perp < 1\times10^{-3}$, indicating excellent lattice matching and thus high crystalline quality. A typical calibration procedure for the growth of AlGaAsSb lattice matched to GaSb is to grow first AlGaSb with the targeted Al-

content and then to add As to the alloy by increasing the As flux up to the point where lattice matching is achieved.

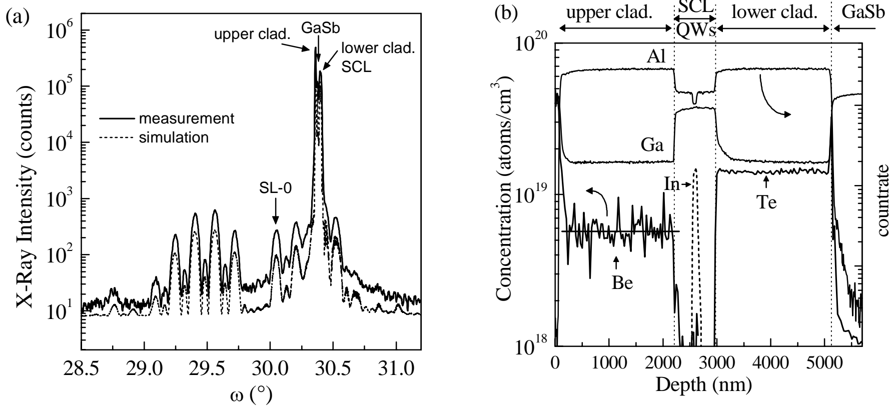


Figure 6. (**a**) Experimental HRXRD profile (*upper curve*) and simulated profile (*lower curve*) covering the (004) reflection range of GaSb for a triple-QW diode laser structure. (**b**) SIMS depth profiles of individual lattice constituents Al, Ga and In (*right axis*), and of the p- and n-dopants Be and Te respectively (*left axis*), for the same laser structure.

In the case of strained quaternary $Ga_{1-x}In_xAs_ySb_{1-y}$ as used for the QW active region, the layer composition cannot be determined from the HRXRD analysis alone. Additional input is needed, such as, *e.g.* the Ga-to-In concentration ratio which can be obtained from group-III flux calibrations. The concentrations of the group-V elements As and Sb are then calculated from the lattice mismatch $(\Delta a/a_0)_\perp$ as obtained from HRXRD. An alternative approach is to use the PL-emission wavelength as an additional input parameter besides the lattice mismatch to characterize the QW active layer. In Figure 7 the lattice mismatch $(\Delta a/a_0)_\perp$ as well as the resulting strain ε_{zz} is plotted vs the calculated PL emission wavelength for different $Ga_{1-x}In_xAs_ySb_{1-y}$ layer compositions (solid lines). Quantisation effects (assuming 10 nm wide QWs and $Al_{0.3}Ga_{0.7}As_{0.02}Sb_{0.98}$ barriers) and strain effects were included in the calculation. The material data are taken from [18] and the correlated function expansion has been used to interpolate these data to represent the quaternary compound [21]. A mesh of iso-composition contour lines is plotted in Figure 7, computed for either constant In- or As-content for In concentrations from 10 to 50% and As-contents from 0 to 30%. Further included in Figure 7 are experimental strain / PL-wavelength data from actual laser structures together with their composition as determined from MBE calibration curves. There is satisfactory agreement between the calculated and experimental data, given the uncertainties in alloy composition and input parameters for the calculations.

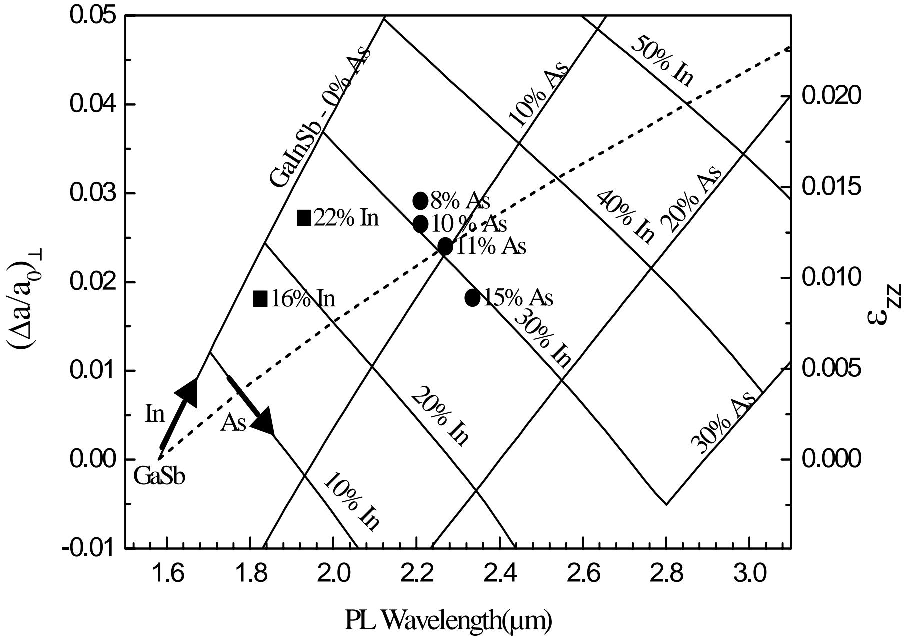


Figure 7. Lattice mismatch $(\Delta a/a_0)_\perp$ as well as the resulting strain ε_{zz} vs. the PL emission wavelength for different $Ga_{1-x}In_xAs_ySb_{1-y}$ QW compositions. The *solid lines* are calculated dependencies, including quantisation (assuming 10 nm wide QWs and $Al_{0.3}Ga_{0.7}As_{0.02}Sb_{0.98}$ barriers) and strain effects. The *symbols* are measured data from laser structures with 30 % In-content, varying As-content (*circles*), 0 % As-content and varying In-content (*squares*), respectively. The *dashed curve* is derived with the condition that the As-content y is 0.32 times the In content x (see Section 6).

Because of the native acceptors in GaSb [36], there is no semi-insulating substrate material available, making Hall-measurements to calibrate the doping concentration a difficult task. A different approach is to use secondary ion mass spectroscopy (SIMS) as shown in Figure 6b. The Te- and Be-countrate were calibrated using AlGaAsSb standards with well defined implanted Te and Be concentrations. The free electron and free hole concentration can be calculated from the measured Te and Be concentrations [43] using the effective density of states at the band edge [18] and the ionization energy of the Te-donor (68meV [20]) and Be-acceptor (38meV [44]), respectively, neglecting electrical compensation.

3.3 Device Processing

Depending on the required output power, beam quality and spectral purity, the GaSb-based epitaxial layer sequences are processed into different types of laser structures. Narrow ridge-waveguide (RW) lasers provide a single lateral mode and quasi single longitudinal mode lasers [45,46] which can be used for spectroscopic applications. In order to increase the side-mode suppression ratio (SMSR) as well as the range for stable single-mode operation, distributed feedback (DFB) lasers have been fabricated, employing a lateral metal grating at the side of a narrow ridge-

waveguide structure. In this way, a SMSR of 33 dB together with a maximum output power of 8.5 mW has been achieved with an emission wavelength of 2.4 μm [47]. A conventional DFB-processing scheme with a Bragg grating etched into the top SCL followed by overgrowth of the top cladding is not suitable in the GaSb-based material system because of the high Al-content materials used which results in the formation of stable aluminium oxide [41,47].

Another type of semiconductor laser, which provides single-frequency emission, is the vertical cavity surface emitting laser (VCSEL). Electrically pumped as well as optically pumped GaSb-based VCSEL structures have been reported, operating in pulsed mode and quasi-CW at room temperature respectively with an emission wavelength of 2.1-2.2 μm and output powers of a few mW [48,49].

For high output power, gain-guided broad-area (BA) lasers are used, either as single emitters or as laser bars, the latter emitting up to 10 W in CW operation at 0°C [50]. To improve the beam quality of high power lasers, the concept of tapered lasers, composed of a ridge-waveguide section and a trapezoidal gain region has been employed also for GaSb-based lasers. Single emitters of this type of device have been shown to be able to emit 0.6 W of diffraction limited output power at 2.05 μm in CW operation [51]. Tapered laser arrays made out of these devices with anamorphic collimating lenses were able to emit 1.7 W peak power in pulsed mode within a 65 mrad cone [52].

4 Gain and Loss Mechanisms in III-Sb-based Lasers Emitting at Around 2 μm

In the following section, we will discuss the optical gain and dominant electric and optical loss mechanisms for III-Sb based Type-I QW lasers with an emission wavelength of around 2 μm at room temperature, before addressing the high power performance of these lasers in Section 5.

4.1 Optical Gain and Electric Loss in the QW Active Region

GaSb-based lasers with an emission wavelength of 2.0 μm at room temperature can be grown with ternary, compressively strained QWs using $Ga_{1-x}In_xSb$ with an In concentration of 22 %. These layers exhibit a lattice mismatch perpendicular to the growth plane of $(\Delta a/a)_\perp = 2.8$ % and a compressive strain of $\varepsilon_\perp = 1.3$ % [35]. In order to investigate the optical gain, three different laser structures were examined with one, two and three 10 nm wide $Ga_{0.78}In_{0.22}Sb$ QWs. The other parts of the structure were left unchanged as displayed in Figure 3. The modal gain Γg_{th} at threshold can be calculated using the internal losses α_i, measured by analysing lasers of different resonator length L, and the mirror losses α_m, given by $\alpha_m = -1/2L \ln(R_1R_2)$ with the reflectivity of the facets R_1 and R_2. The thus obtained value of Γg_{th} is plotted in Figure 8a vs the current density in the active region $\eta_i \cdot J_{th}$, together with a fit of the logarithmic gain model $\Gamma g(J) = \Gamma g_0 \cdot \ln(\eta_i \cdot J/J_{tr})$ [53] to the experimental data. Using $\eta_i \cdot J_{th}$ instead of J_{th}, ensures that the current losses outside of the active region are separated and that only the current density in the active region itself is taken into

account [53]. It can be seen that the transparency current density scales with the number of QWs and that the differential gain increases with increasing number of QWs, leading to an intersection of the different gain curves at a modal gain of approx. 50 cm^{-1}.

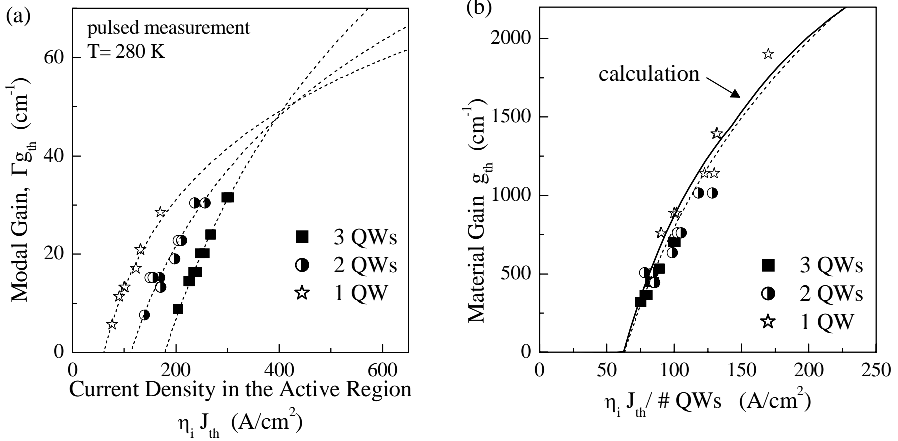


Figure 8. (**a**) Modal gain at threshold vs current density in the active region for a set of lasers with different numbers of QWs. (**b**) Material gain at threshold vs current density per QW. The data points are fitted, using a logarithmic gain model (*dotted lines* in (a) and (b)) and a detailed gain calculation (*solid line* in (b)).

Using the calculated confinement factor Γ_{QW} of the QWs, the material gain can be deduced, which is displayed in Figure 8b versus the current density per QW. All experimental data points can be fitted by the logarithmic gain model (dotted line) using a single set of parameters, yielding a transparent current density of J_{tr}= 63 A/cm^2 and a gain coefficient of g_0= 1700 cm^{-1}. These curves represent the optical gain of the $Ga_{0.78}In_{0.22}Sb$-QW material over a wide range of threshold gain and thus current density, showing an almost linear behavior up to a gain of approximately 2000 cm^{-1} without any severe gain saturation.

Fundamentally, the carrier density N determines the optical gain of the QW-layer, which can be calculated using Fermi's golden rule [27]. The required current density J to establish this carrier density N is then determined by the electric loss mechanisms in the QWs: the defect assisted recombination, the spontaneous emission and the Auger-recombination. With the corresponding phenomenological recombination coefficients A, B, and C, respectively, the threshold current density can be expressed as

$$J_{th} = qd(AN + BN^2 + CN^3) \quad (1)$$

where q is the elementary charge and d the width of the QW. For the $Ga_{0.78}In_{0.22}Sb$-QW material, the gain g(N) was calculated using the band structure and gain relationships described in [26,53] and material parameters given in [18]. A calculated spontaneous emission coefficient of $B=2.5\times10^{-10}$ cm^3s^{-1} was used,

together with an Auger coefficient of $C=5\times10^{-29}$ cm^6s^{-1}, taken from [54]. The coefficient for the defect assisted recombination A was used as a fit parameter. In this way, the measured gain g(J) in Figure 8b could be reproduced fairly well (solid line in Figure 8b), yielding a defect assisted recombination coefficient of $A=8\times10^8$ s^{-1}.

Comparing the different recombination mechanisms, one can conclude, that the defect assisted recombination is dominant for these 2.0 μm GaSb lasers close to room temperature (280 K): with a rate of 8×10^8 s^{-1}, this loss mechanism exhibits the highest rate compared to the spontaneous emission (1.4×10^8 s^{-1}) and Auger-recombination rate (0.3×10^8 s^{-1}; all values are calculated for a threshold gain of g=640 cm^{-1}). At elevated temperatures, however, the carrier density N, required to provide the threshold gain, is increased due to the broadening of the Fermi-function. As the Auger-recombination rate exhibits a cubic dependence on the carrier density, the increase in N has the most pronounced effect on this loss mechanism, leading to a strong increase of the threshold current at higher temperatures. For GaSb-based lasers emitting at 2.3 μm, this effect has been observed by measuring the spontaneous emission intensity through the side of the diode laser in dependence of current and temperature [55]: below 65°C, the threshold current was found to be dominated by a monomolecular nonradiative recombination process, whereas above 65°C, the Auger recombination is the dominant effect.

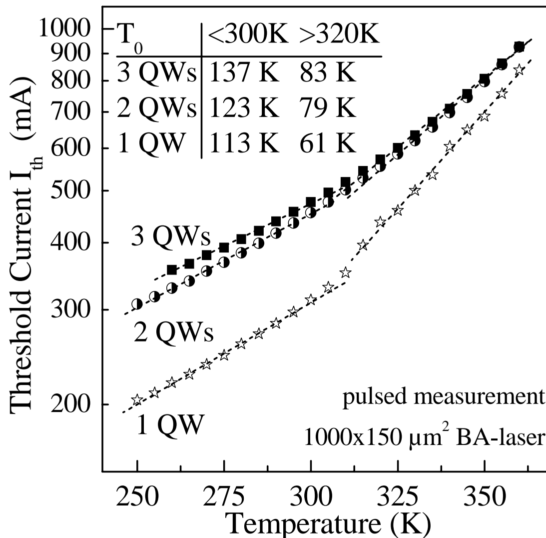


Figure 9. Threshold current vs temperature for a set of lasers with different numbers of QWs. Data are shown for 1000×150 μm² BA lasers which were measured in pulsed operation.

Considering the modal gain of these GaSb-based 2.0 μm lasers shown in Figure 8a, it can be seen that the lowest threshold current densities J_{th} can be achieved with single-QW (SQW) lasers, as typical values for the total losses $\alpha_i+\alpha_m$ do not exceed 50 cm^{-1}, even for short resonators around 500 μm length. The lowest value for J_{th} reported so far was 50 A/cm² for a 3000 μm long SQW-laser, emitting at 2.05 μm at

room temperature [54]. But the drawback of these SQW-devices is the higher temperature sensitivity of the threshold current, as reflected by a lower value of the characteristic temperature T_0. In Figure 9, the temperature-dependent threshold current of 1000 μm long devices is displayed for the laser structures discussed above with one, two and three QWs in the active region. In the high temperature regime of 320 – 360 K, the SQW lasers exhibit a T_0 value of 61 K, rising to 79 K and 83 K for the two and three QW laser, respectively. Therefore, for the optimization of the laser structure for a given application, electric, optical and thermal effects have to be included in order to decide, if a higher temperature stability (multiple-QW-laser, MQW) or a lower threshold current (SQW) is more favorable. The higher temperature stability of the MQW-lasers is attributed to the reduced carrier density in the QWs, leading to a reduced Auger recombination. Additionally, the thermal leakage current out of the QWs is reduced, as this current does not scale with the number of QWs and is therefore more pronounced in SQW lasers [53, 56].

4.2 Optical Losses in the Laser Structure

The most important optical loss mechanisms for the photons generated in the laser cavity are the free-carrier absorption and the intervalence band absorption. In the case of the free-carrier absorption, the photon is absorbed by a free electron or hole, which is excited to a higher energy state in the conduction or valence band respectively. Due to momentum conservation, this intraband transition can only take place with the help of an additional interaction with a phonon or an impurity. This loss mechanism increases in strength with increasing wavelength proportional to λ^k with k in the range of 2 to 3.5 [41,57]. In the case of the intervalence band absorption, the absorbed photon energy excites an electron from a filled state in a lower lying valence band to an empty state in a higher lying valence band (interband transition). As both optical loss mechanisms are assumed to be proportional to the free carrier density [58], they can be combined into the free-carrier (FC)-losses, given by

$$\alpha_{FC} = \sigma_N \Gamma_N N + \sigma_P \Gamma_P P \tag{2}$$

where N and P are the free electron and hole concentration, Γ the confinement factor of the optical mode with the respective region and σ_N, σ_P the cross-section for free electron or hole absorption, respectively. In the GaAs and InP based material systems, these losses are more pronounced in p-doped, than in n-doped regions [58]. In order to analyze and quantify the losses in the highly doped cladding layers of GaSb-based lasers, five different laser structures were examined, all with three 10 nm wide $Ga_{0.78}In_{0.22}Sb$-QW in the active region, emitting at 2.0 μm at room temperature. The 20 nm wide barriers and the separate confinement layers consist of $Al_{0.29}Ga_{0.71}As_{0.02}Sb_{0.98}$, the n- and p-doped cladding layers are made up of $Al_{0.84}Ga_{0.16}As_{0.06}Sb_{0.94}$. For the reference structure (A), 400 nm wide SCLs were used and the lower and upper cladding layers were doped with 1.4×10^{19} cm^{-3} Te and 6.0×10^{18} cm^{-3} Be respectively. The differences of these five samples are illustrated in Figure 10a, b. In the first set of samples (A, B, C) (Figure 10a), the width of the SCLs were fixed at 400 nm while the doping concentration in the cladding layers

was varied. Starting from the reference structure (A), the doping was lowered in the inner 500 nm wide portions of the 2 µm thick cladding layers to 5×10^{18} cm^{-3} Te and 8.0×10^{17} cm^{-3} Be for structure (B), whereas in sample (C), a higher p-doping of 1×10^{19} cm^{-3} Be was employed in the top p-cladding layer, while the bottom n-cladding layer was as for sample (A). For the second set of samples (Figure 10b), the width of the SCLs was varied from 400 nm for the reference structure (A) to 155 nm (D) and 600 nm (E), resulting in a total width of the waveguide core region of 870, 380 and 1270 nm, respectively [35].

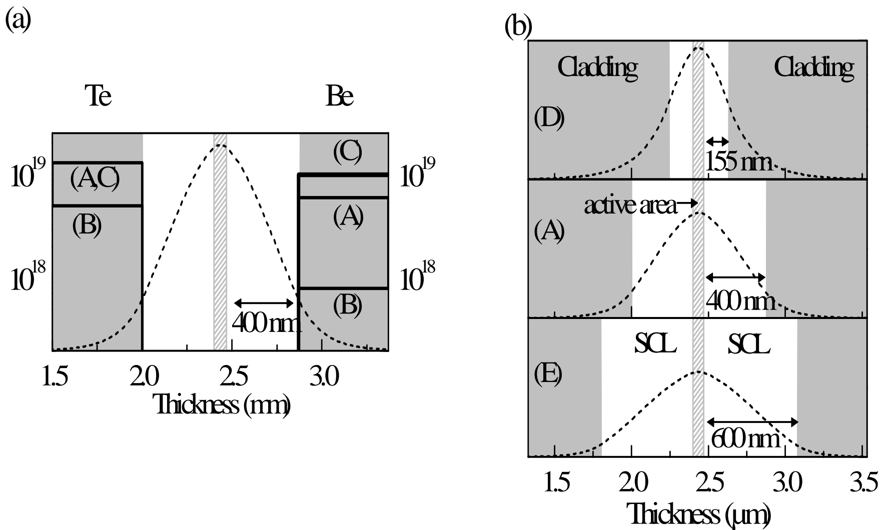


Figure 10. Optical mode intensity profile within the laser structure (*dashed lines*) for the five different samples A to E. In (**a**), the Te- and Be doping levels in the lower and upper cladding layers of samples A, B and C are indicated, (**b**) shows the effect on mode confinement.

To examine the internal parameters of these laser structures, the inverse of the differential quantum efficiency η_d is plotted in Figure 11a, b vs the cavity length. For the first set of samples (A, B, C) with different doping profiles (Figure 11a), the intersection of the linear fit to the experimental data with the ordinate, which yields the internal efficiency η_i, is almost identical for all three samples, giving high values for η_i in the range of 82 % to 86 %. In contrast to that, the slope, and thus the internal losses α_i, shows a strong dependence on the doping profile with values ranging from 10.8 cm^{-1} down to 5.9 cm^{-1}. In Figure 11 b, equivalent data are shown for the second set of samples (A, D, E) with different widths of the separate confinement layers. Reducing the width of the waveguide core (structure D) leads to a dramatic increase of the internal losses to $\alpha_i = 18.4$ cm^{-1} compared to $\alpha_i = 8.8$ cm^{-1} for reference structure (A), whereas the sample with the widest waveguide core (E) exhibits the lowest losses of $\alpha_i = 6.5$ cm^{-1} and an unchanged high value for the internal efficiency of $\eta_i = 85$ %..

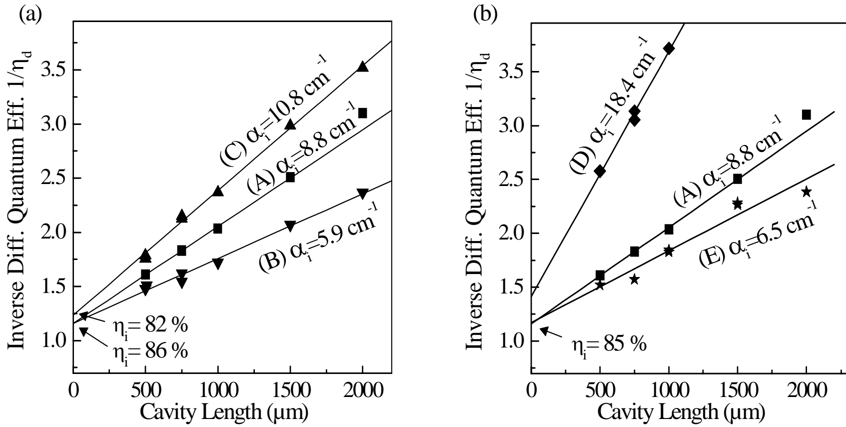


Figure 11. Inverse differential quantum efficiency (η_d^{-1}) vs cavity length for the set of samples with different doping-profiles in the cladding layers (**a**) and with different widths of the SCL (**b**)

The observed strong variation of α_i in these sets of five samples arise from the different absorption losses due to free carriers in the cladding layers (see Equation 2). In the first series (A,B,C) the doping level, and thus the free carrier concentration N and P in the cladding layers is varied, while in the second series (A,D,E) the overlap of the optical mode with the cladding layers Γ_N and Γ_P is changed. Assuming that the losses in the p-doped region are dominant compared to the n-doped region, the cross-section σ_P can be determined by plotting the measured internal losses α_i versus the product $\Gamma_P \cdot P$ (Figure 12a). The experimental data for all five different samples (A) to (E) can be fitted by one linear dependence, from which the cross-section for free hole absorption in $Al_{0.84}Ga_{0.16}As_{0.06}Sb_{0.94}$ can be deduced to be $\sigma_P = 4.6\times10^{-17}$ cm^2. This cross-section for the free hole absorption in 2.0 μm GaSb-based lasers is comparable to values reported for (AlGaIn)(AsP)-based 1.55 μm lasers, ranging from 2×10^{-17} cm^2 to 6×10^{-17} cm^2 [59-61].

The analysis shown in Figure 12a further confirms the initial assumption, that the dominant losses originate from the p-doped layer: In sample (C), only the p-doping level was varied, whereas the changes in all other samples affect the overlap of the optical mode with the free carriers in both, the n- and p-doped cladding. If the losses in the n-doped region had a significant effect on α_i, the experimental value for sample (C) should be well below the linear fit to the other samples, which is not the case. Extrapolating the linear fit in Figure 12a to $\Gamma_P \cdot P \to 0$ determines the remaining losses apart from the contribution of the p-cladding, yielding a value of 4.8 cm^{-1}. These losses originate from the QW active region and the SCLs.

To analyze these remaining losses further, we examined a third set of samples, where the number of QWs was reduced from three for the reference structure (A) to two (F) and one (G). From the experimental values for the internal losses α_i, we subtracted the known contributions of the p-cladding layer α_c as obtained from the

above analysis (see Figure 12a), and plotted the remaining optical losses α_i - α_c in Figure 12b vs the number of QWs.

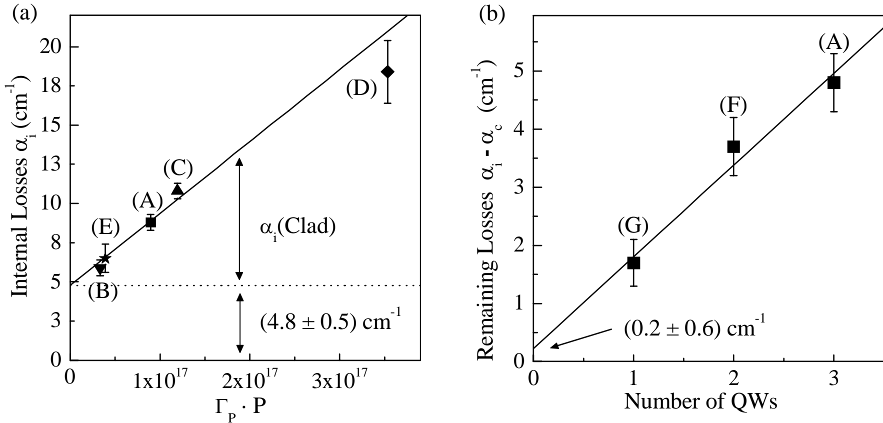


Figure 12. (**a**) Internal losses α_i for the samples (A) to (E) vs confinement factor with the p-doped cladding layers Γ_P times the free hole concentration P in this layer. All experimental data can be fitted by one linear curve (*solid line*), yielding a cross-section for free hole absorption in $Al_{0.84}Ga_{0.16}As_{0.06}Sb_{0.94}$ of $\sigma_P = 4.6\times10^{-17}$ cm^2. (**b**) Remaining internal losses α_i - α_c, where α_c denotes the losses due to free hole absorption in the p-cladding layer, vs the number of QWs. From the linear fit (*solid line*) an average optical loss of 1.5 cm^{-1} per QW is deduced.

The data points can be described very well by a linear fit with an average optical loss of 1.5 cm^{-1} per QW. These losses arise from a combination of free carrier absorption due to the threshold carrier concentration in the QWs and from scattering losses at the QW-barrier interface. The present experiments do not allow to differentiate between these two possible loss mechanisms. The intersection of the linear fit in Figure 12b with the ordinate is close to zero, with a fitted value of 0.2 ± 0.6 cm^{-1}. This indicates that the losses in the only remaining part of the laser structure, the SCLs, are negligible.

5 High-power Performance of 2 µm III-Sb-based Diode Lasers

The above analysis of 2 µm III-Sb based QW diode lasers revealed a high enough gain and sufficiently low distributed losses for high-power applications of this type of laser. Design and fabrication of high-power diode lasers, however, requires not just the optimization of one laser parameter, such as, *e.g.* threshold current, but the simultaneous optimization of a multitude of interrelated parameters, such that, for a given output power, the overall power- (or wall plug-) efficiency is maximized. To that end, high reflection / antireflection (HR/AR) coatings have to be applied to the rear / front facet of the laser to achieve a single-ended output as well as to enhance

the differential efficiency η_d at the expense of a certain increase in threshold current. But as high-power diode lasers are operated at current densities well above threshold, this increase can be tolerated. Further, as the maximum power efficiency η_{Power} of 2 μm III-Sb based diode lasers is in the 25 - 30% range, and typical values of η_{Power} at high output powers are more in the 15 - 20% range (see below), efficient heat removal from the active region by appropriate thermal management becomes an important issue.

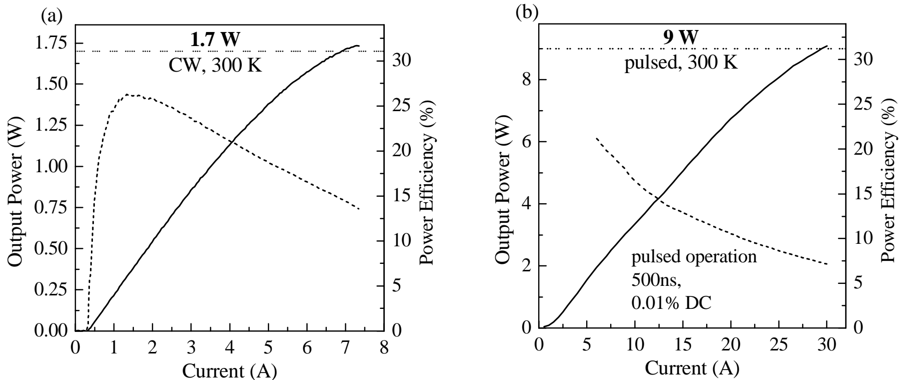


Figure 13. (**a**) CW- and (**b**) pulsed (0.5 μs pulse width, 0.01 % duty cycle) output power and power efficiency *vs* current characteristics of a 1000x150 μm^2 broad-area laser, emitting at 2.0 μm, with coated facets and mounted p-side down. Measurements were performed at 300 K heatsink temperature.

Representative output power vs current (P-I) and η_{Power} vs current characteristics of a broadened waveguide 2 μm GaInSb/AlGaAsSb triple-QW diode laser are shown in Figure 13a, b for 300 K CW and pulsed operation, respectively [35]. The facets of the 1000 μm × 150 μm broad-area diode laser were HR/AR-coated and the device was mounted p-side down for improved heat sinking. A maximum CW output power of 1.7 W was achieved, limited by thermal roll-over, and the maximum power efficiency was 27%, dropping to 15% at maximum output. Using short current pulses (500 ns pulse length, 0.01% duty cycle) to overcome thermal limitations, a peak output power of at least 9 W could be achieved (Figure 13b), this time limited by the maximum current of 30 A of the power supply. These data show that the present diode lasers can be operated in pulsed mode at currents up to 90 times the threshold current without indication of any facet degradation. The severe drop in power efficiency arises from an increase in operating voltage due to contact and series resistances in the substrate, the cladding and p-contact layers. Therefore, further optimization will be required while avoiding an increase in distributed losses due to free carrier absorption in the cladding layers. Figure 14 displays P-I and η_{Power} vs current characteristics of three 1000 μm × 150 μm broad area diode lasers with different compositions of the GaIn(As)Sb QWs in the triple-QW active region such that the respective lasing wavelengths are 1.9, 2.0 and 2.1 μm. All devices show 280 K CW output powers exceeding 1 W with maximum power efficiencies around 25%, indicating that a good high power performance can be achieved over an extended wavelength range. Actually, a CW output power of 1 W has been

demonstrated also at 2.5 μm for a heatsink temperature of 285 K, using two 1.6% compressively strained 14.5 nm wide $Ga_{0.59}In_{0.41}As_{0.14}Sb_{0.86}$ QWs as the active region [62]. However, the maximum power efficiency was reduced to 12%.

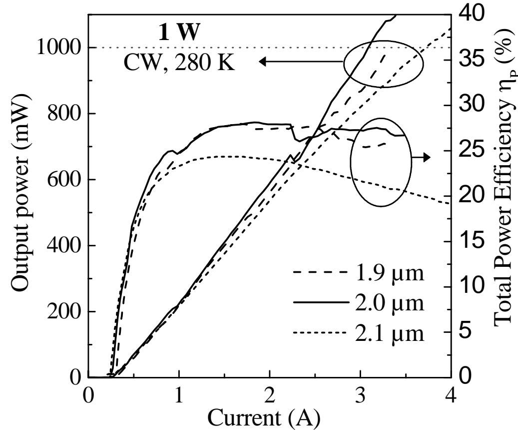


Figure 14. CW output power and power efficiency for three different 1000×150 μm^2 broad-area lasers, with emission wavelengths of 1.9, 2.0, and 2.1 μm

In a situation where heterobarrier leakage is the dominant loss mechanism [31,63], the use of AlGaAsSb barrier and separate confinement layers with a higher Al content of, *e.g.* 40%, compared to 25-30% for the above examples, and thus increased barrier height for electrons and holes, can be advantageous [64]. As has been shown in [31] for GaInAsSb/AlGaAsSb triple-QW diode lasers emitting at 2.2 μm, increasing the Al content from 30 to 40% results in a moderate increase in T_0 from 66 to 82 K for heatsink temperatures >320 K. However, there is a concomitant increase in operating voltage at threshold by 0.15 V, which directly reflects the increase in AlGaAsSb band gap energy. As a consequence, the power efficiency is even lower for 40% Al content barriers than for an Al-content of 30% up to operating temperatures of at least 340 K. On the other hand, using barrier and separate confinement with an Al content of only 20% was found to result in a reduced operating voltage and an improved maximum power efficiency of 30% at 2.2 μm and 280 K [31].

As can be seen from Figure 14, the effect of device self-heating becomes very important when dealing with high-power III-Sb diode lasers, limiting the maximum achievable CW output power to less than 1/5 of the maximum output in pulsed operation. Therefore, to gain a more detailed insight into the device parameters limiting CW performance as well as to get a tool at hand for further optimization of the laser structure as well as the mounting technology, an electro-optical-thermal model has been developed [29]. Input parameters for this model are temperature-dependent data for the threshold current I_{th} and differential quantum efficiency η_d, which are derived from temperature-dependent P-I characteristics measured in pulsed mode to avoid self-heating of the device.

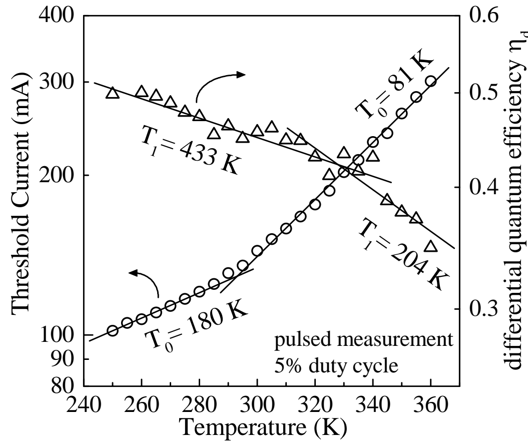


Figure 15. Threshold current and differential quantum efficiency *vs* temperature of a 1.98 μm emitting 1000×64 μm^2 ridge-waveguide laser, measured in pulsed operation

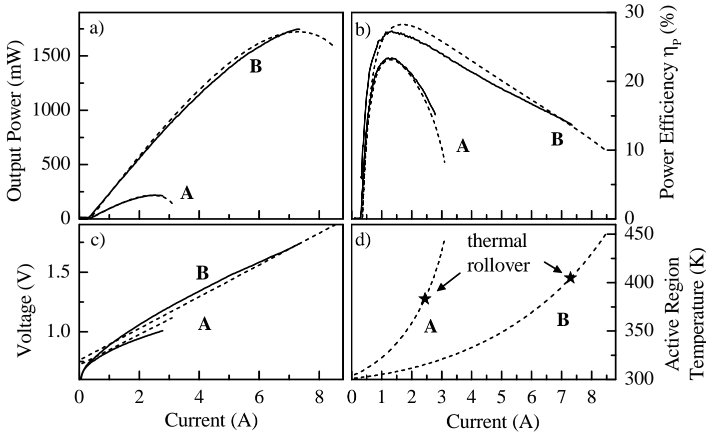


Figure 16. Measured (*solid lines*) and calculated (*dashed lines*) output power- (**a**), power efficiency- (**b**), voltage- (**c**) and (**d**) active region temperature *vs* current characteristics for a 2 μm emitting device mounted substrate-side down with uncoated facets (A) and for a device with HR/AR-coated facets mounted p-side down (B).

Typical experimental data are shown in Figure 15 for a 1000 μm × 64 μm ridge waveguide GaInSb/AlGaAsSb triple-QW diode laser, emitting at 1.98 μm for a heatsink temperature of 300 K. From the semi-logarithmic plot characteristic temperatures T_0 and T_1 can be derived for I_{th} and η_d, respectively, which differ for the low- and high-temperature range. The temperature-dependent thermal resistance Z_T is derived from the change in lasing wavelength at fixed heatsink temperature as a function of injection current, with the temperature tuning coefficient of the diode laser taken again from pulsed reference measurements. To account for different device geometries, Z_T is assumed to be inversely proportional to the cavity length

and to scale logarithmically with the ratio substrate thickness over ridge or stripe width. Finally standard text book equations are taken for the threshold gain condition and to calculate the emitted optical and dissipated thermal power.

As an example for the application of the above model, Figure 16 shows measured (solid lines) and calculated (dashed lines) P-I (a), η_{Power}-I (b), and voltage *vs* current characteristics (c) for a 1000 μm × 150 μm 2 μm GaInSb/AlGaAsSb triple-QW broad area lasers, mounted substrate-side down with uncoated facets (A) and for a device with HR/AR-coated facets mounted p-side down (B). There is excellent agreement between experimental and calculated data. All the input parameters were set to the laser parameters derived from independent measurements (see above). Only α_i and η_i were adjusted within the measurement uncertainties of about 10% for a best overall fit between experimental and modeled characteristics. The degree of agreement achieved demonstrates that this model is capable of accurately reproducing effects of device geometry and mounting on the high-power performance. From this model also the active region temperature is obtained as a function of the injection current (Figure 16d). It is interesting to note that thermal rollover occurs when the active region temperature exceeds the heatsink temperature by about 100 K, irrespective of device mounting [29].

6 Long-wavelength III-Sb-based Type-I QW Lasers

Regarding the long-wavelength limit of GaInAsSb/AlGaAsSb Type-I QW diode lasers, considerable progress has been made recently. Room-temperature pulsed operation at a lasing wavelength of 2.78 μm was reported in 1995 for a 4-QW laser with only slightly compressively strained GaInAsSb QWs [65], followed in 1999 by a report on room-temperature CW operation in the 2.3 to 2.7 μm wavelength range for double-QW devices with highly strained GaInAsSb QWs as the active region [22]. Recently, significant progress has been made regarding the long-wavelength performance of III-Sb based Type-I QW lasers, achieving now a room-temperature CW output of 1 W at 2.5 μm [62], 500 mW at 2.7 μm and 160 mW at 2.8 μm [37]. All these data apply to p-side down mounting with HR/AR-coated facets. The active region is composed of two highly strained GaInAsSb QWs with nominal In- and As-contents of 41 and 14% (2.5 μm devices) and 50 and 19% (2.7-2.8 μm devices), respectively, separated by a 200 nm wide $Al_{0.25}Ga_{0.75}As_{0.02}Sb_{0.98}$ barrier. The maximum power efficiency was found to drop from 12% at 2.5 μm [62] to 9% at 2.7 μm and only 2% at 2.8 μm [37]. The longest room-temperature lasing wavelength reported so far for III-Sb based Type-I QW diode lasers is 3.04 μm, employing again a strained double-QW active region to avoid excessive strain accumulation. The maximum CW output power at that wavelength, however, was just 3 mW at 15°C [15]. It is also interesting to note that the characteristic temperature of the threshold current, T_0, drops from 83 K at 2.5 μm [62] to 71 K (2.7 μm), 59 K (2.8 μm) [37] and 30 K (3.04 μm) [15], all data recorded in pulsed mode.

The decrease in laser performance for longer lasing wavelength has been attributed to increasing heterobarrier leakage of holes out of the increasingly shallower valence band potential well rather than to an increase in Auger

recombination rate [63,66]. As discussed already in Section 2, for $Ga_{1-x}In_xAs_ySb_{1-y}$ lattice matched to GaSb (*i.e.* y = 0.913·x, see Figure 1b), the valence band offset between this active layer and, *e.g.*, $Al_{0.3}Ga_{0.7}As_{0.02}Sb_{0.98}$ as the barrier material decreases with increasing In-content (*i.e.* with increasing wavelength), leading eventually to a negative valence band offset, *i.e.* a Type-II band alignment, for an In-content >35%. Without the addition of As, *i.e.* for $Ga_{1-x}In_xSb$ (see Figure 1a), the valence band offset increases with increasing In-content, but the high compressive strain will lead to a relaxation of long wavelength active QW layers of $Ga_{1-x}In_xSb$. A constant valence band offset of 130 meV between the $Ga_{1-x}In_xAs_ySb_{1-y}$ active layer and the barriers (*i.e.* independent of the In-content) is achieved if the condition y = 0.32 x is fulfilled. That means the As content should not exceed 0.32 of the In-content of the QW layer, in order to allow sufficient hole confinement in the valence band. The correlation between the strain ε_{zz} and the emission wavelength of these active layers satisfying the condition y = 0.32 x is shown by the dashed line in Figure 7, indicating that, as a rule of thumb, the strain-wavelength data of a laser structure should lie on or above that line. It is interesting to note that the data points of the long wavelength lasers with low T_0 values (2.8 µm and 3.04 µm [15,37]) fall below this line, while the shorter wavelength lasers with higher T_0 values (including the 2.7 µm laser in [37] with higher strain) are located above.

Another issue of concern for long-wavelength lasers is the increasing spread of the optical mode spreading in the waveguide of the laser structure. Given the standard laser structure shown in Figure 3, the confinement factor Γ_{QW} of the optical mode with the QWs is reduced by a factor of 0.76 and the confinement factor $\Gamma_{clad.}$ with the doped cladding layer is increased by a factor of 2.5, when the emission wavelength is increased from 2.0 to 3.0 µm. This leads to an increase of the threshold current for longer wavelength lasers because of the lower modal gain $\Gamma_{QW} \cdot g$ and the higher optical losses in the doped cladding layers (see Section 4), which results in turn in a reduced slope of the PI-characteristic. This issue cannot be solved by simple scaling of the optical waveguide, as this would result in a by far too low confinement factor Γ_{QW}.

7 Outlook

III-Sb based Type-I diode lasers covering the 2.0 to 2.7 µm range have already reached a considerable level of maturity, as evidenced by the low threshold currents, good power efficiencies and high output powers achieved by several groups (see Section 5). Thus current R&D focuses increasingly on the optimization of these lasers towards specific application.

Efficient coupling of the diode laser output into optical systems, including coupling into an optical fiber, requires a reduction of the typically large fast axis far-field beam divergence (see Figure 4b). As demonstrated very recently [67], this can be achieved by a careful design of the optical waveguide without sacrificing output power and power efficiency. Figure 17 shows the fast axis far-field intensity profile of such an improved laser structure (full curve) with a $1/e^2$ full width angle of 77°. For comparison, the far-field characteristic of a conventional laser structure (dashed

curve, see also Figure 4b) with a $1/e^2$ angle of 120° is also shown. Such a reduction in beam divergence without a drop in power and efficiency will significantly enhance the efficiency of, *e.g.* fiber coupled 2.X µm laser systems employing such improved III-Sb Type-I diode lasers.

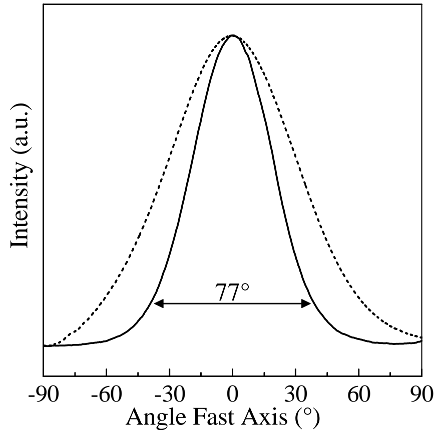


Figure 17. Fast axis far field intensity profiles of a conventional laser structure (*dashed line*) and a laser structure with an improved optical waveguide and thus reduced beam divergence of 77° (*solid line*)

For spectroscopic sensing of substances in solution, such as for non-invasive blood glucose monitoring [3], tuning ranges exceeding 100 nm are required. This can be achieved, *e.g.* by placing the diode laser into an external cavity setup, provided the gain spectrum of the active region is wide enough. For (AlGaIn)(AsSb) based Type-I diode lasers, a wide optical gain spectrum extending from 2.2 to 2.5 µm at room-temperature has already been demonstrated based on a wide QW design, and thus small energy separation between the first and second quantized QW states [68].

Another interesting approach is that of single-mode CW GaInAsSb/AlGaAsSb QW diode lasers emitting in the 2.3-2.4 µm range which can be operated up to temperatures of 130°C [46], suitable, *e.g.* for tunable diode laser absorption spectroscopy (TDLAS) of, *e.g.,* CO, CH_4, and NO_2. Such devices allow the realization of TDLAS trace gas sensors without the need of thermoelectric coolers, as these devices can be temperature stabilized and wavelength tuned entirely by heating with respect to ambient temperature, while the obtained output powers in the 10 mW range are still sufficient for sensitive detection.

Another challenge related to GaInAsSb/AlGaAsSb Type-I QW diode lasers is the exploration of their long wavelength limit. As discussed in the preceding section, the efficiency and temperature performance of these lasers drops significantly when approaching the 3 µm limit, mainly due to insufficient hole confinement. Improving on this requires the use of even higher In-content, and thus even more heavily strained GaInAsSb QWs in the active region. It remains to be seen to what extend further improvements in epitaxy and material quality can push

the limit for compressive strain relaxation further out, thus allowing a larger valence band offset for a given wavelength. Other approaches include the use of strain-compensating, *i.e.* tensile strained, AlGaAsSb barrier layers with highly compressively strained GaInAsSb QWs to avoid the onset of relaxation due to strain accumulation with increasing numbers of QWs [33]. A more radical approach is that of employing a metamorphic AlInSb buffer layer between the GaSb substrate and the laser structure, in order to generate an "artificial" substrate with a lattice constant larger than that of the GaSb [69]. This way, optically pumped GaInSb/AlGaInSb multiple-QW lasers with an In-content in the GaInSb QW layers of up to 76% and a compressive strain of only 1.1% have been realized, showing room-temperature lasing up to 3.26 µm and a maximum lasing wavelength of 3.45 µm at 170 K [69].

Acknowledgments

The authors would like to thank M. Walther, M. Mikulla, and F. Fuchs for valuable discussions and contributions as well as G. Weimann for continuous support and encouragement. Financial support by the German Ministry for Education and Research (BMBF) is gratefully acknowledged.

References

1. L.S. Rothman, C.P. Rinsland, A. Goldman, S.T. Massie, D.P. Edwards, J.M. Flaud, A. Perrin, C. Camy-Peyret, V. Dana, J.Y. Mandin, J. Schroeder, A. McCann, R.R. Gamache, R.B. Wattson, K. Yoshino, K.V. Chance, K.W. Jucks, L.R. Brown, V. Nemtchinov, P. Varanasi: The HITRAN molecular spectroscopic database and HAWKS (HITRAN atmospheric workstation): 1996 edition. J. Quant. Spectrosc. Radiat. Transfer **60**, 665 (1998)
2. B. Jean, T. Bende: Mid-IR Laser Applications in Medicine. In: I.T. Sorokina, K.L. Vodopyanov (ed) Solid-State Mid-Infrared Laser Sources, Springer-Verlag, Berlin, 2003, pp. 511-544 (Topics in Applied Physics no. 89)
3. J.T. Olesberg: Noninvasive blood glucose monitoring in the 2.0-2.5 µm spectral range. In: 2001 IEEE/LEOS Conf. Proc. Vol. 2, p. 529
4. see for example http://www.lisalaser.com
5. M. Mond, D. Albrecht, E. Heumann, G. Huber, S. Kück, V.I. Levchenko, V.N. Yakimovich, V.G. Shcherbitsky, V.E. Kisel, N.V. Kuleshov, M. Rattunde, J. Schmitz, R. Kiefer, J. Wagner : 1.9-µm and 2.0-µm laser diode pumping of Cr^{2+}:ZnSe and Cr^{2+}:CdMnTe. Optics Lett. **27**, 1034 (2002)
6. C.D. Nabors, J. Ochoa, T.Y. Fan, A. Sanchez, H.K. Choi, G.W. Turner: Ho:YAG Laser Pumped By 1.9-µm Diode Lasers, IEEE J. Quantum Electron. **31**, 1603 (1995)
7. I. Melngailis: Maser action in InAs Diodes. Appl. Phys. Lett. **2**, 176 (1963)
8. C. Caneau, A.K. Srivastava, A.G. Dentai, J.L. Zyskind, M.A. Pollack: Room temperature GaInAsSb/AlGaAsSb DH injection lasers at 2.2 µm. Electron. Lett. **21**, 815 (1985)
9. A.N. Baranov, C. Fouillant, P. Grunberg, J.L. Lazzari, S. Gaillard, A. Joullié: High temperature operation of GaInAsSb/AlGaAsSb double heterostructure lasers emitting near 2.1 µm. Appl. Phys. Lett. **65**, 616 (1994)

10. J.P. van der Ziel, T.H. Chiu, W.T. Tsang: Optically pumped laser oscillation at 3.83 μm from $InAs_{1-x}Sb_x$ grown by molecular beam epitaxy on GaSb. Appl. Phys. Lett. **47**, 1139 (1985)
11. H.K. Choi, G.W. Turner, Z.L. Liau: 3.9-μm InAsSb/AlAsSb double-heterostructure diode lasers with high output power and improved temperature characteristics. Appl. Phys. Lett. **65**, 2251 (1994)
12. A.N. Baranov, A.N. Imenkov, V.V. Sherstnev, Y.P. Yakovlev: 2.7-3.9 μm InAsSb(P)/InAsSbP low threshold diode lasers. Appl. Phys. Lett. **64**, 2480 (1994)
13. H.K. Choi, S.J. Eglash: High-power multiple-quantum-well GaInAsSb/ AlGaAsSb diode lasers emitting at 2.1 μm with low threshold current density. Appl. Phys. Lett. **61**, 1154 (1992)
14. G.W. Turner and H.K. Choi: Antimonite-based mid-infrared quantum well diode lasers. In: M.O. Manasreh (ed) Optoelectronic Properties of Semiconductors and Superlattices. Gordon and Beach, Amsterdam, 1997, pp. 369-431
15. C. Lin, M. Grau, O. Dier, M.C. Amann: Low threshold room-temperature continuous-wave operation of 2.24-3.04 μm GaInAsSb/AlGaAsAb quantum well lasers. Appl. Phys. Lett. **84**, 5088 (2004)
16. J.R. Meyer, C.A. Hoffman, F.J. Bartoli: Type-II quantum-well lasers for the mid-wavelength infrared. Appl. Phys. Lett. **67**, 757 (1995)
17. J.R. Meyer, C.L. Felix, W.W. Bewley, I. Vurgaftman, E.H. Aifer, L.J. Olafsen, J.R. Lindle, C.A. Hoffman, M.J. Yang, B.R. Bennett B.V. Shanabrook, H. Lee, C.H. Lin, S.S. Pei, R.H. Miles: Auger coefficients in Type-II $InAs/Ga_{1-x}In_xSb$ quantum wells. Appl. Phys. Lett. **74**, 2857 (1998)
18. I. Vurgaftman, J.R. Meyer, L.R. Ram-Mohan: Band parameters for III-V compound semiconductors and their alloys. Appl. Phys. Rev. **89**, 5815 (2001)
19. C. Alibert, A. Joullié, A.M. Joullié : Modulation-spectroscopy study of the $Ga_{1-x}Al_xSb$ band structure. Phys. Rev B **27**, 4946 (1983)
20. O. Madelung, W. Osten, U. Rössler: Intrinsic Properties of Group IV Elements and III-V, II-VI and I-VII Compounds. In: O. Madelung (ed.) Landolt-Börnstein New Series, Springer-Verlag, Berlin,1987
21. K. Shim, H. Rabitz, P. Dutta: Band gap and lattice constant of $Ga_xIn_{1-x}As_ySb_{1-y}$. J. Appl. Phys. **88**, 7157 (2000)
22. D.Z. Garbuzov, H. Lee, V. Khalfin, R. Martinelli, J. Connolly, G. Belenky: 2.3-2.7 μm Room Temperature CW Operation of InGaAsSb-AlGaAsSb Broad Waveguide SCH-QW Diode Lasers. IEEE Photon. Technol. Lett. **11**, 794 (1999)
23. C. Mermelstein, S. Simanowski, M. Mayer, R. Kiefer, J. Schmitz, M. Walther, J. Wagner: Room-temperature low-threshold low-loss continuous-wave operation of 2.26 μm GaInAsSb / AlGaAsSb quantum-well laser diodes. Appl. Phys. Lett **77**, 1581 (2000)
24. M. Grau, C. Lin, M.C. Amann: Room-Temperature 2.81-μm Continuous-Wave Operation of GaInAsSb-AlGaAsSb Laser. IEEE Photon. Technol. Lett. **16**, 383 (2004)
25. T. Chong, C. Fonstad: Theoretical gain of strained-layer semiconductor lasers in the high-strain regime. IEEE. J. Quantum Electron. **25**, 171 (1989)
26. S. Chuang: Efficient band-structure calculations of strained quantum wells. Phys. Rev. B **43**, 9649 (1991)
27. S. Corzine, R. Yan, L. Coldren: Optical gain in III-V bulk and quantum well semiconductors. In: P.S. Zory (ed.) Quantum well lasers. Academic Press, San Diego, 1986, pp.17-93
28. A. Ghiti, E. O'Reilly: Antimoni-based strained-layer 2-2.5 μm quantum well lasers. IEEE J. Quantum Electron. **27**, 1347 (1991)

29. M. Rattunde, C. Mermelstein, J. Schmitz, R. Kiefer, W. Pletschen, M. Walther, J. Wagner: Compehensive modeling of the electro-optical-thermal behavior of (AlGaIn)(AsSb)-based 2.0 µm diode lasers. Appl. Phys. Lett. **80**, 4085 (2002)
30. A. Salhi, Y. Rouillard, A. Pérona, P. Grech, M. Garcia, C. Sitori: Low-threshold GaInAsSb/AlGaAsSb quantum well laser diodes emitting near 2.3 µm. Semicond. Sci. Technol. **19**, 260 (2004)
31. M. Rattunde, C. Mermelstein, S. Simanowski, J. Schmitz, R. Kiefer, N. Herres, F. Fuchs, M. Walther, J. Wagner: Temperature sensitivity of high power GaSb based 2 µm diode lasers. In: M. Ilegems, G. Weimann, J. Wagner (eds.) Procceedings of the 29th Int. Symp. on Compound Semiconductors, IOP Publishing LTD, Bristol (UK) 2003, p.347 (Inst. Phys. Conf. Ser. no 174)
32. S. Simanowski, N. Herres, C. Mermelstein, R. Kiefer, J. Schmitz, M. Walther, J. Wagner, G. Weimann: Strain adjustment in (GaIn)(AsSb)/(AlGa)(AsSb) QWs for 2.3-2.7 µm laser structures. J. Cryst. Growth **209**, 15 (2000)
33. W. Li, J.B. Héroux, H. Shao, W.I. Wang : Strain-compensated InGaAsSb/ AlGaAsSb mid-infrared quantum-well lasers. Appl. Phys. Lett. **84**, 2016 (2004)
34. D.Z. Garbuzov, R.U. Martinelli, H. Lee, P.K. York, R.J. Menna, J.C. Connolly, S.Y. Narayan: Ultralow-loss broadened-waveguide high-power 2 µm AlGaAsSb/InGaAsSb/GaSb separate-confinement quantum-well lasers. Appl. Phys. Lett. 69, 2006 (1996)
35. M. Rattunde, J. Schmitz, R. Kiefer, J. Wagner: Comprehensive analysis of the internal losses in 2.0 µm (AlGaIn)(AsSb) quantum-well diode lasers. Appl. Phys. Lett. **84**, 4750 (2004)
36. A. Baraldi, C. Grezzi, R. Magnanimi, A. Parigini, L. Tarricone, A. Cosacchi, S. Franchi, V. Avanzino, P. Allegri: Preparation of GaSb by molecular beam epitaxy and electrical and photoluminescence characterisation. Mat. Science and Engineering B **28**, 174 (1994)
37. J.G. Kim, L. Shterengas, R.U. Martinelli, G.L. Belenky: High-power room-temperature continuous wave operation of 2.7 and 2.8 µm In(Al)GaAsSb/GaSb diode lasers. Appl. Phys. Lett. **83**, 1926 (2003)
38. C. Mermelstein, M. Rattunde, J. Schmitz, S. Simanowski, R. Kiefer, M. Walther, J. Wagner: Sb-Based Mid-Infrared Diode Lasers. Mat. Res. Soc. Symp. Proc. 692, 365 (2002)
39. C.A. Wang, H.K. Choi: GaInAsSb/AlGaAsSb multiple-quantum-well diode lasers grown by organometallic vapor phase epitaxy. Appl. Phys. Lett. 70, 802 (1997)
40. S. Simanowski, C. Mermelstein, M. Walther, N. Herres, R. Kiefer, M. Rattunde, J. Schmitz, J. Wagner, G. Weimann: Growth and layer structure optimization of 2.26 µm (AlGaIn)(AsSb) diode lasers for room temperature operation. J. Cryst. Growth **227-228**, 595 (2001)
41. A. Joullié, P. Christol, A.N. Baranov, A. Vicet: Mid-Infrared 2-5 µm Heterojunction Laser Diodes. In: I.T. Sorokina, K.L. Vodopyanov (ed.) Solid-State Mid-Infrared Laser Sources, Springer-Verlag, Berlin, 2003, pp. 1-59 (Topics in Applied Physics no. 89)
42. K. Onabe: Unstable region in quaternary $In_{1-x}Ga_xAs_{1-y}Sb_y$ calculated using strictly regular solution approximation. Jpn. J. Appl. Phys. **21**, 964 (1982)
43. S.L. Chuang: Physics of optoelectronic devices. John Wiley & Sons, Inc, New York, 1995
44. B. Bennett, W. Moore, M. Yang, B. Shanabrook: Transport properties of Be- and Si-doped AlSb. J. Appl. Phys. **87**, 7876 (2000)
45. H.K. Choi, S.J. Eglash, M.K. Connors: Single-frequency GaInAsSb/AlGaAsSb quantum-well ridge-waveguide lasers emitting at 2.1 µm. Appl. Phys. Lett. **63**, 3271 (1993)
46. D.A. Yarekha, G. Gastre, A. Perona, Y. Rouillard, F. Genty, E.M. Skouri, G. Boissier, P. Grech, A. Joullié, C. Alibert, A.N. Baranov: High temperature GaInSbAs/GaAlSbAs

quantum well single mode continuous wave lasers emitting near 2.3 μm. IEEE Electron. Lett. **36**, 537 (2000).
47. M. Hümmer, K. Rößner, A. Benkert, A. Forchel: GaInAsSb-AlGaAsSb Distributed Feedback Lasers Emitting Near 2.4 μm. IEEE Photon Technol. Lett. 16, 380 (2004)
48. A.N. Baranov, Y. Rouillard, G. Boissier, P. Grech, S. Gaillard, C. Alibert: Sb-based monolithic vertical cavity surface emitting laser operating near 2.2 μm at room temperature. Electron. Lett. **34**, 281 (1998)
49. L. Cerutti, A. Garnache, F. Genty, A. Ouvrad, C. Alibert: Low threshold, room-temperature laser diode pumped Sb-based VECSEL emitting around 2.1 μm. Electron. Lett. **39**, 290 (2003)
50. R.J. Menna, D.Z. Garbuzov, H. Lee, R.U. Martinelli, S.Y. Narayan, J.C. Connolly: High-power broadened-waveguide InGaAsSb/AlGaAsSb quantum-well diode lasers emitting at 2 μm. Proc. SPIE 3284, 238 (1998)
51. H.K. Choi, J.N. Walpole, G.W. Turner M.K. Connors L.J. Missaggia, M.J. Manfra: GaInAsSb–AlGaAsSb Tapered Lasers Emitting at 2.05 μm with 0.6-W Diffraction-Limited Power. IEEE Photon. Technol Lett. **10**, 938 (1998)
52. J.N. Walpole, H.K. Choi, L.J. Missagia Z.L. Zia, M.K. Connors, G.W. Turner, M.J. Manfra, C.C. Cook: High-Power High-Brightness GaInAsSb-AlGaAsSb Tapered Laser Arrays with Anamorphic Collimating Lenses Emitting at 2.05 μm. IEEE Photon. Technol. Lett. **11**, 1223 (1999)
53. L.A. Coldren, S.W. Corzine: Diode Lasers and Photonic Integrated Circuits. John Wiley & Sons, New York, 1995
54. G. Turner, H. Choi, M. Manfra: Ultralow-threshold (50 A/cm^2) strained single-quantum-well GaInAsSb/AlGaAsSb lasers emitting at 2.05 μm. Appl. Phys. Lett. **72**, 876 (1998)
55. D. Garbuzov, M. Maiorov, H. Lee, V. Khalfin, R. Martinelli, J. Connolly: Temperature dependence of continuous wave threshold current for 2.3-2.6 μm InGaAsSb/AlGaAsSb separate confinement heterostructure quantum well semiconductor diode lasers. Appl. Phys. Lett. **74**, 2990 (1999)
56. T. Newell, L. Lester, X. Wu, Y. Zhang: Gain and threshold current density characteristics of 2 micron GaInAsSb/AlGaAsSb MQW lasers with increased valence band offset. Proc. SPIE **3284**, 258 (1998)
57. P. Yu, M. Cardona: Fundamentals of Semiconductors. Springer-Verlag, Berlin (1996)
58. J. Piprek: Semiconductor Optoelectronic Devices. Academic Press, San Diego (2003)
59. I. Joindot, J. Beylat: Intervalence Band Absorption coefficient measurements in bulk layer, strained and unstrained multiquantum well 1.55 μm semiconductor lasers. Electron. Lett. **29**, 604 (1993)
60. G. Fuchs, J. Hörer, A. Hangleiter, V. Härle, F. Scholz, R. Glen, L. Goldstein: Intervalence band absorption in strained and unstrained InGaAs multiple quantum well structures. Appl. Phys. Lett. **60**, 231 (1992)
61. J. Piprek, P. Abraham, J.E. Bowers: Cavity Length Effects on Internal Loss and Quantum Efficiency of Multiquantum-Well Lasers. IEEE J. Sel. Top. Quantum Electron. **5**, 643 (1999)
62. J.G. Kim, L. Shterengas, R.U. Martinelli, G.L. Belenky, D.Z. Garbuzov, W.K. Chan : Room-temperature 2.5 μm InGaAsSb/AlGaAsSb diode lasers emitting 1 W continuous waves. Appl. Phys. Lett. **81**, 3146 (2002)
63. M. Rattunde, C. Mermelstein, S. Simanowski, J. Schmitz, R. Kiefer, N. Herres, F. Fuchs, M. Walther, J. Wagner: Temperature dependence of threshold current for 1.8 to 2.3 μm (AlGaIn)(AsSb)-based QW diode lasers. Proceedings of the 27th Int. Symp. on Compound Semiconductors, Institute of Electrical and Electronics Engineers, Inc. 2001, p.437

64. T. Newell, X. Wu, A.L. Gray, S. Dorato, H. Lee, L.F. Lester: The Effect of Increased Valence Band Offset on the Operation of 2 μm GaInAsSb-AlGaAsSb Lasers. IEEE Photon. Technol. Lett. **11**, 30 (1999)
65. H. Lee, P.K. York, R.J. Menna, R.U. Martinelli, D.Z. Garbuzov, S.Y. Narayan, J.C. Connolly: Room-temperature 2.78 μm AlGaAsSb/InGaAsSb quantum-well lasers. Appl. Phys. Lett. **66**, 1942 (1995)
66. L. Shterengas, G.L. Belenky, J.G. Kim, R.U. Martinelli: Design of high-power room-temperature continuous-wave GaSb-based Type-I quantum-well lasers with $\lambda > 2.5$ μm. Semicond. Sci. Technol. **19**, 655 (2004).
67. M. Rattunde, J. Schmitz, G. Kaufel, J. Wagner, to be published
68. D.V. Donetsky, D. Westerfeld, G.L. Belenky, R.U. Martinelli, D.Z. Garbuzov, J.C. Connolly: Extraordinarily wide optical gain spectrum in 2.2-2.5 μm In(Al)GaAsSb/GaSb quantum-well ridge-waveguide lasers. J. Appl. Phys. **90**, 4281 (2001)
69. E.A. Pease, L.R. Dawson, L.G. Vaughn, P. Rotella, L.F. Lester: 2.5-3.5 μm optically pumped GaInSb/AlGaInSb multiple quantum well lasers grown on AlInSb metamorphic buffer layers. J. Appl. Phys. **93**, 3177 (2003).

VCSELs Emitting in the 2-3 µm Wavelength Range

F. Genty, A. Garnache and L. Cerutti
CEM2, UMR n°5507, University of Montpellier 2
Montpellier, France

1 Introduction

The aim of this chapter is not to offer an exhaustive presentation of Vertical Cavity Surface Emitting Lasers (VCSELs) but to introduce the existing structures and their applications in the particular 2 -3 µm wavelength range. VCSELs have indeed been the subject of many publications and several books [see for example 1-3] where the interested reader will be able to find very detailed descriptions concerning design and/or operating conditions.

The VCSEL has become one of the most important devices for optical interconnects or high speed data transfers. Historically, the very first surface-emitting laser was fabricated during the first years of the semiconductor laser's history but the structure of this device was far from the vertical-cavity lasers that we know today [4]. In this InSb injection device, the optical cavity was delimited by the two surfaces of the wafer. Multimode laser operation at 5.2 µm was reported at 10K and for such operation a magnetic field was required. The Vertical Cavity Surface Emitting Lasers that we know appeared much later. Suggested [5] then demonstrated for the first time at the end of the 1970s [6], the first GaAs-based VCSELs operating in continuous wave (CW) mode at room temperature (RT) were demonstrated at the end of the 1980s [7] and InP-based devices at the beginning of the 1990s [8]. These devices were then extensively developed on GaAs and InP substrates with emitting wavelengths varying between 0.8 and 1.6 µm primarily for data communications and high speed optical interconnects. Indeed, very interesting properties, such a high potential modulation rate combined with a low threshold and a low divergence circular output beam, made VCSELs ideal laser sources for these kind of applications.

For a number of years, mid-IR VCSEL technology has attracted a very large scientific interest. VCSELs from 2 µm [9] up to 8 µm [10, 11] have been developed, operating both by electrical pumping and optical pumping. In this chapter, we will focus on the 2 - 3 µm range. This wavelength range is of particular interest for gas detection and spectroscopy. In fact, in this wavelength region, strong absorption lines of atmospheric pollutants such as CH_4, CO, NO_2 and H_2CO exist while the CO_2 and H_2O absorptions are very low. Several published works have shown that VCSEL devices offer a variety of advantages when compared to edge-emitting diode lasers for absorption spectroscopy [12, 13]. Their single mode operation, the broad continuous tuning range (without mode-hops) and the high repetition rate of VCSELs allow rather good time resolution measurements even in very difficult experimental conditions. Moreover, VCSELs have additional advantages over edge-emitting diodes such as low cost manufacturing and easy testing capabilities

(without the need of facet cleavage). But gas detection is not the unique application that VCSELs can find in the 2 -3 μm range. Such devices can also be used in domains as diverse as laser surgery, fiber amplifier pumping, free space communications or military countermeasures for example [14 - 18].

2 A Short Description of Vertical-Cavity Lasers

The vertical-cavity structure was conceived in 1977 and was implemented in 1979 with the realization of the first operating device produced by liquid phase epitaxy [6]. In this structure, the mirrors were simple metal deposits : Au for the n-doped side and Au-Zn for the p-doped side (which was also used as an electrode). A bulk layer of InGaAsP produced laser emission at 1.18 μm at 77K in pulsed regime with a threshold current density of 11 kA/cm^2. Then, many improvements were realised in the following years until the demonstration, by the same team, of the first VCSEL emitting at 1.3 μm in continuous wave (CW) mode operation near 14°C in 1993 [8].

2.1 Conventional Microcavity VCSELs

To obtain laser emission from the surface, the device must contain a vertical cavity where the stimulated radiation no longer propagates in a plane parallel to the epitaxial layers (as in edge-emitting sources) but in a direction perpendicular to the growth plane. The more commonly developed VCSEL structures are therefore microcavities (cavity length ~ emission wavelength) where the optical cavity is defined by two high reflectivity mirrors epitaxially grown to sandwich the gain region (Figure 1). These high reflectivity mirrors are typically semiconductor Distributed Bragg Reflectors (DBR) lattice-matched to the substrate (GaAs/AlGaAs on GaAs, AlGaInAs/AlInAs, InGaAsP/InP, AlGaAsSb/AlAsSb on InP, AlGaAsSb/GaSb on GaSb, ...) but dielectric stacks, metamorphic materials, oxidized mirrors and wafer-bonded structures have also been used successfully [19-28].

With such a short cavity length, the optical gain per round trip is low. As a consequence, radiation losses in the device must be drastically reduced to reach lasing threshold. Therefore, to compensate the weak gain per one round trip, it is necessary to incorporate very high reflectivity mirrors in the structure (> 99%). So, the fundamental difference of VCSELs, when compared to other laser structures, originates from its low volume active layer. The VCSEL is therefore a structure exhibiting a potentially very low laser threshold. Threshold currents on the order of a few tens of μA for small emission diameter, threshold voltage as low as 1.3 V and wall-plug efficiency higher than 50% have been demonstrated [29 – 31]. Moreover, a very low active volume combined with a very high density of photons provides a high modulation rate. Fiber transmission rates provided by a VCSEL as high as 6 Gb/s have already been observed [32, 33]. But this low active volume also induces a low output power, lower than the output power of an edge-emitting device. Nevertheless, more than 1 Watt with only one VCSEL structure and several Watts with a bi-dimensional matrix have already been demonstrated in CW mode of operation. A record power of 5 W was measured from a bi-dimensional matrix in pulsed operation [34–37].

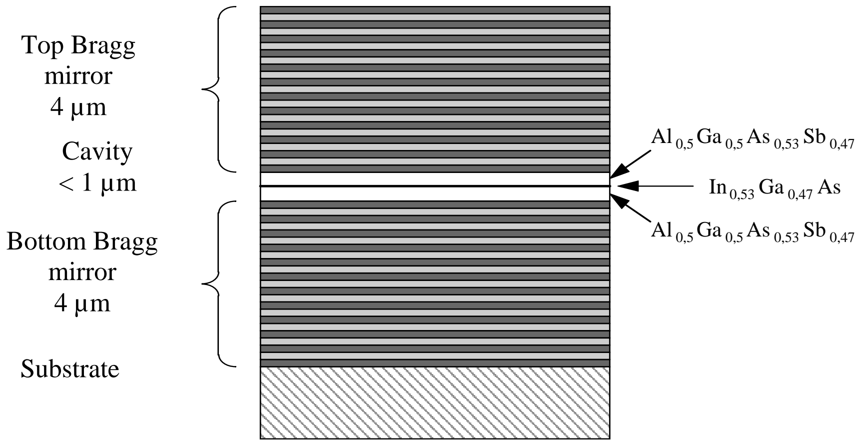


Figure 1. Typical structure of a VCSEL microcavity

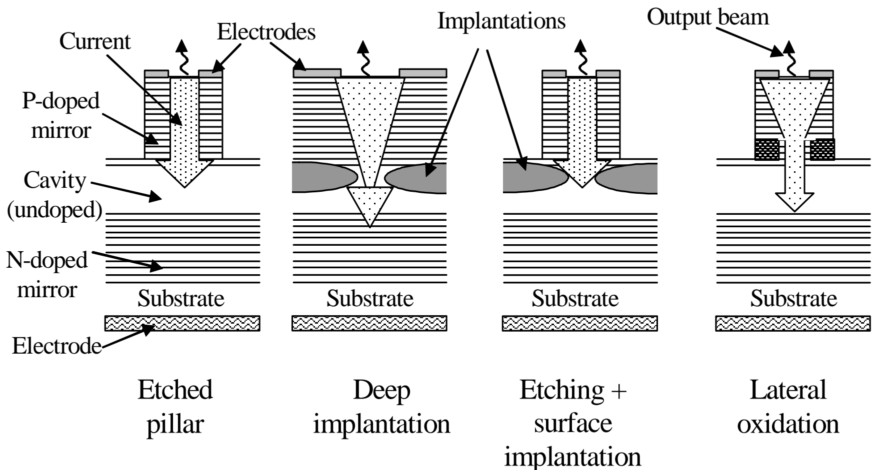


Figure 2. Several VCSEL structures with different processing. These designs were essentially developed for electrically-pumped devices.

Another distinct advantage of a VCSEL is its excellent output beam. Indeed, single-mode operation of a VCSEL can be obtained in the longitudinal as well as the transverse direction. While longitudinal single-mode operation is inherent to the short VCSEL optical cavity, transverse mode behaviour depends strongly on device design and the technological operations performed. Several basic VCSEL structures are presented in Figure 2. Generally, most of them can be divided into etched devices, implanted devices, oxidized devices or some appropriate combination.

These technological treatments have essentially been developed for electrically-pumped devices to provide both current and optical mode confinement.

However, either electrical or optical pumping can be used in a VCSEL. Figure 3 illustrates the band diagram of an active layer under optical (a) and electrical (b) pumping. With electrical pumping, current leakage can be generated which lowers the laser performance. When optical pumping is used, the size of the output beam is determined by the size of the pump beam diameter. Moreover, an optical source of power allows one to avoid doping, etching and deposition of electrical contacts during VCSEL fabrication, so that sophisticated technological processes are no longer necessary in the fabrication of high performance devices.

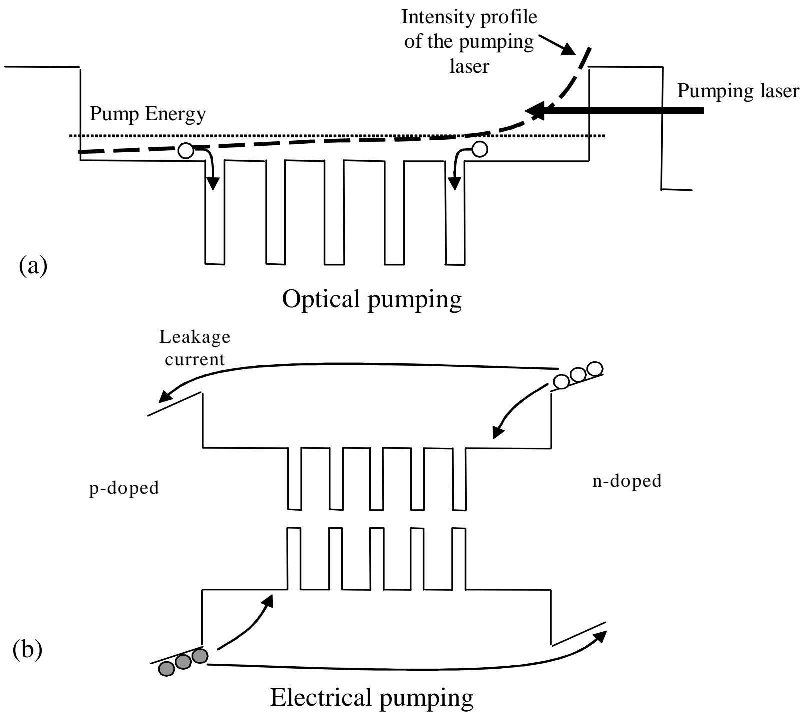


Figure 3. Schematic band diagram of Type-I MQW active layer under optical (**a**) and electrical (**b**) pumping

The performance of electrically pumped VCSELs is limited by optical losses and resistive self-heating. Conversely, in an optically-pumped system, the carriers are directly generated in the active layer, limiting self-heating. The undoped Bragg mirrors also reduce optical losses and thermal resistance. So, optical pumping seems to be an advantageous way to fabricate low threshold devices with high operating temperature without post-growth processing. Nevertheless, one of the Bragg mirrors must be transparent to the pumping wavelength to allow a direct excitation in the active layer. Therefore, a top dielectric mirror, deposited post-growth, is often used. In the mid-infrared wavelength range the required alternating layers can be made of

for example; SiO_2/Si, SiO_2/TiO_2, YF_3/ZnS [38-40]. Polarization and incident angle of the pump beam must also be optimized to minimize the reflectivity at the surface of the VCSEL.

2.2 External-Cavity VCSELs (VECSELs)

2.2.1 Structure Description

The VCSEL's microcavity output beam properties depend on device design. For small device diameters (<10 μm), VCSELs can operate in a single transverse mode but with a low output power (<10 mW) [41]. For larger device diameters, higher ouput power can be obtained, however the laser operates in multiple transverse mode [34, 36, 42]. VCSEL emission generally has a linear polarization. When it is grown on a (100) substrate, the polarization is oriented in (110) or an equivalent direction. But polarization switching and multimode operation can occur due to spatial hole burning or temperature/excitation variations.

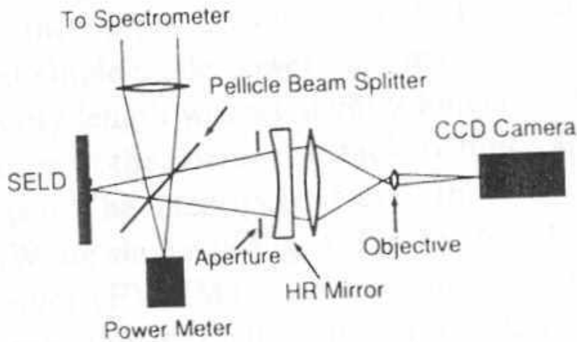


Figure 4. Schematic diagram of the SELD experimental setup reprinted with permission from M. A. Hadley, G. C. Wilson, K. Y. Lau, and J. S. Smith, Applied Physics Letters, 63, 1607 (1993). Copyright 1993, American Institute of Physics. [44]. The HR mirror has a transmission of 0.1% and radius of curvature of 10 cm.

Much work has been undertaken to control the polarization and to obtain high power emission while still maintaining single-transverse mode operation. Another objective of research was also to increase the tuning range of VCSELs. In 1992, Kenichi Iga's team [43] developed the first microcavity device with an extended cavity to improve the tuning range. This structure was composed of a bottom evaporated Si/SiO_2 Bragg mirror, an InGaAsP bulk active layer and an external plane Si/SiO_2 reflector (similar to the bottom one). The external mirror was set in contact with the chip surface, then the cavity length was slightly changed mechanically by adjusting the spacing between the chip and the external mirror. CW laser operation at 77K with a continuous tuning range of 40 Å and a threshold

current of 7.5 mA for a 10 μm-diameter circular emitting surface was achieved, using electrical pumping. This innovative structure demonstrated very interesting emission properties and was the first step towards the Vertical-External-Cavity-Surface-Emitting-Laser (VECSEL). Therefore, the use of a vertical external cavity appeared to be an advantageous way to control the VCSEL's output beam properties and stabilize the single mode operation and polarization state. The VECSEL, (also called the semiconductor disk laser) was then demonstrated as the most adapted system which combines a high tuning range, a high output power and a good beam quality (TEM_{00}). In this device, there is no upper Bragg mirror as it is replaced by an external concave mirror. So, the laser oscillation occurs in an optical cavity formed by the semiconductor active layer (spacers and multi-quantum-well region) and the air gap between the surface of the semiconductor and the external mirror. These VECSELs combine the advantages of edge-emitting devices (high power) and VCSEL lasers (single transverse mode operation, low threshold). This kind of structure can be either optically or electrically-pumped.

Many systems have been developed with different window layers (the window is the layer or the stack of layers at the air/semiconductor interface of the VECSEL). For example, the SELD (Surface Emitting Laser Diode) system can be considered [44] (see Figure 4). The SELD is of standard microcavity construction except that the top window is a semiconductor Bragg mirror with reduced reflectivity (~ 85%) and the laser is coupled to an external cavity. The external cavity is defined by a HR (High Reflectivity) concave mirror (R ~ 99.9%). In this configuration, the transverse modes are controlled both by using an intracavity aperture and by adjusting the cavity length which changes the mode diameter. The aperture and cavity length combinations increase the losses of the higher-order modes and therefore allow single-transverse mode laser operation. Electrically-pumped TEM_{00} laser operation with ouput powers as high as 100 mW pulsed and 2.4 mW CW have been reported using this configuration [45]. A fiber Bragg grating (FBG) can also be used as an external selective wavelength reflector. With such an external cavity, stable single-mode laser operation at room temperature was reported for modulation frequencies ranging from 500 Mb/s up to 4 Gb/s [46].

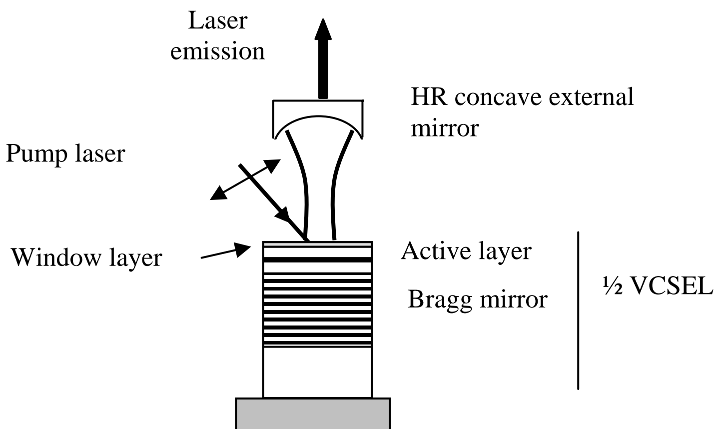


Figure 5. Schematic of an optically-pumped VECSEL configuration

The VECSEL's structure and properties were then improved by the use of a modified top window (an anti reflection stacks for example) and an optical pump source (Figure 5). With these modifications, Sandusky *et al.* [47] reported the first CW optically-pumped VECSEL operating at room temperature and, a few months later, Kuznetsov *et al.* [48,49] demonstrated CW single transverse mode (TEM_{00}) VECSEL operation at room temperature with an output power of 0.52 W near 1 µm. Up to now, the highest output power was reported by Lutgen *et al.* [50] who obtained 6 W in TEM_{00} mode operation at room temperature (8 W at 273K) near 1 µm. A single-frequency VECSEL emitting at 1 µm has been successfully used as the laser source for Intra Cavity Laser Absorption Spectroscopy (ICLAS) [51]. High performance VECSELs emitting near 1.5 µm have also been developed [52,53].

2.2.2 Geometrical Stability of a Concave/plane Laser Cavity

To allow a better comprehension of the VECSEL's operation, the stability conditions of an optical laser cavity formed by a concave mirror and a plane mirror are described briefly below.

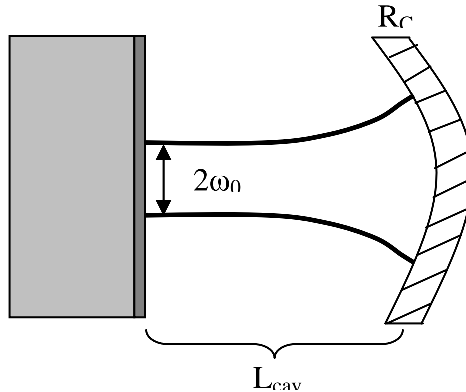


Figure 6. Scheme of operation of a VECSEL with a plane/concave laser cavity

To operate in a stable laser mode, the wave front of the gaussian beam in the cavity must be superimposed on itself after one round trip. For this, it is necessary that the radii of curvature of the gaussian beam $R(L_{cav})$ and the cavity mirrors coincide well. Thus, in a microcavity VCSEL where the cavity is plane/plane, the device does not operate on a stable mode. In a concave/plane cavity (Figure 6), the waist (ω_0) of the beam is located at the surface of the plane mirror. The radius of curvature of the beam must be the same as R_c, the radius of curvature of the concave mirror

$$R(L_{cav})=R_C \tag{1}$$

Therefore, a plane/concave resonator is stable only when the cavity length is less than or equal to the radius of curvature of the concave mirror. So, the waist at the surface of the plane mirror, with λ the wavelength of operation, is given by :

$$\omega_0 = \sqrt{\frac{\lambda}{\pi}\sqrt{L_{cav}\left(R_C - L_{cav}\right)}} \quad (2)$$

For a given radius of curvature (R_c) of the concave external mirror, the waist depends only on the cavity length and its maximum value occurs when the cavity length is half the radius of curvature.

Figure 7 shows the evolution of a plane/concave cavity waist vs the cavity length. This stability zone defines the range in which the fundamental gaussian transverse mode (TEM_{00}) will oscillate in the cavity.

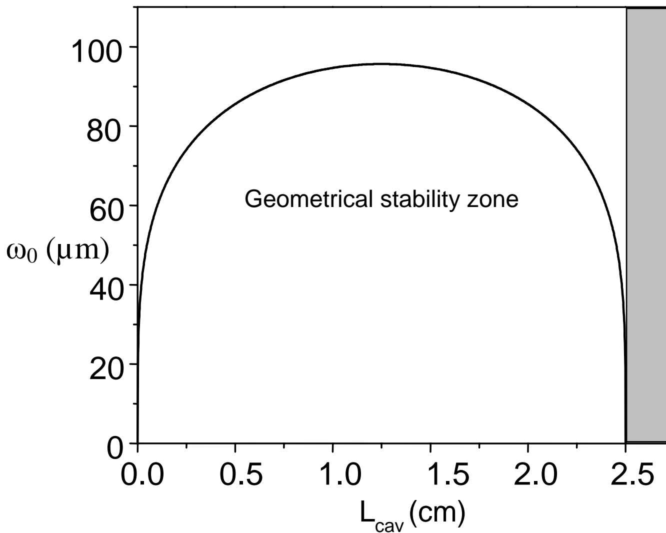


Figure 7. Evolution of the waist vs cavity length for a plan/concave cavity Rc = 2.5 cm and $\lambda = 2.3$ µm

But the surface of the excited zone produced by the pumping system (electrical or optical) is also an important parameter. This surface delimits the excited zone of the active layer and thus the surface of emission in a particular mode. For laser operation, the laser waist must be less than the radius of the pumping zone in order to provide gain in all the transverse dimensions of the mode. The effective stability zone of the same cavity previously described and for a 60 µm-diameter pumping spot is drawn in Figure 8. In this particular case, the length of the effective stability zone is 70 µm.

When a ½ VCSEL structure is coupled with an external plane/concave optical cavity, operation in the circular single transverse mode TEM_{00} is enforced. The transverse mode distribution in a VECSEL adapted to emission at $\lambda = 2.3$ µm is shown in Figure 9. Conversely, as it occurs in a microcavity, the waist of the fundamental transverse mode is lower than the waist of higher order transverse modes. So, with a 60 µm-diameter pump spot, this device will theoretically operate

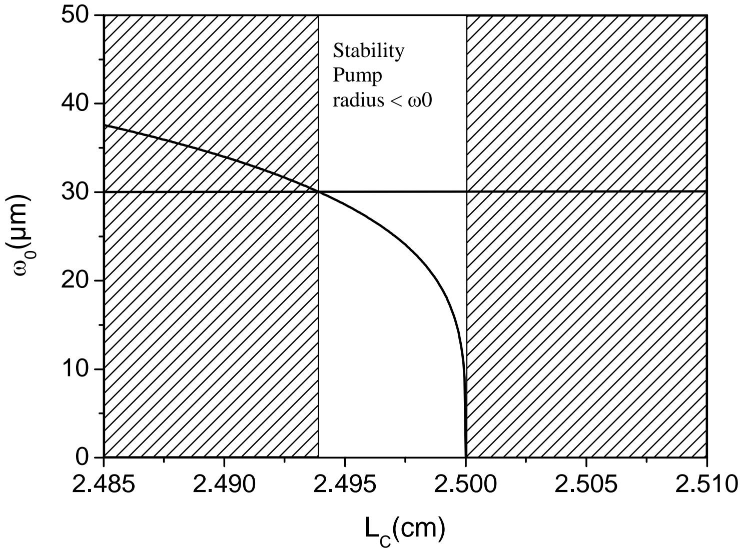


Figure 8. Effective stability zone for a plan/concave laser cavity at $\lambda = 2.3$ µm with an external mirror of Rc = 2.5 µm and a 60 µm-diameter pumping spot

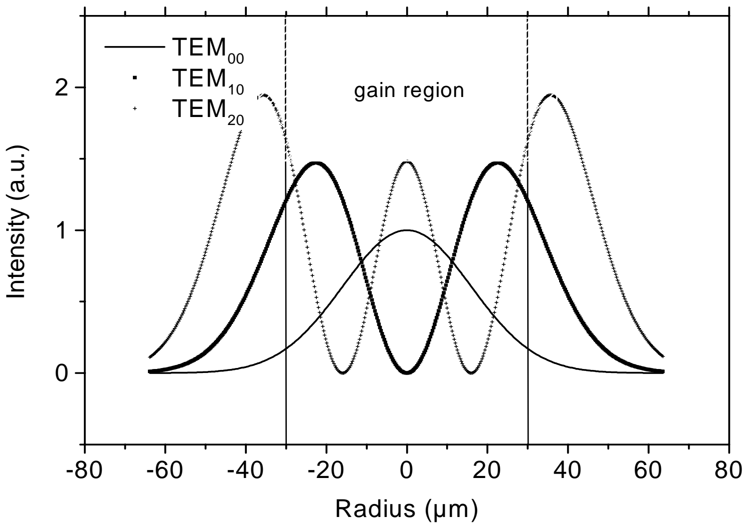


Figure 9. Transverse mode distribution for a plane/concave laser cavity at $\lambda = 2.3$ µm. Gain region is delimited by a 60 µm-diameter pump spot.

in a single transverse mode TEM_{00}. Higher order transverse modes would appear if the diameter of the pump spot is increased.

3 VCSELs Emitting in the 2-3 μm Range

Since their invention, VCSELs have been extensively studied and developed on GaAs and InP substrates with near infrared wavelengths of emission ranging from 0.8 μm up to 1.6 μm. These devices were mainly developed for data transmission applications. More recently, the intrinsic properties of operation and emission of the VCSEL device have opened new fields of applications for such a laser source, including medicine, military countermeasures or spectroscopy for example. Thus, a great research effort has been made to develop mid- and far- infrared emitting VCSELs, typically near 3 μm and above. This established fact can be illustrated by the review work of Heiss *et al.* [54] on microcavities for the mid-infrared. As described earlier in this chapter, a highly reflective Bragg mirror is a key requirement in the VCSEL structure.

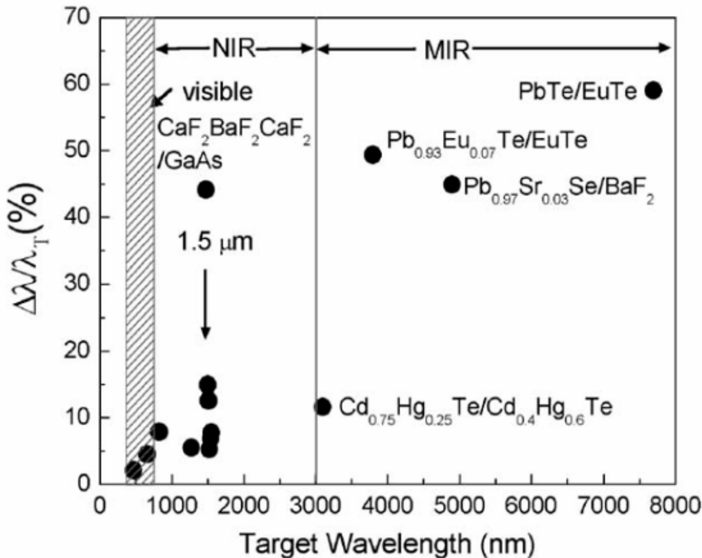


Figure 10. Relative stop band widths for various Bragg mirrors (reprinted from Progress in Quantum Electronics, 25, Heiss W., Schwarzl T., Roither J., Springholz G., Aigle M., Pascher H., Biermann K., Reimann K., "Epitaxial Bragg mirrors for the mid-infrared and their applications", 212,.Copyright 2001, with permission from Elsevier [54]).

Without a good Bragg mirror, it is not possible to fabricate an efficient VCSEL. Heiss *et al.* have reported the main results on Bragg mirrors in the infrared that have been published to date (Figure 10).

According to Figure 10 it does not appear that there are any significant results concerning Bragg mirrors in the wavelength range between 2 and 3 μm. However, several works were undertaken in this spectral range, but these remain far from numerous. To our knowledge, only five different research groups have reported results concerning electrically or optically pumped VCSEL structures emitting between 2 and 3 μm.

Table 1. State of the art of published works concerning VCSELs emitting in the 2 – 3 μm wavelength range

Research team	Active layer	Bragg mirror	Pump source	Tmax	Threshold
CEA - France	CdHgTe 2.63 μm	YF3/ZnS (6 pairs and 4 pairs)	OP	190K	1.8 kW/cm^2 (80K) 5 kW/cm^2 (190K)
Walter Schottky Institute, Munich - Germany	AlGaInAs/InP (InP) 2.012 μm	GaInAs/AlInAs (32 pairs) Si/CaF2 (2 pairs)	EP	300K (CW)	
Naval Research Laboratories / Hughes Research Laboratories-USA	InAs/GaSb/InAs/AlSb (GaSb) 2.9 μm	AlAsSb/GaSb (12.5 pairs) SiO2/Si	OP	160K (CW) 280K (pulsed)	1 kW/cm^2 (80K) 10.5 kW/cm^2 (160K)
University Montpellier 2 - France	GaInAsSb/GaSb (GaSb) 2.2 μm	AlAsSb/GaSb (17.5 pairs)	EP	300K (pulsed)	2 kA/cm^2
University Montpellier 2 - France	GaInAsSb/AlGaAsSb (GaSb) 2.3 μm – 2.36 μm (external cavity, microcavity)	AlAsSb/GaSb (23.5 pairs) SiO2/TiO2 (7 pairs)	OP	305K (microcavity) 350K (external cavity)	5 kW/cm^2 (microcavity) 1 kW/cm^2 (external cavity)

A group at CEA, Grenoble (France) has fabricated optically-pumped microcavity VCSELs with a CdHgTe-based active layer and the group at the Walter Schottky Institute, Munich (Germany) has realised and optimised very efficient electrically-pumped InP-based microcavity VCSELs emitting up to 2.012 μm. Meanwhile, collaborators from Naval Research Laboratory and Hughes Research Laboratory (USA) have published results concerning 2.9 μm emitting GaSb-based microcavity VCSELs, and the two teams of the University of Montpellier 2 (France) have realised, on GaSb substrates, optically-pumped VECSELs and both electrically and optically pumped microcavity VCSELs emitting near 2.3 μm. These different results are summarized in Table 1. These different structures and their respective performances will be decribed below.

3.1 CdHgTe-based Structures – CEA, Grenoble (France)

The first structure reported by CEA was a II-VI (CdHgTe) optically-pumped VCSEL emitting at 3.06 μm but only at very low temperature and with high threshold pump intensities [55]. Then, this research group improved this structure into an optimized optically-pumped VCSEL emitting at 2.63 μm at 190K [56]. This device was fabricated in several stages with technological processing to remove the epitaxial substrate. The active layer was first grown by molecular beam epitaxy

(MBE) on a nominal (100) $Cd_{0.96}Zn_{0.04}Te$ substrate. It consisted of five 16 nm-thick $Cd_{0.4}Hg_{0.6}Te$ quantum wells embedded in 10 nm-thick $Cd_{0.52}Hg_{0.48}Te$ barriers. This multi-quantum-well (MQW) zone was sandwiched between two 344 nm-thick spacer layers to obtain a one-wavelength-long cavity. All layers of this active region were nominally undoped. A six pair YF_3/ZnS dielectric Bragg mirror was then evaporated on top of the active layer. (Individual Bragg mirror pairs were composed of a 441 nm-thick YF_3 layer and a 289 nm-thick ZnS layer). Based on the high refractive index contrast between these two materials, a reflectivity of 99.2% was estimated for this dielectric Bragg mirror. The sample was then glued face down on a suprasil quartz plate to facilitate substrate removal, which was done by mechanical polishing followed by an appropriate selective chemical etching. Finally, the microcavity structure was completed by the evaporation of a second dielectric Bragg mirror consisting of five pairs of YF_3/ZnS. A corresponding reflectivity of 98.1% was estimated.

Figure 11 (a) plots the reflectivity spectrum of the fabricated microcavity together with the photoluminescence emission at 300K and the laser emission at 20K, and Figure 11 (b) plots the peak output intensities vs pump power recorded for different temperatures.

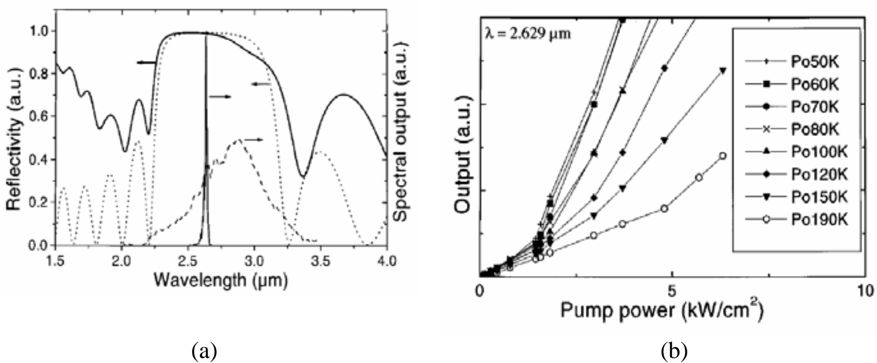


Figure 11. (**a**) Measured (*solid line*) and stimulated (*dotted line*) reflectivities of the bottom Bragg mirror, photoluminescence intensity at 300K (*dot-dashed line*), and lasing intensity (*thin solid line*) at 20K vs wavelength (**b**) Peak output intensities vs pump power at several temperatures for the five period structure (top mirror reflectivity of 98.1%). The maximum lasing temperature is 190K. These Figures are reprinted with permission from C. Roux, E. Hadji, and J.-L. Pautrat, Applied Physics Letters, 75, 3763 (1999). Copyright 1999, American Institute of Physics [56].

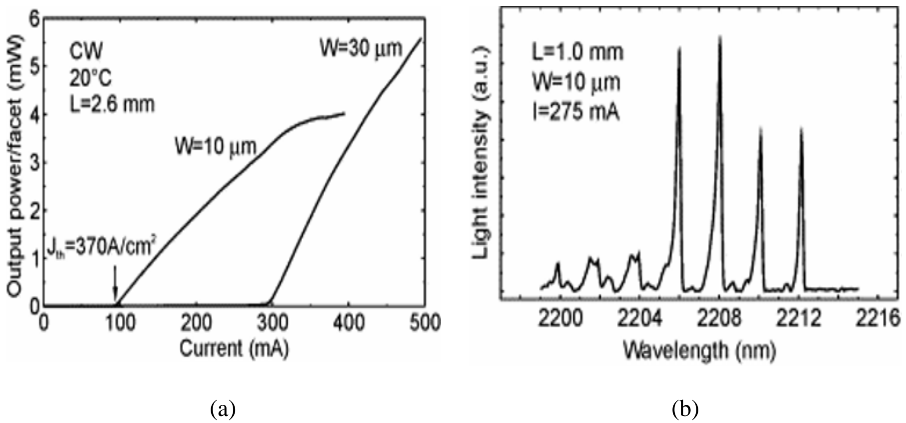
Multimode transverse laser operation in pulsed regime was observed up to 190K with a threshold ranging from 1.7 kW/cm^2 at 80K up to 4.9 kW/cm^2 at 190K. A commercially available microchip Nd:YAG laser (1.064 µm) was used as the excitation source. The Full Width at Half Maximum (FWHM) of the pump beam was ~750 µm. A reflection of 31% of the pump beam on the VCSEL surface was considered for threshold calculations. A characteristic temperature (T_0) of 113K in the range 50-190K was deduced from these measurements. The linewidth of the multimode transverse laser spectrum was estimated to be 2.8 meV.

3.2 InP-based Structures – Walter Schottky Institute, University of Munich, Garching (Germany)

The AlGaInAs/InP materials system is known to be very efficient for the fabrication of laser gain regions emitting at telecommunications wavelengths, *i.e.* 1.3 μm – 1.55 μm [57 – 59]. However, the same system was used by a group of the Walter Schottky Institute at the University of Munich (Germany) to fabricate VCSELs and edge emitting lasers emitting above 2 μm at room temperature. Owing to a MQW active region comprising two 8 nm-thick compressively strained $In_{0.84}Ga_{0.16}As$ QWs separated by a 30 nm-thick InP lattice-matched InGaAs barrier, this group demonstrated CW operation up to 373K for an edge emitting device emitting near 2.2 μm [60]. In this structure lattice-matched InGaAs and $In_{0.48}Ga_{0.42}Al_{0.10}As$ layers were used to fabricate a separate confinement heterostructure. Figure 12 (left) shows the typical CW output power-current characteristics for 2.6 mm lasers and Figure 12 (right) displays the typical spectrum of a 1.0 mm long and 10 μm wide device.

Similar active regions were inserted in microcavity structures in order to fabricate electrically-pumped VCSELs. However, the highest wavelength of operation reported for such a VCSEL was 2.014 μm at room temperature [9]. The operation of this VCSEL depends upon a buried tunnel junction (BTJ). According to the authors, the use of the BTJ concept in an InP-based VCSEL has significantly improved performance with respect to single-mode power and operation temperature. A schematic cross section of a typical BTJ-VCSEL is displayed in Figure 13. The structure is fabricated in several steps. In a first epitaxial step, the top Bragg mirror is grown followed by the active region in which compressively strained QWs are separated by weakly tensile strained barriers. This first epitaxial step is finished by growing the heavily doped (~10^{20} cm^{-3}) tunnel junction layers. After recessing the tunnel junction, it was overgrown with n-doped InP layers in a second epitaxial step. The microcavity was then formed by the deposition of a dielectric Bragg mirror. Then, the device was mounted upside-down in an integrated heatsink and the InP-substrate on the top of the structure was completely removed by selective etching.

The VCSEL emitting beyond 2 μm consisted of 31.5 pairs of InP-lattice-matched GaInAs/AlInAs layers for the top epitaxial Bragg mirror and 2.5 pairs of CaF_2/a-Si dielectric for the bottom Bragg mirror. The active layer consisted of five compressively strained InGaAs/InGaAs QWs separated by weakly strained tensile barriers. For an output emission near 2 μm, an InAs fraction of about 0.78 was used in the wells in order to obtain an average lattice-mismatch of 0.8%. The whole five QWs active region was 80 nm-thick. An output power of 0.3 mW, series resistance of 70 Ω and threshold voltage of 0.85 V was measured using a tunnel junction diameter of 6 μm. Typical P(I) and U(I) characteristics of near 2 μm emitting VCSELs are presented in Figure 14 (a). The VCSEL wavelength of emission was tuned from 2.0105 up to 2.0145 μm [9] by increasing the input current from 2 to 7 mA as shown in Figure 14 (b).



(a) (b)

Figure 12. (**a**) Typical CW output power-current characteristics at 293K for 2.6 mm long edge-emitting lasers at 2.2 µm. (**b**) Emission spectrum at RT of a 1.0 mm long and 10 µm wide edge-emitting laser. The maximum peak is located at 2208 nm. These Figures are reprinted from Journal of Crystal Growth, 227-228, Kuang G. K., Böhm G., Grau M., Rösel G., Amann M-C., "Long wavelength InGaAs-InGaAlAs-InP lasers grown in MBE", 336, Copyright 2001, with permission from Elsevier [60].

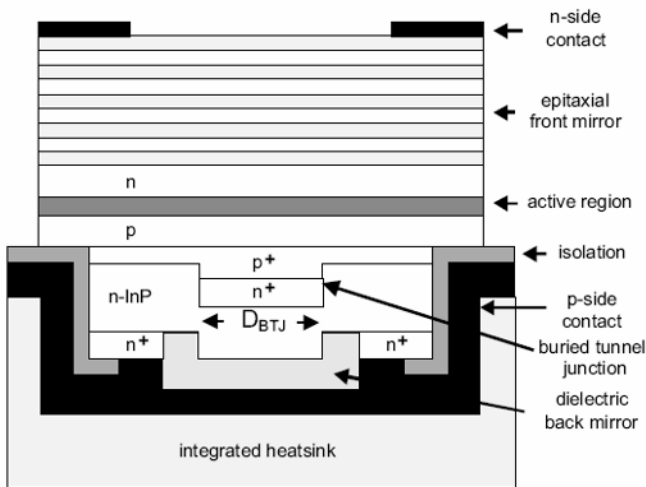


Figure 13. Schematic cross-section of a BTJ-VCSEL. Figure reprinted from Journal of Crystal Growth, 251, Boehm G., Ortseifer M., Shau R., Rosskopf J., Lauer C. Maute M., Köhler F., Mederer F., Meyer R., Amann M-C., "InP-based VCSEL technology covering the wavelength range from 1.3 to 2.0 µm", 749-752, Copyright 2003, with permission from Elsevier [9].

Currently, it seems that 2.0145 μm and 2.208 μm are the highest wavelengths of emission at room temperature of BTJ-VCSELs and edge-emitting laser diodes respectively that have been reported by this team with an AlInGaAs based active layer on InP substrate.

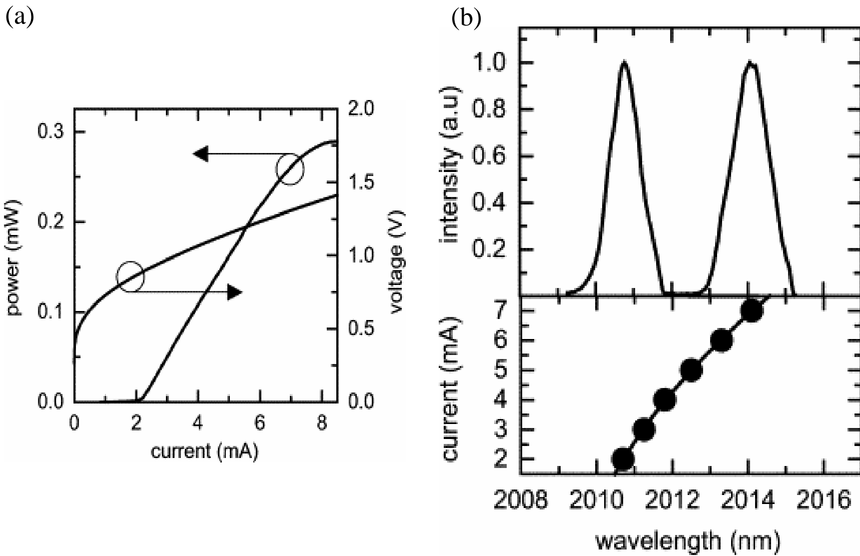


Figure 14. (**a**) Output power – current and voltage – current characteristics of a BTJ-VCSEL emitting at 2.0 μm. (**b**) Tuning characteristic of a VCSEL emitting at 2 μm : single mode emission and tuning range of about 4 nm when increasing current from 2 to 7 mA. Figures reprinted from Journal of Crystal Growth, 251, Boehm G., Ortseifer M., Shau R., Rosskopf J., Lauer C. Maute M., Köhler F., Mederer F., Meyer R., Amann M-C., "InP-based VCSEL technology covering the wavelength range from 1.3 to 2.0 μm", 749-752, Copyright 2003, with permission from Elsevier [9].

3.3 GaSb-based Structures I – Naval Research Laboratory/ Hughes Research Laboratory (USA)

Chronologically, these collaborators from Naval Research Laboratory and Hughes Research Laboratory were the first to report the fabrication and operation of a VCSEL structure emitting in the 2-3 μm wavelength range. It was also the first work realized on a GaSb substrate with Sb-based active region. The primary result dates from October 1997 with the description of the operation of a VCSEL emitting at 2.9 μm up to 280 K in pulsed regime [61]. Later, NRL/HRL collaborators published results concerning the same structure operating in the CW regime up to 160K [62].

The microcavity VCSEL structure contained a 12 ½ $GaSb/AlAs_{0.08}Sb_{0.92}$ bottom semiconductor Bragg mirror lattice-matched to the GaSb substrate. This mirror was grown in several steps. First an 11 pair Bragg mirror was grown. Then, the reflectivity spectrum of this mirror was measured in order to adjust the output

emission wavelength of the active region to the centre of the stop-band. A second epitaxial step was then performed on this sample. An additional semiconductor Bragg mirror (1½ pairs) was overgrown followed by a 79 period InAs/GaSb/InAs/AlSb (18.5 Å, 28 Å, 18.5 Å, 40 Å) so-called "W"- structure which constituted a one-wavelength-long active region adapted for high temperature laser emission at the centre wavelength of the Bragg mirror. The schematic band offsets of this "W" system are reported in Figure 15. This active region was capped with 160 Å of AlSb (to block the escape of holes from the active region), followed by 150 – 200 Å of GaSb as protection to avoid oxizidation. The total epitaxial thickness was ~ 6 µm.

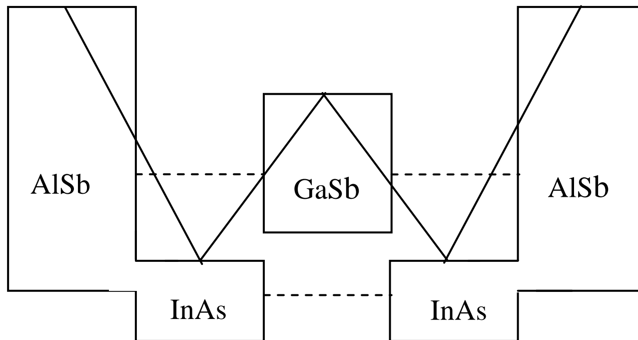


Figure 15. Schematic representation of band offsets and carriers levels (*dashed*) in VCSEL with "W" active layer (AlSb/InAs/GaSb/InAs)

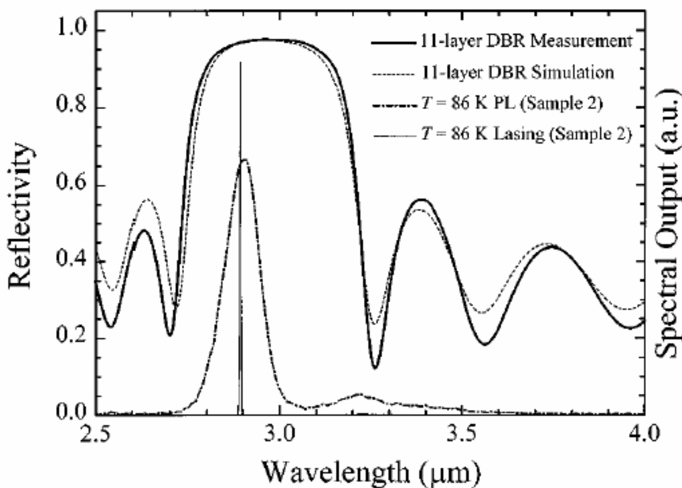


Figure 16. Reflectivity of the bottom Bragg mirror (measured and simulated), photoluminescence intensity, and pulsed lasing intensity (at 86K) vs wavelength of the GaSb-based VCSEL structure developed by Felix *et al.* (reprinted with permission from C. L. Felix, W. W. Bewley, I. Vurgaftman, J. R. Meyer, L. Goldberg, D. H. Chow, and E. Selvig, Applied Physics Letters, 71, 3483 (1997). Copyright 1997, American Institute of Physics) [61].

The reflectivity of the 11 pairs Bragg mirror before overgrowth is presented in Figure 16 together with the photoluminescence emission of the active layer at 86K. The Bragg mirror stop-band was centered at 3 µm. From this measurement, the internal reflectivity (incident from the active layer) of the total Bragg mirror (12 ½ pairs) can be estimated to be ~ 95%. However, the cavity mode appeared at a shorter wavelength (2.9 µm) than 3 µm and which would limit the laser performance.

The microcavity was completed by the deposition of a dielectric Bragg mirror with ~ 99% reflectivity at 3 µm and ~ 80% transmission at 1.06 µm. The substrate was then mechanically polished to a thickness of 50 – 70 µm and metallized with Ti – Pt – Au, and Pb – Sn soldered to the copper cold finger of a dewar.

This microcavity was first optically pumped by 85 ns pulses from a TEM_{00} Q-switched 2.1 µm Ho:YAG laser with a 1 Hz repetition rate. Several sizes of pump beam diameters were used depending on the requirement for a low threshold or a high output power. The lasing spectrum observed at 86K is presented in Figure 15. At T=86K, the pulsed threshold power is as low as 22 mW using a 30 µm excitation spot. Multimode laser operation was observed near 2.9 µm up to 280K (Figure 16), with output powers higher than 2 W up to 260K for a 600 µm excitation spot. Conversion efficiencies were as high as 4% at 120K and > 1% at 220K.

This sample was later optically pumped in the CW regime with a TEM_{00} 1.06 µm Nd:YAG laser. Multimode CW laser operation was obtained near 2.9 µm up to 180K. At T=78 K CW output power of 45 mW was obtained for a 50 µm-diameter pump spot. A low threshold pump power of 25 mW was observed at 160K for a 10 µm-diameter pump spot (Figure 17 (b)). The FWHM of the 2.9 µm laser emission peak at 78 K was measured to be 2.9 nm (with a spectrometer resolution of 0.5 nm) (Figure 17 (a)). The researchers from NRL/HRL only published results concerning optically pumped structures operating in both pulsed and CW regime.

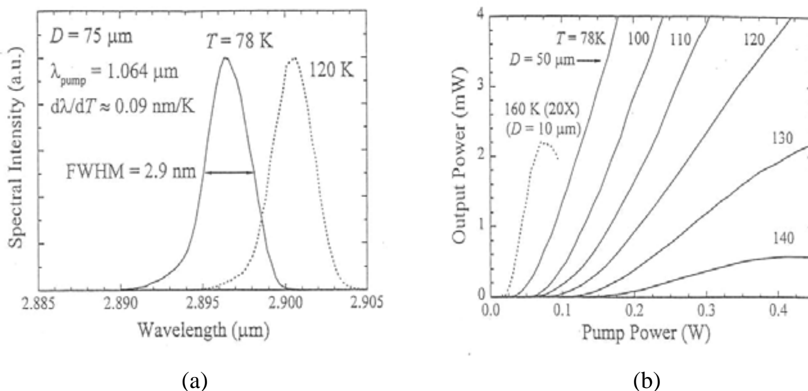


Figure 17. (**a**) Spectral output vs wavelength at T=78K and 120K of the GaSb-based VCSEL structure operating in CW. (**b**) CW output power vs pump power characteristics at several temperatures of the GaSb-based VCSEL structure. *Solid curves* are for a 50 µm-diameter pump spot, *dashed curve* is for a 10 µm-diameter pump spot. Reprinted with permission from Bewley W. W., Felix C. L., Vurgaftman I., Aifer E. H., Meyer J. R., Goldberg L., Lindle J. R., Chow D. H., Selvig E., IEEE Photonics Technology Letters, 10, 661 (1998). Copyright 1998. The Institute of Electrical and Electronics Engineers Inc. [62].

3.4 GaSb-based Structures II – University of Montpellier 2 (France)

3.4.1 Electrically-pumped Microcavity VCSEL Emitting Near 2.2 µm

The first electrically-pumped VCSEL emitting in the 2–3 µm wavelength range was demonstrated in 1998 by a group from the University of Montpellier 2. This monolithic VCSEL operated near 2.2 µm at room temperature in the pulsed regime [63]. This structure was grown in one MBE run on a (100) n-doped GaSb substrate. It was a classical monolithic VCSEL structure (as described in Figure 1) with bottom (Te-doped, 17.5 pairs) and top (Be-doped, 15 pairs) GaSb/$AlAs_{0.07}Sb_{0.93}$ Bragg mirrors. These two Bragg mirrors sandwiched a one-wavelength-long cavity containing a Type-II active zone. This MQW Type-II active layer was made of six 6.5 nm-thick $Ga_{0.65}In_{0.35}As_{0.1}Sb_{0.9}$ compressively strained quantum wells embedded in 22 nm-thick GaSb barriers. This MQW zone was surrounded by 225 nm of GaSb spacer layers to form the one-wavelength-long optical cavity. The schematic band diagram of such a Type-II active layer is shown in Figure 18. In this case, only electrons are confined and the radiative recombination occurs between the e_1 level of the well and the holes at the top of the GaSb barrier valence band. The last p-doped GaSb layer of the top semiconductor Bragg mirror was also used as the contact layer. Mesa diodes (200 µm-diameter) were fabricated using wet chemical etching. Edges of the mesa were passivated with anodic oxide. Before processing, the optical transmission was measured at different points on the grown wafer. The spectral position of the centre of the stop-band varied from 2.28 µm in the middle of the wafer to 2.15 µm at a distance of 2 cm away (at room temperature).

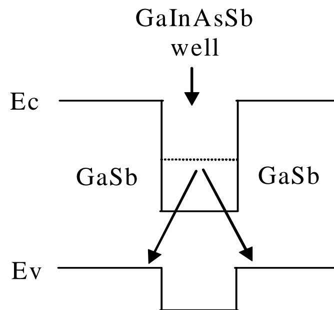


Figure 18. Schematic band diagram of the Type-II GaInAsSb/GaSb QW. The radiative recombination occurs between an electron level in the well and the highest hole level in the valence band of the GaSb barriers.

Despite inhomogeneous growth, multimode lasing operation in the electrically pulsed regime (repetition rate 10 KHz, pulsewidth 1 µs) was achieved at room temperature with emission wavelength ranging from 2.22 µm up to 2.28 µm with devices fabricated from different parts of the wafer. The FWHM of the cavity lasing mode was measured to be narrower than the spectrometer resolution (6nm) (Figure 19 (a)). Lasing threshold current densities were higher than 2 kA/cm^2 (Figure 19

(b)). The threshold in the light-current curves and the narrowing of the emission line became less pronounced when the duration of the current pulse exceeded 500ns, even at a duty factor of less than 1%. This behaviour was attributed to overheating originating from the high potential barriers in the GaSb/AlAsSb Bragg mirrors.

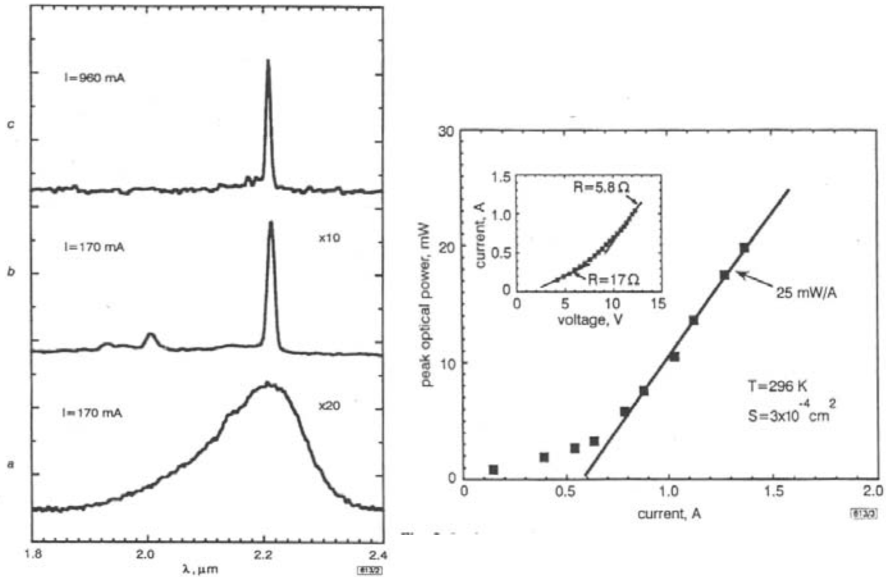


Figure 19. (**a**) Room temperature spontaneous and lasing emission spectra of Sb-based EP-VCSEL a – spontaneous spectrum measured from the edge of the device (20 μs, 20 KHz); b – spontaneous spectrum measured from the surface (20 μs, 20 KHz); c – surface emission laser spectrum (1 μs, 10 KHz). (**b**) Peak optical power of VCSEL vs pulsed current. In inset : current-voltage characteristic of VCSEL. Figures reprinted with permission from Baranov A. N., Rouillard Y., Boissier G., Grech P., Gaillard S., Alibert C., Electronics Letters, 34, 282 (1998). Copyright 1998. The Institution of Electrical Engineers [63].

3.4.2 Optically-pumped GaSb-based VCSEL Structures

3.4.2.1 VCSEL and VECSEL Fabrication

The University of Montpellier 2 team fabricated several OP-VCSELs to demonstrate a CW single transverse mode TEM_{00} laser operating near 2.3 μm. Diode pumping of these devices with an inexpensive 830 nm diode laser allows one to achieve a low divergence beam and high power laser operation without the need for post-growth processing and at low cost. In these structures there was no need to dope the Bragg mirror which also reduced optical losses and simplified the growth. All the structures were manufactured using a similar process. First, an epitaxial ½ VCSEL was grown by MBE on an undoped GaSb (100) substrate using As and Sb valved cracker cells. It consisted of a bottom DBR comprising 23.5 pairs of $GaSb/AlAs_{0.07}Sb_{0.93}$ surmounted by a Type-I MQW active zone (Figure 20).

Considering the high refractive index contrast occurring around 2.3 µm between these two alloys (Δn~0.6), a reflectivity higher than 99.7% was estimated. The gain region was formed by a one-wavelength long cavity, containing five compressively-strained 10nm-thick $Ga_{0.65}In_{0.35}As_{0.1}Sb_{0.9}$ Type-I quantum wells, separated by 20nm-thick $Al_{0.35}Ga_{0.65}As_{0.04}Sb_{0.96}$ barriers. The schematic band diagram of this MQW zone is drawn in Figure 21. In contrast to the previous electrically pumped VCSEL, aluminium was added in the barriers which allows one to obtain Type-I MQWs instead of Type-II. With this structure, high electron and hole confinement in the well was ensured at 2.3µm, *i.e.* around 490meV for the electrons and 170meV for the heavy holes, which should warrant a high quantum efficiency and operation at high temperature [64]. The barrier bandgap was 1.05µm, allowing efficient optical pumping with a commercial 830 nm diode laser.

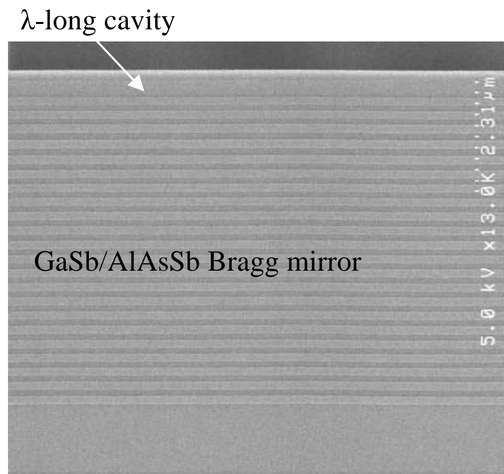


Figure 20. Scanning electron micrograph of a typical epitaxial ½ VCSEL structure grown by MBE at University of Montpellier 2

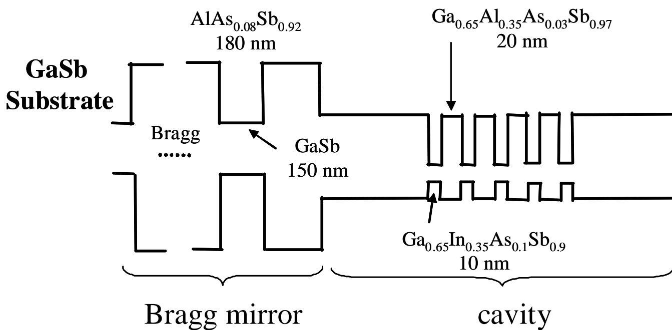


Figure 21. Schematic band diagram of the ½ VCSEL structure containing a cavity with 5 Type-I GaInAsSb/AlGaAsSb quantum wells

From this ½ VCSEL structure, the Montpellier team fabricated both microcavity and VECSEL sources. To manufacture a microcavity, a top dielectric Bragg mirror (SiO_2/TiO_2 or YF_3/ZnS) was evaporated on the ½ VCSEL by ion beam assisted deposition. A typical scanning electron micrograph of a microcavity is shown in Figure 22.

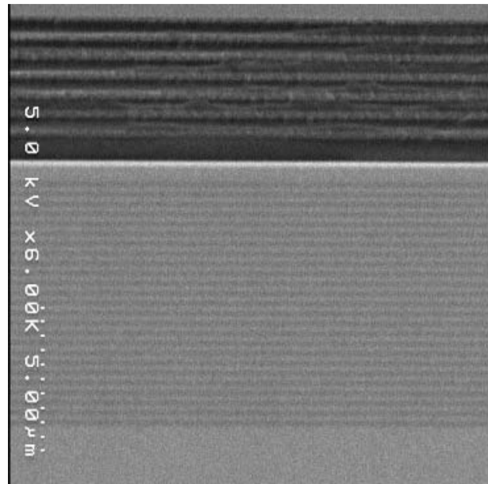


Figure 22. Scanning electron micrograph of a complete microcavity structure with a bottom AlAsSb/GaSb DBR and a top SiO_2/TiO_2 DBR

One of the main limitations of VCSELs comes from overheating due to a bad dissipation of the heat generated by the pumping. A solution generally used is to remove the substrate and to bond the device on a heatsink or to add a heat-spreader layer at the top of the active region [65–68]. The first VECSEL structure fabricated by University of Montpellier 2 operated at 2.1 µm at room temperature [69]. Unfortunately, this device suffered from a low characteristic temperature, T_0 (~33K) which did not allow CW mode operation at room temperature. This low T_0 was believed to be caused by a lower internal quantum efficiency due to dislocations/impurities and also by the apparition of a thermal lens due to the overheating at the surface of the structure which increases losses by diffraction. To improve the heat dissipation, a heat-spreader layer was added at the top of the active region. The best candidates for heat dissipation layers are binary alloys such AlSb because of their high thermal conductivities. With regard to this ½ VCSEL structure, this heat-spreader layer must be lattice-matched to GaSb, so AlSb cannot therefore be used. However, the thermal conductivity of AlAsSb lattice-matched to GaSb is around 10 –20 W/mK, a value much lower than that of AlSb or GaSb but higher than AlGaAsSb, the material of the barrier [70], and the AlAsSb/GaSb multilayer Bragg structure. So, instead the active MQW structure was covered with a heat-spreader made of a multiple of half-lambda thick AlAsSb layers. The improved ½ VCSEL structure was therefore composed of a bottom GaSb/AlAsSb Bragg mirror,

a one-wavelength-long MQW GaInAsSb/AlGaAsSb cavity capped by a three-wavelength-long AlAsSb heatspreader layer and a 5 nm-thick GaSb surface layer to prevent oxidation. This reduced the thermal impedance of the device by about 25% for a 30µm pump spot diameter, leading to a typical value of about 600 K/W relative to the incident pump power and averaged over the laser beam.

3.4.2.2 Characterization of Optically-pumped Microcavity and External Cavity VCSELs Emitting in CW Regime Near 2.3 µm at Room Temperature - Single Frequency Operation and Frequency Tunability

The main difficulty occuring during the fabrication of a vertical cavity laser is to match the MQW gain peak with the cavity mode. In early work, the University of Montpellier 2 group fabricated a microcavity VCSEL structure emitting at 2.2µm [71] with an evaporated ZnS/YF_3 stack as top mirror. The semiconductor part of this VCSEL (bottom Bragg mirror + MQW zone) was grown in only one MBE run. Figure 23 shows the reflectivity spectrum of the semiconductor part of this VCSEL combined with its edge-photoluminescence properties. The calculated reflectivity spectrum (not shown) is in good agreement with the experimental one. As exhibited in Figure 23, this device unfortunately suffered from a large wavelength red shift of the gain peak compared to the micro-cavity mode wavelength. Such a shift limited its operating temperature at 100K in pulsed regime operation.

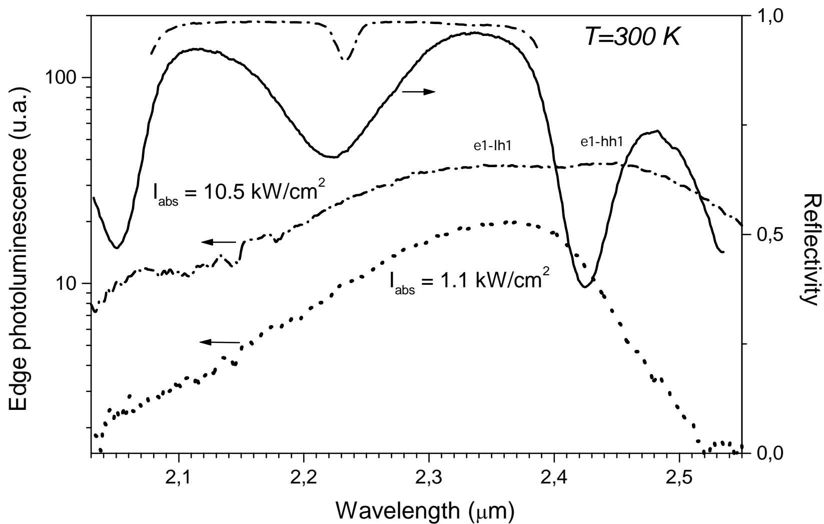


Figure 23. Room-temperature reflectivity and edge-emitting photoluminescence spectra at two different pump powers for the first ½ VCSEL structure . The *dash-dot* line shows the reflectivity spectrum of the full microcavity structure after deposition of the dielectric coating

An improved optically-pumped VCSEL structure emitting near 2.3 µm [72] and operating in CW regime above room temperature, in both microcavity and external-cavity configurations, was demonstrated following optimisation. The reflectivity and photoluminescence spectra are shown in Figure 24. This shows the good matching

between the peak gain and the designed Bragg mirror wavelength. The calculated reflectivity spectrum also agrees very well with the experimental one. The broadband absorption/diffraction losses in the ½ VCSEL structure were measured to be below 2 cm^{-1}. At 300 K, the internal quantum efficiency, in the spontaneous regime, of the quantum well active region was measured to be about 20 – 30 %, in good agreement with the theory.

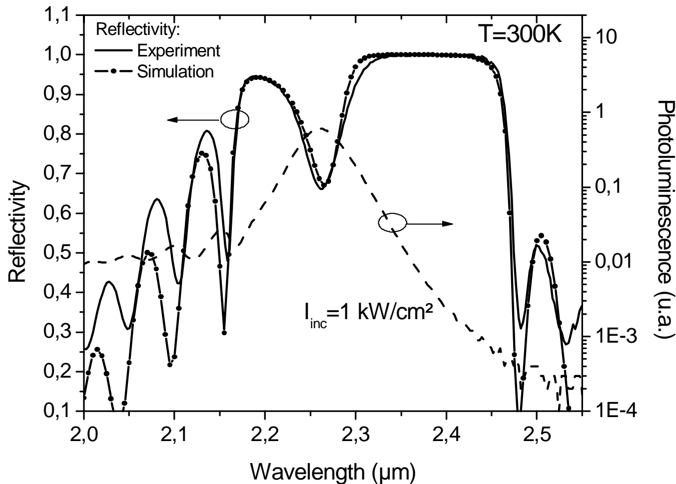


Figure 24. Room-temperature reflectivity (measured and calculated) and top emitting photoluminescence spectra of the improved ½ VCSEL structure. I_{inc} is the incident pump density.

A. Ouvrard *et al.* [73], have demonstrated two kinds of diode-pumped laser cavity: (i) a dielectric Bragg mirror made of seven pairs of Ti_2O_3/SiO_2 (reflectivity R=99 %) was formed by sputtering on the ½ VCSEL structure to form a microcavity [74]; (ii) an external-cavity (VECSEL) formed by the ½ VCSEL structure and an external concave dielectric mirror (reflectivity R = 99.5 %), placed in the air at a distance slightly shorter than the 15 mm mirror radius of curvature. VECSELs show higher laser frequency stability and lower drift, as the cavity mode frequencies are almost decoupled from the structures temperature and pump power fluctuations. The single mode 830 nm low power pump diode was focused to a 30 µm diameter (at $1/e^2$) spot (with TM incidence) close to the Brewster angle, to produce a circular beam and reduce the pump reflectivity. Laser operation was obtained near 2.3 µm in CW up to 350K with threshold pump densities as low as 600 W/cm² near 273 K, as low as GaAs-based diode pumped VCSEL. The microcavity device shows a higher threshold of about 8.5 kW/cm^2, due to a slight mismatch between the microcavity mode and the gain peak wavelength. The characteristic temperature, T_0, of the VECSEL varied from 74 K to 44 K between 275 K and 350 K. Both lasers exhibit a TEM_{00} circular beam (Figure 25) with a low divergence, but with the laser beam becoming rapidly multimode in the case of the microcavity device. Increasing the pump spot size and pumping with a high power diode laser, allowed an increase in the maximum output power of the device up to 10mW at 274K, but the transverse

beam remained TEM_{00} only in the case of the VECSEL configuration. The external quantum efficiency, relative to the incident pump power at 830 nm, was measured to be about 11 % for the VECSEL and 15 % for the microcavity VCSEL, showing the high internal quantum efficiency under stimulated emission (close to unity), in agreement with the low threshold densities measured in these devices.

VCSELs exhibit single-frequency operation and a wide tuning range free from mode-hops and so are particularly well adapted for trace gas spectroscopy applications, especially in the 2 - 2.5 μm atmospheric transparent window.

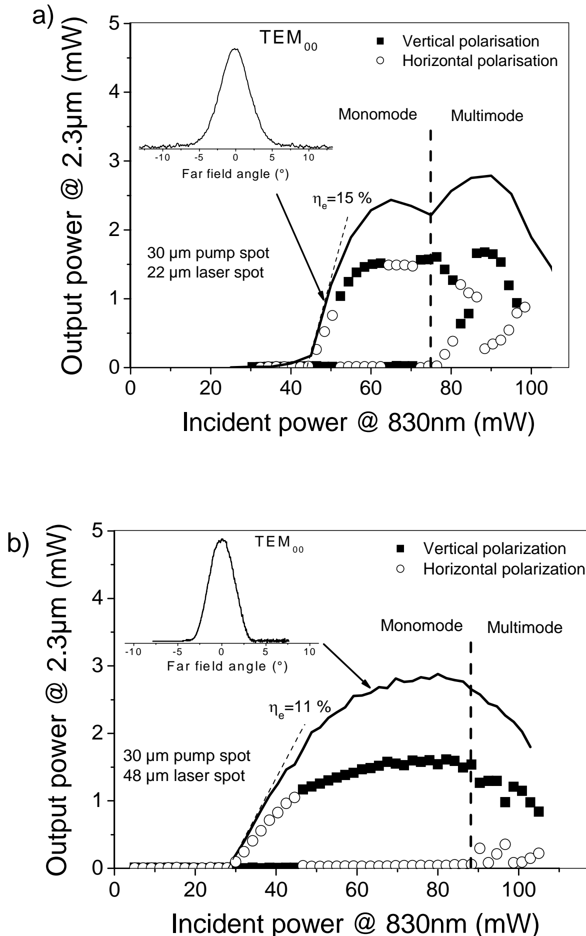


Figure 25. Output power vs incident pump power in CW of the VCSEL (**a**) at 273 K and the VECSEL (**b**) at 283 K. *Inset:* output beam far field distribution [73]

The VCSEL *(VECSEL)* device developed by A. Ouvrard *et al.* [73] exhibits single frequency operation up to 2.5 mW *(3 mW)* output power at 273 K *(283 K)* with a low divergence circular TEM_{00} beam (Figure 25). While operating in single transverse mode, the polarisation is linear along the [110] or [1-10] crystal axis, with

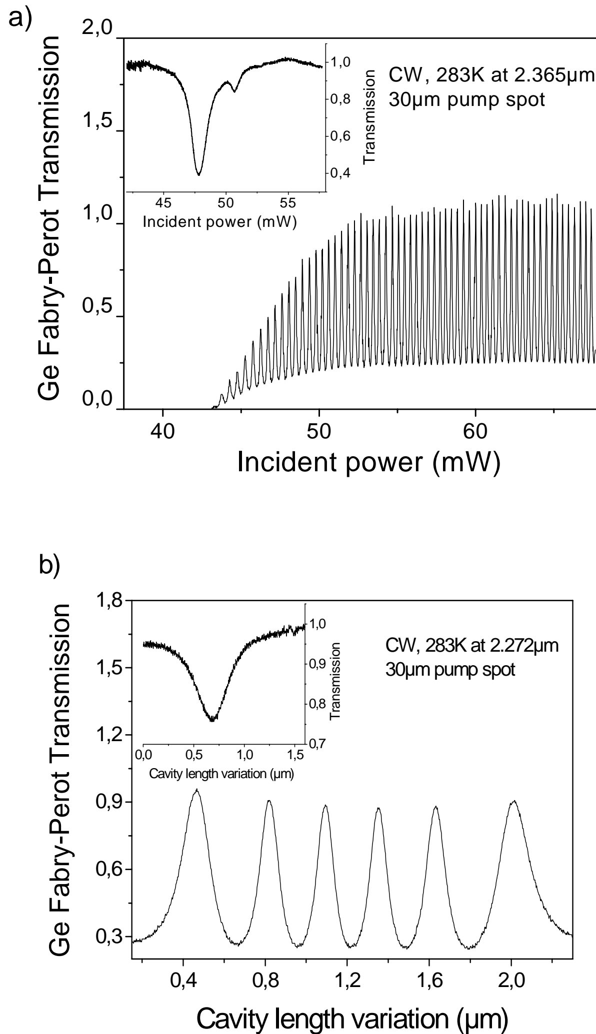


Figure 26. Output power transmitted through a germanium Fabry-Perot etalon vs incident pump power (VCSEL (**a**), VECSEL (**b**)). *Inset*: CH_4 absorption spectrum [74].

an extinction ratio of 1:130 *(1:2500)* (Figure 25). These lasers exhibit complete polarisation switching between these crystal axes as the pump is varied. In the case of the microcavity, when a polarization switch occurs a laser frequency jump of 23 GHz is measured, leading to an average birefringence of ~10^{-4} in the ½ VCSEL structure. VECSEL devices show a higher polarisation stability and are less sensitive to feedback compared to microcavity devices. For both devices, a side mode suppression ratio better than 20 dB is measured (apparatus noise limited).

For the VCSEL device, continuous frequency tuning is achieved by modulating the pump power. At best, mode hop free tuning over 138 GHz was measured

(Figure 26), but with a typical value around 40 GHz; the tuning rate is - 4.8 GHz/mW. For the VECSEL device, continuous frequency tuning is achieved by modulating the cavity length with a piezo mounted output mirror, and the maximum continuous tuning achieved (not shown) is 25 GHz for the 15 mm long cavity, still limited by the piezo (example in Figure 26). The temperature wavelength tuning rate is -8 GHz/K for the microcavity device and -50 GHz/K for the VECSEL device. This is a great advantage of VECSELs in that one is able to achieve broad wavelength tuning from a single device.

Thanks to the long extended cavity, VECSELs have a laser linewidth much narrower than a micro-cavity (< 1/10,000). The VECSEL linewidth was measured to be << 20 kHz (apparatus limited), using a high finesse scanning Fabry-Perot interferometer. The weak noise on the measurement leads to the conclusion that the real laser linewidth was much narrower than the linewidth of the interferometer. For this VECSEL device, the theoretical limit of Schallow-Townes gives about 1 Hz for 1 mW output power, but the practical limit of around 1 kHz is mainly determined by mechanical fluctuations. For the microcavity device, the theoretical limit is above 10 MHz for a 1 mW output, but pump jitter could limit the laser linewidth above this value. In conclusion, 2.3 µm single-frequency tunable VECSEL and microcavity devices operating at room-temperature with a circular TEM_{00} beam have been developed by the Montpellier 2 team. Methane detection has been reported (100% saturated cell) with a signal to noise ratio of 1000 (Figure 26). This work shows the suitability of Sb-based diode-pumped VCSEL devices for high resolution spectroscopy and gas detection applications.

4 Conclusion

In this chapter, we tried to give a precise picture of the current state of the art of VCSEL technology in the 2-3 µm spectral range. We described classical microcavities but also structures with external-cavity geometry which allows one to fabricate devices with improved output beam properties (single transverse mode TEM_{00} operation) at high power. The electrically-pumped microcavities with highest performance in the 2-3 µm range are obtained from the AlGaInAs/InP materials system. But $\lambda = 2.1$ µm seems to represent the high wavelength limit of such devices. Concerning optically-pumped devices, Sb-based structures seem to have the best properties. Circular TEM_{00} output beam and single frequency laser operation was obtained above room temperature from both microcavities and external cavity devices with Type-I GaInAsSb/AlGaAsSb MQW active regions.

These mid-infrared emitting VCSELs were developed essentially for spectroscopic applications. The single mode operation, together with the broad continuous tuning range (without mode-hops) and the high repetition rate of VCSELs, are particularly interesting and highly suitable for gas detection or environmental monitoring. However, the recent development of vertical-cavity structures combining high power and high output beam quality opens supplementary fields of applications, for example in medicine, fiber amplifier pumping, free-space communications or military infrared countermeasures.

Acknowledgments

We are indebted to A. Ouvrard, G. Boissier and P. Grech for MBE fabrication, processing and characterization of devices. We also gratefully acknowledge C. Alibert for his encouragement and his technical co-operation.

References

1. Li H, Iga K. Vertical-Cavity Surface-Emitting Laser Devices, Springer Series in Photonics, 2002
2. Wilmsen C, Temkin H, Coldren LA. Vertical-Cavity Surface-Emitting Lasers, Cambridge University Press, 1999
3. Choi HK. Long-wavelength infrared semiconductor lasers. Wiley Series in Lasers and Applications, 2004
4. Melngailis I. Longitudinal injection-plasma laser of InSb. Appl. Phys. Lett. 1965; 6: 59-61
5. Iga K, Koyama F, Kinoshita S. Surface emitting semiconductor-lasers. IEEE J. Quantum Electron. 1988; 24: 1845-1855
6. Soda H, Iga K, Kitahara C, Suematsu Y. GaInAsP-InP surface emitting injection-lasers. Japon. J. of Appl. Phys. 1979; 18: 2329-2330
7. Lee YH, Tell B, Brown-Goebeler KF, Leibenguth RE, Mattera VD. Deep-red continuous wave top-surface-emitting vertical-cavity AlGaAs superlattice lasers. IEEE Photon. Technol. Lett. 1991; 3: 108-109
8. Baba T, Yogo Y, Koyama F, Iga K. Near room temperature continuous-wave lasing characteristic of GaInAsP/InP surface emitting laser. Electron. Lett. 1993; 29: 913-914
9. Boehm G, Ortsiefer M, Shau R, *et al*. InP-based VCSEL technology covering the wavelength range from 1.3 to 2.0 μm. J. Crystal Growth 2003; 251: 748-753
10. Schwarzl T, Heiss W, Springholz G, Aigle M., Pascher H. 6μm vertical cavity surface emitting laser based on IV-VI semiconductor compounds. Electron. Lett. 2000; 36: 322-324
11. Fürst J, Schwarzl T, Böberl M, Pascher H, Springholtz G, Heiss W. Vertical-cavity surface emitting lasers in the 8-μm midinfrared spectral range with continuous-wave and pulsed emission. IEEE J. Quantum Electron. 2004; 40: 966-969
12. Totschnig G, Lackner M, Shau R, *et al.* 1.8-μm vertical-cavity surface-emitting laser absorption measurements of HCl, H_20 and CH_4. Meas. Sci. Technol. 2003; 14: 472-478
13. Wang J, Sanders ST, Jefries JB, *et al.* Oxygen measurements at high pressures with verticl caviy surface-emitting lasers. Appl. Phys. B. 2001; 72: 865-872
14. Kaspi R, Ongstad A, Dente GC, *et al*. High power and high brightness from an optically pumped InAs/InGaSb Type-II midinfrared laser with low confinement. Appl. Phys. Lett. 2002; 81: 406-408
15. Waynant RW, Ilev IK, Gannot I. Mid-infrared laser applications in medicine and biology. Phil. Trans. of the Royal Society of London. Series A-Math. Phys. and Engineering Sciences 2001; 359: 635-644
16. Werle P, Slemr F, Maurer K, Kormann R, Mucke R, Janker B. Mid-infrared laser-optical sensors for gas analysis. Optics and lasers in engineering 2002; 37: 101-114
17. Jelinkova H, Pasta J, Nemec M, *et al*. Different influence of long and short mid-infrared laser pulses on eye tissue. Laser Physics 2003; 13: 735-742
18. Schade W, Willer U, Wondraczek L. Mid-infrared laser sensors for mapping environment and combustions. Glass Science and Technology 2003; 76: 109-114

19. Debray JP, Sagnes I, Le Roux G, *et al*. MOVPE growth of a monolithic VCSEL at 1.56 µm in the InGaAlAs-InAlAs system lattice matched to InP. IEEE Photon. Technol. Lett. 1999; 11: 770-772
20. Streubel K, André J, Wallin J, Landgren G. Fabrication of 1.5 µm optically pumped $Ga_{1-x}In_xAs_yP_{1-y}$/InP vertical-cavity surface-emitting lasers. Mater. Sci. Eng. 1994; B28: 289-292
21. Almuneau G, Hall E, Mukaihara T, Nakagawa S, Luo CY, Clarke DR, Coldren LA. Improved electrical and thermal properties of InP-AlGaAsSb Bragg mirrors for long-wavelength vertical-cavity lasers. IEEE Photon. Technol. Lett. 2000; 12: 1322-1324
22. Harmand JC, Ungaro G, Sagnes I, *et al*. Room temperature continuous wave operation under optical pumping of a 1.48 µm vertical cavity laser based on AlGaAsSb mirror. J. Crystal Growth 1999; 201/202: 837-840
23. Boucart J, Starck C, Plais A, *et al*. RT pulsed operation of metamorphic VCSEL at 1.55 µm. Electron. Lett. 1998; 34: 2133-2135
24. Boucart J, Starck C, Gaborit F, *et al*. 1-mw CW-RT monolithic VCSEL at 1.55 µm. IEEE Photon. Technol. Lett. 1999; 6: 629-631
25. Babic DI, Streubel K, Mirin RP, *et al*. Room-temperature continuous-wave operation of 1.54 µm vertical-cavity lasers. IEEE Photon. Technol. Lett. 1995; 7: 1225-1227
26. Genty F, Almuneau G, Chusseau L, Wilk A, Gaillard S, Boissier G, Grech P, Jacquet J. Growth and characterization of vertical cavity structures on InP with GaAsSb/AlAsSb Bragg mirrors for 1.55 µm emission. J. Crystal Growth 1998; 201/202: 1024-1027
27. Anni M, Gigli G, Cingolani R, Patané S, Arena A, Allegrini M. Organic µ cavities based on thermally evaporated $Te0_x$-LiF distributed Bragg reflectors. Appl. Phys. Lett. 2001; 79: 1381-1383
28. Blum 0, Hafich MJ, Klem JF, Baucom K, Allerman A. Wet thermal oxidation of AlAsSb against As/Sb ratio. Electron. Lett. 1997; 33: 1097-1099
29. Langenfelder T, Grothe H. Optimisation of λ = 850 nm hybrid-mirror vertical-cavity surface-emitting laser with 37 µA threshold current. IEE Proc. Optoelectron. 2000; 147: 56-60
30. Lear KL, Choquette KD, Schneider RP, Kilcoyne SP, Geib KM. Selectively oxidized vertical-cavity surface-emitting lasers with 50-percent power conversion efficiency. Electron. Lett. 1995; 31: 208-209
31. Weigl B, Grabherr M, Jung C, *et al*. High-performance oxide-confined GaAs VCSEL's. IEEE J. Select. Topics Quantum Electron. 1997; 3: 409-415
32. Zhang SZ, Margalit NM, Reynolds TE, Bowers JE. 1.54-µm vertical-cavity surface-emitting laser transmission at 2.5 Gb/s. IEEE Photon. Technol. Lett. 1997; 9: 374-376
33. Keating A, Black A, Karim A, *et al*. High-temperature optically pumped 1.55 µm VCSEL operating at 6 Gb/s. IEEE Photon. Technol. Lett. 2000; 12: 116-118
34. Grabherr M, Jäger R, Miller M, *et al*. Bottom-emitting VCSEL's for high-CW optical output power. IEEE Photon. Technol. Lett. 1998; 12: 1061-1063
35. Grabherr M, Miller M, Jäger R, *et al*. High-Power VCSEL's: single devices and densely packed 2-D-arrays. IEEE J. Select. Topics Quantum Electron. 1999; 5: 495-502
36. Yan C, Ning Y, Qin L, *et al*. A high power InGaAs/GaAsP vertical-cavity surface-emitting laser and its temperature characteristics. Semicond. Sci. Technol. 2004; 19: 685-689
37. Francis D, Chen HL, Yuen W, Li G, Chang-Hasnain C. Monolithic 2D-VCSEL array with > 2W CW and > 5W pulsed output power. Electron. Lett. 1998; 34: 2132-2133
38. Salet P, Plais A, Derouin E, Fortin C, Starck C, Jacquet J, Brillouet F. Undercut ridge structures : A novel approach to 1.3/1.55 µm vertical-cavity lasers designed for continuous-wave operation. IEE Proc. –Optoelectron. 1998; 145: 125-131
39. Sun D, Fan W, Kner P, *et al*. Sub-mA threshold 1.5 µm VCSELs with epitaxial and dielectric DBR mirrors. IEEE Photon. Technol. Lett. 2003; 15: 1677-1679

40. Zanatta JP, Noel F, Ballet P, Hdadach N, Million A, Destefanis G, Mottin E, Kopp C, Picard E, Hadji E. HgCdTe molecular beam epitaxy material for microcavity light emitters : Application to gas detection in the 2 – 6 μm range. J. Electron. Mat. 2003; 32: 602-607
41. Weigl B, Grabherr M, Michalzik R, Reiner G, Ebeling KJ. High-power single mode selectively oxidized vertical-cavity surface-emitting lasers. IEEE Photon. Technol. Lett. 1996; 8: 971-973
42. Young DB, Scott JW, Peters FH, *et al.* Enhanced performance of offset-gain high-barrier vertical-cavity surface-emitting lasers. IEEE J. Quantum Electron. 1993; 29: 2013-2021
43. Yokouchi N, Miyamoto T, Uchida T, Inaba Y, Koyama F, Iga K. 40 Å continuous tuning of a GaInAsP/InP vertical-cavity surface emitting laser using an external mirror. IEEE Photon. Technol. Lett. 1992; 4: 701-703
44. Hadley MA, Wilson GC, Lau KY, Smith JS. High single-transverse-mode output from external-cavity surface emitting laser diodes. Appl. Phys. Lett. 1993; 63: 1607-1609
45. Wilson GC, Hadley MA, Smith JS, Lau KY. High single-mode output power from compact external microcavity surface-emitting laser diode. Appl. Phys. Lett. 1993; 63: 3265-3267
46. Giudice GE, Kuksenkov DV, Grave de Peralda L, Temkin H. Single-mode operation from an external cavity controlled vertical-cavity surface-emitting laser. IEEE Photon. Technol. Lett. 1999; 11: 1545-1547
47. Sandusky JV, Brueck SR. A CW external-cavity surface-emitting laser. IEEE Photon. Technol. Lett. 1996; 8: 313-315
48. Kuznetsov M, Hakimi F, Sprague R, Mooradian A. High-power (>0.5-W CW) diode-pumped vertical-external-cavity surface-emitting semiconductor lasers with circular TEM_{00} beams. IEEE Photon. Technol. Lett. 1997; 9: 1063-1065
49. Kuznetsov M, Hakimi F, Sprague R, Mooradian A. Design and characteristics of high-power (>0.5-W CW) diode-pumped vertical-external-cavity surface-emitting semiconductor lasers with circular TEM_{00} beams. IEEE J. Select. Topics Quantum Electron. 1999; 5: 561-573
50. Lutgen S, Albrecht T, Brick P, Reill W, Luft J, Späth W. 8-W high-efficiency continuous-wave semiconductor disk laser at 1000 nm. Appl. Phys. Lett. 2003; 82: 3620-3622
51. Garnache A, Kachanov AA, Stoeckel F, Houdre R. Diode-pumped broadband vertical-external-cavity surface-emitting semiconductor laser applied to high-sensitivity intracavity absorption spectroscopy. J. Opt. Soc. Amer. B. 2000; 17: 1589-1598
52. Garnache A, Hwang WY, Hoogland S, *et al.* 1.5 μm high-power circular TEM00 surface-emitting laser operating in CW at 300K. Proc. IEEE IPRM 2002; Post-deadline paper #3
53. Symonds C, Dion J, Sagnes I, *et al.* High performance 1.55 μm broadband integrated dielectric-metal mirror. Electron. Lett. 2004; 40: 734-735
54. Heiss W, Schwarzl T, Roither J, *et al.* Epitaxial Bragg mirrors for the mid-infrared and their applications. Progress in Quantum Electronics. 2001; 25: 193-228
55. Hadji E, Bleuse J, Magnea N, Pautrat JL. Photopump infrared vertical-cavity surface-emitting laser. Appl. Phys. Lett. 1996; 68: 2480-2482
56. Roux C, Hadji CE, Pautrat JL, 2.6 μm optically pumped vertical-cavity surface-emitting laser in the CdHgTe system. Appl. Phys. Lett. 1999; 75: 3763-3765
57. Mondry MJ, Tarsa EJ, Coldren LA. Molecular beam epitaxial growth of strained AlGaInAs multi-quantum well lasers on InP. J. of Electron. Mater. 1996; 25: 948-954
58. Chen TR, Chen PC, Ungar J, Newkirk MA, Oh S, BarChaim N. Low-threshold and high-temperature operation of InGaAlAs-InP lasers. IEEE Photon. Technol. Lett. 1997; 9: 17-18

59. Ohnoki N, Okazaki G, Koyama F, Iga K. Record high characteristic temperature (T0 = 122K) of 1.55 µm strain-compensated AlGaInAs/AlGaInAs MQW lasers with AlAs/AlInAs multiquantum barrier. Electron. Lett. 1999; 35: 51-52
60. Kuang GK, Böhm G, Grau M, Rösel G, Amann MC. Long wavelength InGaAs-InGaAlAs-InP lasers grown in MBE. J. Crystal Growth 2001; 227-228: 334-337
61. Felix CL, Bewley WW, Vurgaftman I, *et al.* Midinfrared vertical-cavity surface-emitting laser. Appl. Phys. Lett. 1997; 71: 3483-3485
62. Bewley WW, Felix CL, Vurgaftman I, *et al.* Continuous-wave mid-infrared VCSELs. IEEE Photon. Technol. Lett. 1998; 10: 660-662
63. Baranov AN, Rouillard Y, Boissier G, Grech P, Gaillard S, Alibert C. Sb-based monolithic VCSEL operating near 2.2 µm at room temperature. Electron. Lett. 1998; 34: 281-282
64. Yarekha DA, Glastre G, Perona A, *et al.* High temperature GaInSbAs/GaAlSbAs quantum well singlemode continuous wave lasers emitting near 2.3 µm. Electron. Lett. 2000; 36: 537-539
65. Alford WJ, Raymond TD, Allerman AA. High power and good beam quality at 980 nm from a vertical external-cavity surface-emitting laser. J. Opt. Soc. Am. B 2002; 19: 663-666
66. Lindberg H, Strassner M, Bengtsson J, Larsson A. High-power Optically pumped 1550-nm VECSEL with a bonded silicon heat spreader. IEEE Photon. Technol. Lett. 2004; 16: 1233-1235
67. Hastie JE, Hopkins J-M, Calvez S, *et al.* 0.5-W single transverse-mode operation of an 850-nm diode-pumped surface-emitting semiconductor laser. IEEE Photon. Technol. Lett. 2003; 15: 894-896
68. Symonds C, Sagnes I, Oudar J-L, *et al.* Room temperature CW operation at 1.55 µm of a monolithic InP-based optically-pumped Vertical-External-Cavity Surface-Emitting lasers grown by MOCVD. Proc. IEEE IPRM 2003; Post-deadline paper WB1.7: 259-260
69. Cerutti L, Garnache A, Genty F, Ouvrard A, Alibert C. Low threshold, room temperature laser diode pumped Sb-based VECSEL emitting around 2.1 µm. Electron. Lett. 2002; 39: 290-292
70. Borca-Tasciuc T, Song DW, Meyer JR, *et al.* Thermal conductivity of $AlAs_{0.07}Sb_{0.93}$ and $Al_{0.9}Ga_{0.1}As_{0.07}Sb_{0.93}$ alloys and $(AlAs)_1/(AlSb)_{11}$ digital-alloy superlattices. Appl. Phys. 2002; 92: 4994-4498
71. Genty F, Cerutti L, Garnache A, *et al.* Type-I quantum-well VCSEL structure on GaSb emitting in the 2-2.5 µm range. IEE Proc. Optoelectron. 2002; 149: 22-26
72. Cerutti L, Garnache A, Ouvrard A, Genty F. High temperature continuous wave operation of Sb-based vertical external cavity surface emitting laser near 2.3 µm. J. Crystal Growth 2004; 268: 128-134
73. Ouvrard A, Garnache A, Cerutti L, Genty F, Romanini D. Single Frequency tunable Sb-based Vertical Cavity Surface Emitting Lasers emitting at 2.3µm. Proc. IEEE CLEO San Francisco 2004; paper CPDB9
74. Cerutti L, Garnache A, Ouvrard A, Garcia M, Cerda E, Genty F. 2.36 µm diode pumped VCSEL operating at room temperature in continuous wave with circular TEM_{00} output beam. Electron. Lett. 2004; 40: 869-871

Antimonide Type-II "W" Lasers

I. Vurgaftman, W.W. Bewley, C.L. Canedy, C.S. Kim, J. R. Lindle, M. Kim, and J.R. Meyer
Code 5613, Naval Research Laboratory, Washington DC 20375

1 Introduction

Mid-infrared (mid-IR) lasers have achieved significant advances in performance over the past decade. In parallel with the demonstration of improved devices, ongoing experimental studies have also more fully illuminated the physics that governs their performance, although important unresolved questions still remain. In this chapter, we overview the development and current status of one of the most promising mid-IR sources, the antimonide Type-II "W" laser [1,2]. Both optically- and electrically-pumped devices are reviewed, and the primary challenges inherent in improving their performance will be discussed.

The "W" laser takes its name from the shape of the active region's conduction-band profile, which for a prototypical structure is shown along with the valence-band profile and corresponding wavefunctions in Figure 1. Two InAs electron quantum wells (QWs) sandwich a Ga(In)Sb hole QW and are themselves surrounded by a pair of barriers, which typically employ the Al(Ga)Sb alloy although quaternaries such as AlGaAsSb and GaInAsSb have also been used. The small effective mass for electrons permits them to spread over the Ga(In)Sb hole QW, whereas the high barriers are needed to induce a quasi-two-dimensional density of states. Calculations show that the overlap of the electron wavefunction ψ_n, which has its maximum in the InAs layers, and the hole wavefunction ψ_p, which is centered on the Ga(In)Sb, is sufficient to yield interband optical matrix elements that are ≈ 70% as large as those in typical Type-I laser structures. By comparison with a simpler two-constituent InAs/Ga(In)Sb superlattice, in which the electron density of states is essentially three-dimensional, "W" lasers have a higher differential gain, stronger electrical confinement, and the potential for a greater degree of Auger suppression.

While the "W" configuration was first proposed and demonstrated for the mid-IR antimonide system illustrated in Figure 1, its enhanced wavefunction overlap and 2D density of states can be similarly advantageous when applied to other Type-II laser systems. Reports of wider-gap "W" devices include InGaAs /GaAsSbP/ InGaAs/GaAs, InGaAs/GaSbAs/InGaAs/GaAs, and GaAs/GaAsSb/GaAs/GaAsP structures for emission at 1.1-1.55 μm [3-6]. We recently designed GaAsN /GaAsSb/GaAsN/GaAs "W" active regions for 1.55 μm emission [7] and InAsN/GaAsSb/InAsN/InP structures for λ = 3-6 μm [8].

Mid-IR antimonide Type-II superlattice lasers were first demonstrated by Richard Miles and co-workers at HRL [9]. The "W" refinement was proposed by NRL in 1994 [1], and the first working optically-pumped "W" laser was achieved a year later in collaboration with the University of Houston [2]. Since then a number of milestones have been passed, including near-room-temperature CW operation for optical pumping, room-temperature pulsed operation for "W" diodes, the interband

cascade laser (ICL) with "W" active regions, mid-IR vertical-cavity surface-emitting lasers (VCSELs), and high-brightness photonic-crystal distributed-feedback (PCDFB) lasers. However, this technology remains a work in progress, and further optimization is still required before antimonide "W" devices reach their full potential.

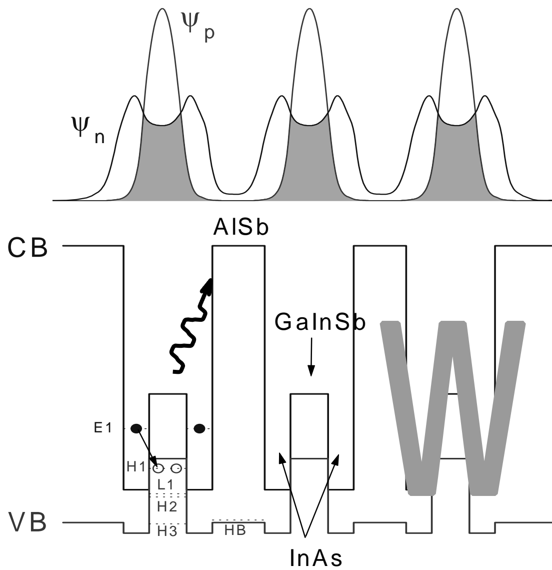


Figure 1. Band profiles for the active region of the antimonide Type-II "W" laser, with electron and hole subbands indicated by *dashed lines*. Also shown are the electron and hole wave functions and their overlap.

Currently, two leading applications are driving the need for improved mid-IR semiconductor lasers. The first is laser absorption spectroscopy, to enable the detection of trace chemicals [10]. This technique exploits the narrow mid-IR spectroscopic signatures associated with fundamental vibrational/rotational transitions in each organic or inorganic molecule (for example, CO, CO_2, NO_2, SO_2, and all hydrocarbons). While overtones at shorter wavelengths may be probed using more mature near-IR diode lasers, sensitivities to the fundamental transitions in the mid-IR tend to be 2-4 orders of magnitude stronger. However, if this Type of detection is to become widely used in such applications as atmospheric pollution monitoring, leak detection, chemical process control, and drug monitoring, it will be necessary to generate moderate CW output powers in a single spectroscopic mode, without any cryogenics.

The second major application is for infrared countermeasures (IRCM) to protect aircraft, ships, and other vehicles against heat-seeking missiles. High CW or quasi-CW powers are needed, and to save weight and power it is also highly desirable to operate at a temperature accessible with a thermoelectric (TE) cooler (above ≈ 250

K). Other potential applications include IR scene projection, IR illumination and range finding, remote chemical sensing, atmospheric monitoring and laser surgery.

In this chapter, we summarize advances in the theoretical understanding and practical development of mid-IR “W” lasers. Section 2 will discuss recent progress in their growth by molecular beam epitaxy (MBE), along with corresponding influences on the photoluminescence (PL) and lasing properties. We then devote separate sections to several important device classes: optically-pumped “W” lasers, some of which employ “integrated absorber” layers (Section 3), single-stage electrically pumped “W” lasers (Section 4), and interband cascade lasers (Section 5). Section 6 treats more elaborate configurations that were developed to improve the spectral characteristics and beam quality. Those include vertical-cavity surface-emitting lasers (VCSEL), angled-grating distributed-feedback (α-DFB) lasers, and photonic-crystal DFB (PCDFB) lasers. Finally, Section 7 reviews some critical issues that must be resolved before the ultimate goal of high-power CW operation at ambient or TE-cooler temperature can be reached.

2 Advances in the MBE Growth of “W” Laser Structures

Until quite recently, there had been surprisingly few detailed studies of the optimal growth conditions for “W” and related Type-II antimonide structures [11,12]. Antimonide MBE is far more challenging [13] in general than the relatively mature growth for GaAs- and InP-based heterostructures. The four layers of the “W” structure in Figure 1 contain three distinct materials, whose ideal growth temperatures differ by more than 100° C. Furthermore, the interface chemistry can be quite complex, since the adjacent layers share neither a common cation nor a common anion. Mixed group-V layers such as AlAsSb can be grown as either a random alloy [13] or a digital alloy [14,15] with compositional modulation along the growth axis. Variables such as growth temperature, growth rate, and substrate preparation further influence the growth quality.

Early PL studies [11,12] at NRL suggested that growth temperatures in the 400-450° C range were optimal, and that the InSb-like interface bond Type (IBT) between the InAs/GaInSb and InAs/AlSb layers was strongly preferred [16,17]. However, later studies at AFRL [18] indicated that it could be beneficial to alternate InSb- and GaAs-like IBT sequences, or to avoid forcing a particular Type so as to obtain a mixture of InSb- and GaAs-like IBT at each interface.

A recent exploration of specific sub-regions of the MBE parameter space [19] produced significant improvements in the performance of optically pumped “W” lasers. This was accomplished by systematically cross-correlating the growth conditions with such figures of merit as the surface morphology, resolution of the X-ray features, PL strength (at both low and high temperatures), PL linewidth, lasing threshold, and laser efficiency. Fortunately these properties were strongly correlated, in that whatever MBE parameters optimized any given figure of merit tended to be favorable for all. Beginning with a Te-doped *n*-GaSb substrate and a GaSb smoothing layer, a lattice-matched $AlAs_{0.08}Sb_{0.92}$ digital-alloy buffer layer was grown at 530-555°C to a thickness of either 0.5 μm for PL samples (insufficient optical confinement to support lasing) or 2.5 μm for laser samples. Next a 5-period

(PL samples) or 10-period (lasers) "W" active region, designed for emission at λ ≈ 3.7 μm at 78 K, was grown. The nominal layer thicknesses were InAs(18Å) /$GaIn_{0.3}Sb_{0.7}$(34Å)/InAs(18Å)/$AlAs_{0.106}Sb_{0.894}$(234Å) when the IBT was mixed, and InAs(15Å)/$GaIn_{0.3}Sb_{0.7}$(34Å)/InAs(15Å)/$AlAs_{0.13}Sb_{0.87}$(234Å) when the InSb-like IBT was forced. The GaSb cap layer of thickness 150 Å (for PL samples) or 5000 Å (for lasers) acted as a separate-confinement region in the case of the lasers.

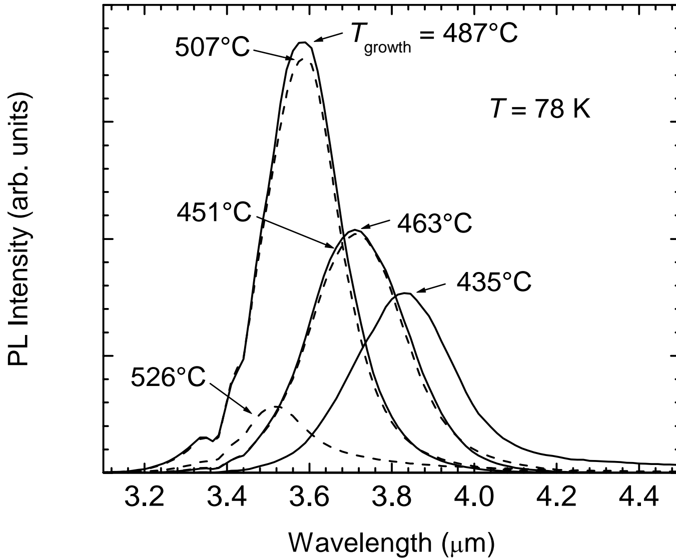


Figure 2. PL spectra at 78 K for W structures with identical five-period active regions [InAs(18 Å)/$Ga_{0.7}In_{0.3}Sb$(34 Å)/InAs(18 Å)/$AlAs_{0.106}Sb_{0.894}$(234 Å)] and identical growth conditions apart from the indicated substrate temperatures

Figure 2 shows PL spectra (T = 78 K) for a number of samples whose active regions were grown at substrate temperatures ranging from T_{growth} = 435° C to 526° C. In distinct contrast to most previous studies [11,12,20] the strongest emission occurred at relatively high T_{growth} (487-507° C), which is also consistent with other recent growths at NRL. The plot of integrated PL intensity *vs* temperature in Figure 3 indicates that the relative ordering of the samples does not change as T varies from 5 to 300 K. The decrease in PL intensity by a factor of ≈ 12 from 78 to 300 K indicates that: (i) the carrier lifetime is dominated by non-radiative recombination processes and (ii) the recombination rate depends strongly on temperature. For the relevant photoexcitation intensity, the active-region carrier density N of mid-10^{10} cm^{-2} at T = 78 K and should decrease somewhat at higher temperatures where the lifetime is shorter. Thus Auger processes can be neglected and the non-radiative lifetime, limited by Shockley-Read defects, yields N proportional to ≈ $T^{-1/2}$ for constant pump intensity. Since the radiative lifetime in a 2D system with parabolic bands is proportional to T/N (the dependence on T becomes superlinear when nonparabolicity is taken into account), we expect the ratio of non-radiative to radiative lifetimes to scale roughly as T^{-2} (dotted curve in Figure 3, which has

arbitrary magnitude). For the range T = 78-300 K, the experimental data do not deviate substantially from this dependence. The much weaker variation at the lowest T probably occurs because the lifetime for spontaneous emission then becomes comparable to the Shockley-Read lifetime.

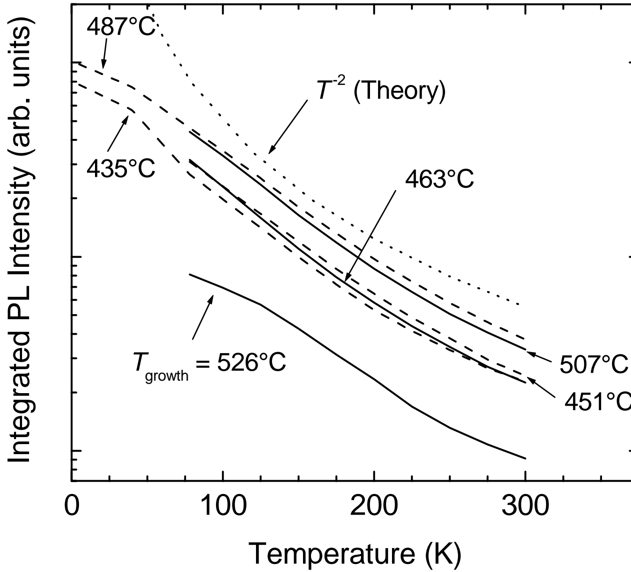


Figure 3. Integrated PL intensities *vs* temperature for the same series of wafers whose spectra are shown in Figure 2. The *dotted curve* represents the theoretical T^{-2} dependence with arbitrary scaling.

It was also found that the PL linewidths at T = 78 K were narrowest when the PL intensity was strongest, and that again the ordering of the samples remained nearly fixed at all T. The smallest full-width at half-maximum (FWHM) linewidth was 138 nm (13 meV) at 5 K for the sample grown at 487° C, which increased to 234 nm (24 meV) at 100 K. In order to accurately compare the linewidth results, we first derived the dispersion relations using an 8-band $k{\bullet}p$ calculation [21]. We then assumed inhomogeneous broadening of the spontaneous emission spectra with a Gaussian shape and linewidth δ_i (= FWHM/1.665). Comparing the theoretical spectra to the data at 5 K yielded an absolute upper bound of δ_i = 7.5 meV on the Gaussian broadening. This value is obtained by ignoring the contributions of band filling and homogeneous broadening. The tentative inclusion of band-filling effects leads to a more realistic upper bound of $\delta_i \leq 4$ meV. Lorentzian broadening may be caused by carrier-carrier collisions or scattering from short-range fluctuations at the interfaces. Taking the electron mobility (probably limited by interface roughness scattering) to be < 10,000 cm^2/Vs, the associated homogeneous broadening is by itself sufficient to account for most of the observed $\Delta\lambda$. On the other hand, we estimate that a single monolayer (ML) fluctuation in the thickness of one of the InAs wells in each period

of the active region would result in a bandgap shift of ≈30 meV. Whereas such ML fluctuations certainly do occur, from the PL data it is clear that the lateral scale is not large enough to permit the formation of bound states (one scanning tunneling microscopy (STM) study yielded a scale of 10-200 Å [17]). The energy levels for individual electrons then depend primarily on the *average* well thickness rather than local fluctuations. It should be noted that holes with their heavier mass are more susceptible to ML fluctuations in the GaInSb thickness. The inhomogeneous broadening would be up to 4-5 meV, which is consistent with the PL data.

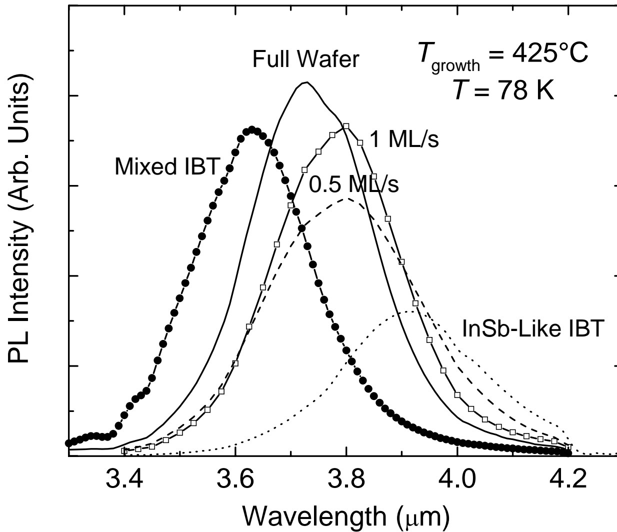


Figure 4. PL intensity for W laser samples with ten periods of the same active region as the series of PL samples. All had mixed interface bonds except the one labeled "InSb-Like IBT", and all had InAs QWs grown at 1 ML/s except the one labeled "0.5 ML/s". The samples labeled "Mixed" and "InSb-Like IBT" were grown back-to-back on the same day, as were the samples labeled "1 ML/s" and "0.5 ML/s". All were quarter wafers except the one labeled "Full Wafer".

A series of 10-period "W" laser samples was also grown, although at a sub-optimal temperature (425° C) because the PL study had not yet been performed [19]. PL spectra for five representative laser structures $T = 78$ K are illustrated in Figure 4. Here, forcing the InSb-like IBT clearly led to a reduction in the PL intensity. In fact, nearly all the samples grown with mixed IBT after this comparison displayed higher PL intensities than earlier samples grown with InSb-like IBT. Note also that the sample grown at a higher deposition rate (1 ML/s *vs* 0.5 ML/s) displayed stronger PL, although further study of that effect is required. The sample labeled "Full Wafer" combined mixed IBT with the higher growth rate. Almost without exception, the wafers displaying the highest PL intensities also exhibited the lowest lasing thresholds and highest power conversion efficiencies at all temperatures. When 2-mm-long laser cavities were pumped by short optical pulses, the threshold

pump intensities at 78 K ranged from 170 W/cm^2 for the "Full Wafer" sample to 380 W/cm^2 for the sample with InSb-like IBT. We also quantified the internal losses and carrier lifetimes in these "W" lasers by correlating thresholds measured as a function of cavity length to calculated threshold carrier densities [24]. The resulting lifetime of 24 ns at 78 K decreased to 1.4 ns at $T = 275$ K. When correlated with a calculated threshold carrier density of 1.2×10^{12} cm^{-2}, the latter implied a 2D Auger coefficient of 5×10^{-16} cm^4/s. This is consistent with previous studies of "W" lasers emitting in the 3.2-6.0 μm range [22,25,29,35] which yielded Auger coefficients between 4×10^{-16} and 3×10^{-15} cm^4/s. Figure 5 shows that the internal loss for a laser grown at the optimal temperature (500° C) ranged from only ≈ 2 cm^{-1} at 78 K to ≈ 50 cm^{-1} at 300 K [23], which compares to much higher losses (*e.g.*, 130 cm^{-1} at 300 K) in a laser grown at 425° C [19]. The next section examines output powers and spectral characteristics for optically-pumped "W" lasers in somewhat greater detail.

Structural information gained from an STM study of several of the samples in the PL series made it possible to account for the pronounced peak in device quality at T_{growth} in the 487-507° C range [25]. At too low T_{growth} (*e.g.*, 435° C), there is significant AlAs clustering in the AlAsSb digital-alloy barrier layers. On the other hand, if T_{growth} is too high (*e.g.*, 526° C), the digital alloy layers are quite distinct but there is In clustering in the GaInSb hole QWs. The best compromise, which minimizes both types of clustering, occurs at intermediate temperatures in the optimal T_{growth} range.

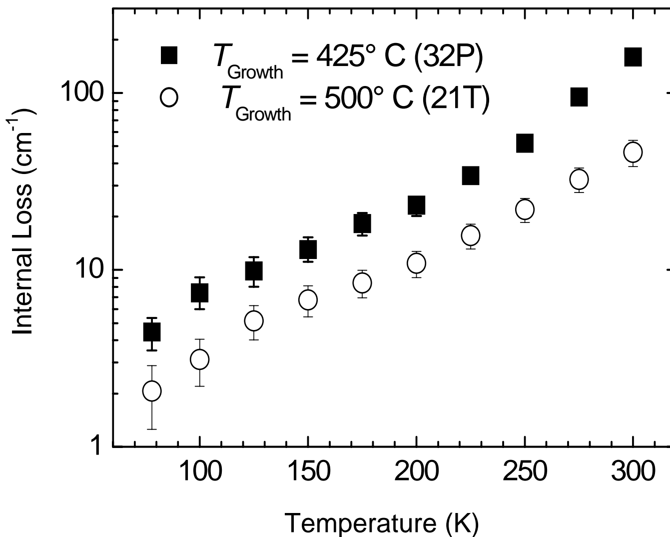


Figure 5. Internal loss *vs.* temperature, from cavity-length studies of two samples grown at 425°C on the Riber 32P MBE machine (*filled squares*) and at 500°C on the Riber 21T MBE machine (*open circles*)

3 Optically Pumped "W" Lasers

Optically pumped "W" lasers have advanced considerably since their initial demonstration in the mid-1990's [2,24-28]. The wavelength range was extended to λ = 7.3 μm several years ago [29], and most recently to λ = 9.3 μm [30]. Although the packaging is complicated by the requirement for a pump laser, the growth and fabrication are simplified compared to diodes because there is no need for intentional doping or vertical electrical transport. Optically-pumped "W" lasers therefore provide a valuable testbed for exploring the potential for high-performance Type-II antimonide devices. Of particular concern is that in order to obtain high power-conversion efficiency, the pump beam must be strongly absorbed without excessive internal loss. The optical pumping injection cavity (OPIC) [31] and integrated absorber (IA) [32] approaches to this issue will be discussed in this section.

Optically pumped "W" lasers have operated in CW mode nearly to room temperature by employing the diamond-pressure-bond (DPB) approach [33], in which the epitaxial side of the device is mechanically pressed against a very flat diamond heat sink. Figure 6 shows the light-light characteristics for a DPB-mounted laser (λ = 3.2 μm) operated CW and quasi- CW (25% duty cycle) [34]. The maximum CW power of 0.54 W increased to 0.76 W at the lower duty cycle, although the quasi- CW output fell to 0.27 W at 140 K.

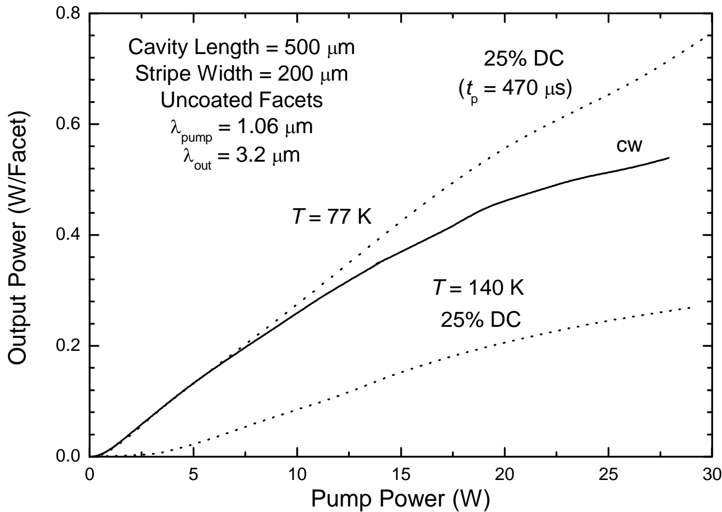


Figure **6.** Output power per facet *vs* pump power for an optically-pumped "W" laser emitting at λ = 3.2 μm. The wavelength of the pump laser is 1.06 μm, the cavity length is 0.5 mm, and the stripe width is 200 μm. Results are shown for CW operation at 78 K (*solid curve*), and quasi- CW (25% duty cycle and 270 μs pulses) operation at 78 and 140 K (*dashed curves*).

In principle, the power-conversion efficiency can be improved by reducing the photon decrement, *i.e.*, employing a longer-wavelength pump source, and by reducing the number of active quantum wells (N_{QW}) so as to minimize free carrier absorption. However, by themselves each of these measures leads to a weaker absorbance of the pump beam and hence a net efficiency decrease. The OPIC approach circumvents this difficulty by forming an etalon cavity for the pump beam along the vertical axis [31], even though the mid-IR laser emission is from the edge. The absorbance is enhanced by multiple passes of the pump beam through the active region, and the GaSb/$AlAs_{0.08}Sb_{0.92}$ distributed Bragg reflectors (DBR) also act as optical cladding layers for the mid-IR lasing mode. With only ten active QWs (compared to N_{QW} = 40-80 in most of the earlier "W" devices), absorbances of 57-71% were measured when the etalon resonance was tuned to 2.1 μm for pumping by a Ho:YAG laser. At 100 K, power conversion efficiencies per facet were as high as 11.4%, which compares to 1.8% for conventional "W" lasers [35]. The OPIC laser's efficiency also degraded much more slowly with increasing temperature, *e.g.*, it still exceeded 7% per facet at 220 K. Recent measurements with a wavelength-tunable source indicate a strong efficiency peak and sharp minimum in the threshold pump intensity (I_{th}) near the etalon resonance [36].

An alternative approach to enhancing the efficiency with optical pumping is the IA configuration, which for antimonide Type-II devices was pioneered by MIT Lincoln Laboratory [32] and further developed by AFRL [37]. In this approach, the Al(As)Sb barrier in the prototypical "W" structure of Figure 1 is replaced by a thick (typically ≈ 1000 Å) layer of $Ga_{1-x}In_xAs_ySb_{1-y}$, where $x \approx$ 15-20% and y is adjusted so as to assure strain compensation to GaSb[20,37,38]. The IA layers assure strong absorption of the relatively-long-wavelength pump beam ($\lambda_{pump} \approx$ 1.8-1.9 μm), while again allowing N_{QW} to be reduced, typically to the 6-10 range. Because a high fraction of the incident radiation is absorbed [20] and the internal losses are low, external quantum efficiencies of up to ≈ 24% have been reported for such "W" IA lasers [30,37]. Because the OPIC and IA configurations are fully compatible with one another, one could reduce N_{QW} even further by combining the two.

The IA approach has also made it possible to develop *Al-free* "W" lasers [20,38]. Whereas optically-pumped "W" devices usually employ $AlAs_{0.08}Sb_{0.92}$ for the bottom optical clad, the refractive index of $Ga_{1-x}In_xAs_ySb_{1-y}$ with the appropriate x tends to be slightly higher ($\Delta n \approx 0.03$) than that of the GaSb substrate [38]. The substrate and a second GaSb layer grown epitaxially on top then provide the two claddings. The reduced optical confinement in the active region narrows the far-field divergence angle along the *fast-axis* (growth direction) to 15° FWHM for a 175-μm-wide stripe and a pump intensity of ≈ 7 × I_{th} [20]. The corresponding *slow-axis* (in the plane) divergence was 6° FWHM. Pumped with 32 μs pulses at 1% duty cycle, peak output powers of at least 5-7.5 W have been generated at $T \approx$ 80 K [20,30], with a corresponding external quantum efficiency of ≈ 60% and power-conversion efficiency of 28%. Hakki-Paoli and cavity-length measurements determined a low internal loss of ≈3 cm^{-1} at 78 K, with no apparent increase up to at least 120 K [39]. A disadvantage of the "W" IA lasers demonstrated to date is that they are highly efficient only at relatively low temperatures [39], probably due to a limited offset between the valence band maximum in the GaInAsSb IA layer and the active H1 subband in the GaInSb QW.

Very little emphasis has been placed over the past decade on developing optically pumped "W" lasers with narrow spectral linewidths. Most devices supported multi-mode lasing because pulsed excitation was used to pump wide stripes that were Gaussian in both the temporal and spatial domains. Many of the spectral envelopes have in fact been even broader ($\Delta\lambda \approx$ 40-150 nm) than expected on the basis of simulations. The PL linewidths discussed in the previous section imply low-temperature lasing envelopes no wider than a few nm, as long as smooth gain spectra are assumed. The dashed curves in Figure 7 illustrate the rather broad spectral envelopes for the "Mixed-IBT" laser from Figure 4, pumped at 20 × I_{th} at a series of temperatures. While further investigation will be required before the linewidths are fully understood, well-to-well fluctuations along the growth axis in the 10-QW devices may be ruled out since a single-QW "W" laser with similar layer thicknesses displayed nearly the same envelope width. Another striking feature in Figure 7 is that each dashed spectrum contains multiple peaks, which are spaced in a roughly periodic manner. This appears to be caused by leakage of the optical mode into the GaSb substrate, which allows reflections from the back surface to destructively interfere with the cavity mode (thereby inducing oscillations in the loss spectrum) [40,41]. The obvious approach to eliminating this interference is to thicken the bottom cladding layer. The solid curves in Figure 7 are spectra for the same laser structure, grown on a different MBE machine at a more optimal temperature (T_{growth} = 500° C) and with the AlAsSb bottom clad thickness increased from 2.5 to 3.5 µm [41]. We find that the multiple peaks are eliminated due to the exponential reduction of the mode leakage, and that the envelope widths have also narrowed considerably.

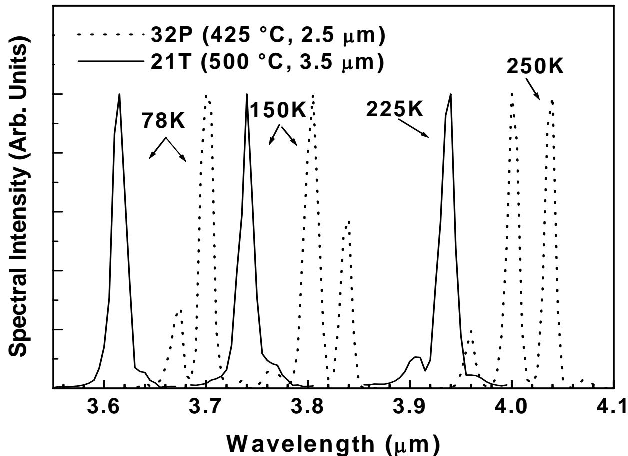


Figure **7.** Lasing spectra at three different temperatures for a sample grown at 425°C with a 2.5 µm cladding layer on the 32P machine, and for a sample grown at 500°C with a 3.5 µm clad on the 21T machine. While all of the lasers grown with the thinner cladding layers displayed broad spectral envelopes with multiple peaks, thickening the clad eliminated the modulation.

Another optically pumped InAs/GaSb/InAs/AlSb "W" structure ($\lambda \approx 3.4$ µm) operated to 320 K, with a characteristic temperature (T_0) of 73.5 K [42]. Auger coefficients were also reported. An $In_{0.85}Ga_{0.15}As/GaSb/In_{0.85}Ga_{0.15}As/AlSb$ "W" laser ($\lambda = 3.5$ µm) was recently grown on an InAs substrate [43]. That device with $N_{QW} = 40$ operated in pulsed mode up to $T_{max} = 265$ K, and displayed a CW differential power conversion efficiency per facet of 4.6% at 30 K.

4 Single-stage "W" Diode Lasers

Our focus in this section will be on the progress of single-stage (non-cascade) "W" diode lasers, and on the challenges presented by electrical pumping in comparison with optical pumping. A first concern is carrier injection, which for holes would be blocked by the combined thicknesses of the AlAsSb barrier and InAs electron QW if one were to employ the structure in Figure 1. "W" diode designs [44] have generally overcome that difficulty by employing quaternary AlGaAsSb barriers (with reduced height), and relying on tunneling by light rather than heavy holes. By bringing the L1 subband to within 100 meV of the valence-band maximum (H1), enough light-hole states are thermally occupied to provide adequate transport. The original design also placed nominally-undoped 0.6-µm-thick AlGaAsSb SCH layers on each side of the active region, to reduce the overlap of the lasing mode with the doped $Al_{0.9}Ga_{0.1}As_{0.07}Sb_{0.93}$ cladding layers. Since the free-hole absorption cross section tends to substantially exceed that for electrons, it is particularly important to minimize the mode overlap with the *p*-cladding.

"W" diode lasers of this design were grown by MBE at Sarnoff and characterized at NRL. A device with $N_{QW} = 10$ operated in pulsed mode to 310 K (where $\lambda \approx 3.3$ µm) [45], which was the first room-temperature lasing of an interband III-V diode emitting beyond 3 µm. The best CW results were obtained for a similar 5-QW structure [44]. A Hakki-Paoli characterization determined that the internal loss in the latter device was nearly constant at ≈ 19 cm^{-1} for the range $T =$ 80-160 K [46]. The electrical characteristics of both 5-QW and 10-QW devices were far from optimal, however, in that the turn-on voltages and series resistances were considerably higher than expected, especially at lower temperatures.

More recent "W" diode lasers, also grown at Sarnoff and characterized at NRL, were found to have improved electrical characteristics [47]. The 3.0-µm-thick bottom optical clad was composed of $Al_{0.5}Ga_{0.5}As_{0.04}Sb_{0.96}$ rather than $Al_{0.9}Ga_{0.1}As_{0.07}Sb_{0.93}$, in order to reduce the high barrier to electron injection from the GaSb substrate. To improve further the transport, a GaSb separate-confinement region replaced the AlGaAsSb on the *n*-side, although unintentional *p*-doping of the GaSb may have inhibited the electron flow. As a result, both the turn-on voltage and differential resistivity for 5-QW lasers with these modifications decreased by ≈ 2. The device emitted CW output powers of 198 mW at 78 K and 40 mW at 160 K, with $T_{max}^{CW} = 195$ K (where $\lambda = 3.56$ µm).

Two other recent "W" diode lasers were grown by MBE at Maxion [48] and NRL [49] (both were characterized at NRL). The sample grown at Maxion employed a Si-doped InAs/AlSb superlattice for the *n*-clad and a Si-doped

InAs/GaInSb superlattice for the *n*-SCH. It also included additional "transition layers" between the various regions, which suppressed the formation of electrostatic barriers by smoothing the voltage steps associated with band offsets. The resulting electrical properties were far superior to those of the previous devices, and the threshold voltage was only slightly higher than that required to overcome the energy gap of ≈ 350 meV in the active region. Figure 8 plots temperature-dependent threshold current densities in pulsed mode for the three recent diodes [47-49]. While the 5-QW sample grown at NRL (λ = 3.4 μm at T_{max}) has the lowest threshold at low T (a sister device had an even lower threshold of only 24 A/cm^2 at 78 K), it also had the lowest characteristic temperature (T_0 = 38 K) and the lowest maximum operating temperature (T_{max} = 190 K). That the Maxion sample (λ = 4.02 μm at T_{max}) had a higher threshold at low temperatures is not surprising, since it had twice as many active QWs (ten). That device had a maximum operating temperature of 315 K, which represents the first room-temperature lasing by a single-stage III-V diode at λ > 4 μm.

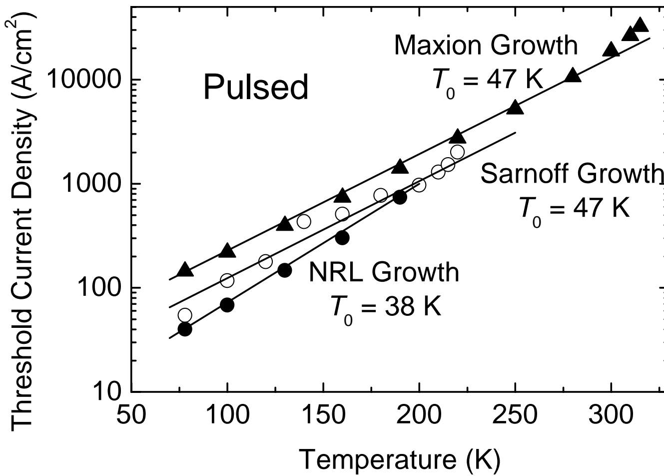


Figure 8. Pulsed threshold current densities *vs* temperature for three recent "W" diode lasers grown at NRL (*filled circles*), Sarnoff (*open circles*) and Maxion (*triangles*). The *lines* are fits corresponding to constant T_0. Pulses were typically 200 ns in duration with a 200 Hz repetition rate.

Figure 9 plots temperature-dependent slope efficiencies for the same three devices operating in pulsed (NRL and Maxion) or CW (Sarnoff) mode. While the NRL and Sarnoff devices had one high-reflectivity (HR, R = 90–95%) facet coating and one anti-reflection (AR, R = 5%) coating, the Maxion device is penalized by having no facet coatings and also a longer cavity (2 mm). The slope efficiency of 186 mW/A for the NRL device at 78 K corresponds to a high external quantum efficiency of 49%. However, the slope degrades relatively rapidly with increasing temperature, for reasons that are still under investigation.

Researchers at the University of Montpellier have grown InAsSb/InAsP/ InAsSb/InAsPSb “W” diode laser structures by metalorganic vapor phase epitaxy (MOVPE) on InAs substrates. The asymmetric cladding layers used InAsPSb on the *n*-side and InPSb on the *p*-side. For pulsed operation, a preliminary 5-QW structure exhibited a threshold current density of 120 A/cm^2 at 90 K, a characteristic temperature of 35 K, and a maximum operating temperature of 135 K, where λ = 3.3 µm [50]. The gain and threshold properties were modeled in some detail [51].

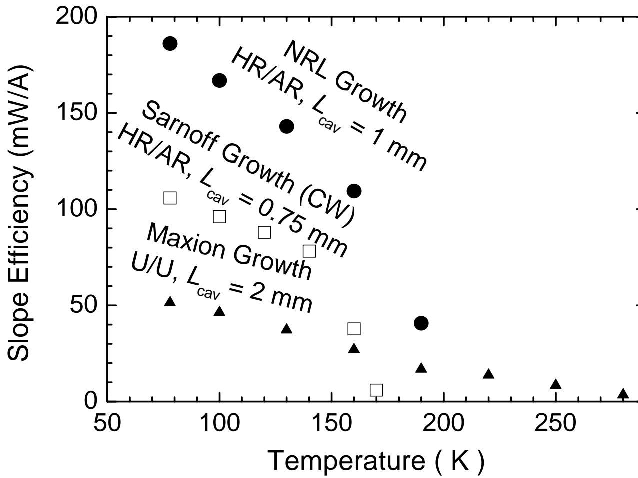


Figure 9. Slope efficiencies *vs* temperature for the three diodes from Figure 8. The NRL (cavity length 1 mm) and Sarnoff (cavity length 0.75 mm) lasers had HR and AR coatings, while the Maxion device (cavity length 2 mm) had uncoated facets.

5 Interband Cascade “W” Lasers

While conventional diode lasers generate at most one photon for each electron-hole pair that is injected, cascade devices can have much higher slope efficiencies because they generate an additional photon at each step of a staircase-like active region. The best-known example is the quantum cascade laser (QCL) [52], which achieves stimulated emission due to a population inversion between conduction subbands. A related concept is the interband cascade laser (ICL) [53], which incorporates multiple stages like a QCL, but nonetheless employs interband lasing transitions like a conventional diode.

A key distinction between ICLs and conventional diode lasers is their contrasting apportionment of the electrical input power ($P_{in} = IV$) between current and voltage. The active QWs in a conventional single-stage diode are effectively connected in parallel, hence in lowest order the current required to reach population inversion scales as N_{QW} while the threshold voltage is independent of N_{QW} (V_{th} must exceed E_g/e, although other potential barriers and the series resistance can also contribute as discussed in the preceding section). On the other hand, the QWs in an ICL are

connected in series, making I_{th} nominally independent of N_{QW} while $V_{th} \propto N_{QW}$. Although P_{in} is proportional to N_{QW} in either case, the ICL tends to operate with lower current whereas the conventional diode requires lower voltage. The ICL arrangement may be especially favorable if the series resistance would otherwise account for a substantial fraction of the total voltage drop.

Figure 10 shows conduction- and valence-band profiles, along with calculated energy levels, for an early ICL design [54]. Lasing occurs when the electrons make Type-II optical transitions between conduction states in the 19-Å InAs QW and valence states in the 40-Å $Ga_{0.7}In_{0.3}Sb$ QW. Electrons in the GaInSb valence band next tunnel to a second hole QW (60Å GaSb) [55], from which they are elastically scattered to conduction states in the 120-Å InAs QW. By taking advantage of the unique antimonide Type-II band alignment, electrons recycle from the valence band of one stage into the conduction band of the next injection region [53], which usually consists of a digitally-graded InAs/Al(In)Sb superlattice. This critical feature makes multi-stage operation (and slope efficiencies greater than unity [56-61]) possible for devices that typically employ 18-35 stages. While early designs such as that in Figure 10 featured a single InAs electron QW, more recent ICLs have routinely incorporated the "W" geometry to increase the electron-hole wavefunction overlap and maximize the differential gain [61].

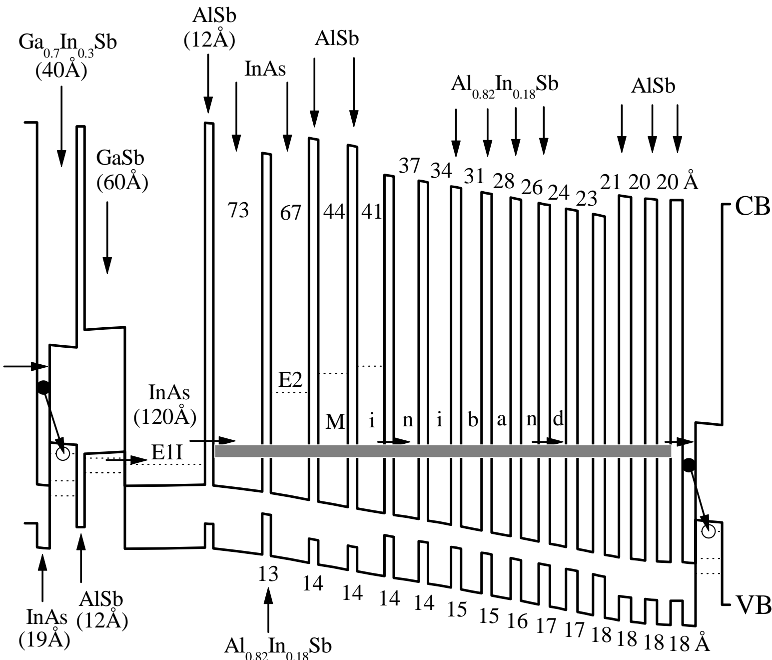


Figure 10. One active period of an early ICL design, under applied bias and including internal electrostatic fields due to charge transfer. The relevant subbands are indicated by *dashed lines.*

The maximum operating temperature for ICLs has increased steadily since the first demonstration of pulsed lasing in 1997 by University of Houston and Sandia (T_{max} = 170 K) [63]. Subsequently, a collaboration between NRL and U. Houston produced pulsed devices operating to T_{max} = 225 K and then 286 K [56,64]. The latter design included double ("W") electron QWs for enhanced gain and three hole QWs rather than two for reduced leakage. Room-temperature pulsed operation was first achieved by the Maxion group ($\lambda \approx 3.5$ μm) [65], who at $T = 80$ K achieved an external differential quantum efficiency exceeding 600%, peak power of ≈ 6 W per facet, and threshold current density as low as 12 A/cm^2 [66]. The characteristic temperature was $T_0 \approx 41$ K for T = 80-200 K. Jet Propulsion Laboratory (JPL) subsequently demonstrated room-temperature pulsed operation at $\lambda = 4.1$ μm [61].

The CW performance has also improved considerably, even though all of the ICLs reported to date have been mounted epitaxial-side-up. Following the first CW operation by the Army Research Lab to T_{max} = 70 K in 1999 [67] and T_{max}^{CW} = 142 K in 2000 [68], recent Maxion "W"-ICL devices pushed T_{max}^{CW} to 214 K [69]. The thermal resistance was reduced by electro-plating a thick (3 μm) layer of gold on top of an evaporated contact to the deeply-etched mesa, which was only 4 μm wide. The same team reported CW wall-plug efficiencies of up to 17% at 80 K [69]. Employing a 110-μm-wide current stripe, JPL obtained 140 mW of CW output power per facet at 80 K, combined with a threshold current density of only 8.9 A/cm^2 [70]. They later reported T_{max}^{CW} = 217 K [71]. For ICLs emitting at much longer wavelengths in the λ = 5.1-5.6 μm range, JPL obtained a maximum CW operating temperature of 165 K [71,73]. Both Maxion and JPL have demonstrated DFB ICLs with one-dimensional line gratings [74,75]. The Maxion device displayed ≈ 0.1 nm/K red-shift with temperature and >30 dB side-mode suppression ratio. The JPL device ($\lambda = 3.3$ μm) had a similar SMSR, and operated CW to 175 K.

6 "W" VCSELs and PCDFB Lasers

The lasers reviewed in this section feature a variety of optical cavity configurations that differ from those of conventional edge emitters. Historically, the first to fall in this class was the "W" vertical-cavity surface-emitting laser (VCSEL) [76], whose advantages in the mid-IR are similar to those documented in the literature for shorter-wavelength VCSELs. Intrinsic features include single longitudinal mode (due to the short cavity), circular and relatively low-divergence output, and very low lasing threshold if the lateral dimensions are also small. NRL and HRL realized the first III-V mid-IR VCSELs by growing a 12.5-period GaSb/$AlAs_{0.08}Sb_{0.92}$ DBR mirror (95% reflectivity) below the active region and depositing a high-reflectivity dielectric coating (99% reflectivity) on top. Transmission experiments indicated that the cavity mode nearly coincided with the "W" active region's low-temperature gain peak near $\lambda \approx 2.9$ μm. For optical pumping at λ_{pump} = 1.06 μm, the maximum operating temperatures were 280 K for pulsed operation and 160 K CW [77]. At a 6-μm-diameter spot and $T = 78$ K, only 4 mW of CW pump power was required to reach threshold.

Since both 1D-DFB and VCSEL structures produce single-mode emission only when the device areas are small, maximum output powers are limited. More

sophisticated geometries are required to circumvent those limitations. One approach is to pattern a 1D line grating that is tilted with respect to the facets, so as to form an angled-grating distributed feedback (α-DFB) device [78,79]. In the ray picture, the α-DFB lasing mode may be viewed as diffracting back and forth in a zigzag pattern. Upon reaching the facet, only those components of the optical wave that impinge nearly at normal incidence have appreciable feedback. The result is very low angular divergence of the output as well as spectral selectivity. The first mid-IR α-DFB lasers were optically-pumped "W" devices that produced near-diffraction-limited output from stripes as wide as 50 μm [80,81].

Optical coherence is enhanced and filamentation suppressed in α-DFB devices because non-collinear propagation of the diffracted beam strongly couples regions of the stripe that are widely-separated spatially. Unfortunately, α-DFBs typically fail to preserve spectral purity under pulsed conditions [79], since the gain margin separating neighboring modes is up to an order of magnitude weaker than in 1D DFB lasers. Even for CW operation, the grating angle with the facets must be highly uniform. These deficiencies may be overcome by generalizing the linear grating geometry of the α-DFB to a 2D pattern defined on a rectangular photonic-crystal (PC) lattice [82], which is similarly tilted with respect to the facets [83,84]. Figure 11 shows a scanning electron micrograph (SEM) top view of a PCDFB structure. Its spectral and spatial selectivities are enhanced by the addition of two diffraction processes that provide intracavity feedback, $\boldsymbol{P}_1 \rightarrow \boldsymbol{P}_{-1}$ and $\boldsymbol{P}_1 \rightarrow \boldsymbol{P}_{-2}$, besides the $\boldsymbol{P}_1 \rightarrow \boldsymbol{P}_2$ process that guides the light along the tilted stripe as in the α-DFB. Thus the PCDFB laser no longer requires facets to operate, although they are still employed to extract the light from the cavity. A time-domain Fourier-transform (TDFT) formalism, developed to model the performance of PCDFB lasers [84], found that both first-order and second-order (with both of the rectangular lattice's periods doubled) gratings can yield similar performance. This is provided that the tilt angle, also equal to the inverse tangent of the ratio of the periods, is ≈ 16° for the second-order grating and ≈ 20° for the first-order grating. The fabrication is considerably less demanding for PCDFBs with the second-order grating, which has both longer periods and a more moderate short-period duty cycle of ≈ 50% (compared to 90-95% for the first-order grating). Optically-pumped "W" PCDFB lasers have been fabricated with both first-order [87,88] and second-order gratings [85,86,89] etched into top GaSb separate confinement regions.

Apart from one initial study of InP-based devices [82], all subsequent PCDFB lasers have employed antimonide "W" active regions emitting in the mid-IR [85-89]. It should be noted that this is an especially challenging system from which to obtain high brightness, because the internal loss and linewidth enhancement factor both tend to be relatively large. The product of those two parameters represents the most important figure of merit for projecting α-DFB and PCDFB performance [84].

The best overall efficiencies and beam qualities were obtained for cavity lengths of ≈ 2 mm. One second-order structure emitted two distinct spectral lines [85,86], which were attributable to distinct modes from different regions of the Brillouin zone that were accidentally degenerate because the tilt angle was too large. Initial structures with first-order gratings were non-optimal because the grating wavelength λ_g did not coincide with the peak of the gain spectrum [87]. However, the

performance improved drastically following laser ablation of areas on both sides of the tilted pump stripe, to eliminate parasitic Fabry-Perot modes [88].

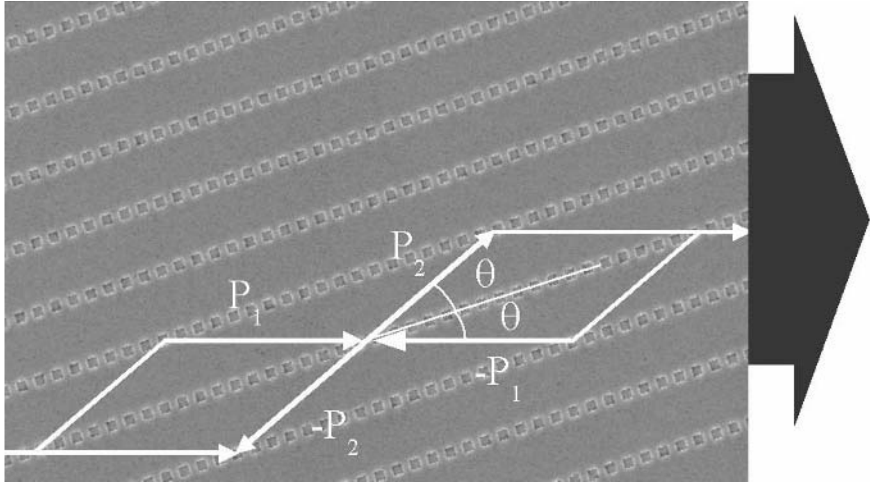


Figure 11. Schematic of the major propagation directions for a rectangular-lattice PCDFB laser, overlaid with a scanning electron micrograph of the etched 2D grating

For a recent device with an optimized second-order grating, temperature tuning of the energy gap was sufficient to bring the peak of the gain spectrum into resonance with λ_g. Spectral characteristics at T = 145 K and a Gaussian pump stripewidth of w = 100 μm ($1/e^2$ half-width) are shown in Figure 12. For pumping at three times the lasing threshold (solid curve), the FWHM spectral linewidth of 2.7 nm is much narrower than the FWHM of 15 nm from a Fabry-Perot laser fabricated from the same material. That the spectrum at $I = 10 \times I_{th}$ (dashed curve) split into several peaks, spanning an envelope of ≈ 25 nm, may have resulted in part from the shortness of the pump pulses. Figure 13 shows angular far-field intensity profiles at T = 145 K and $I = 10 \times I_{th}$ for several stripewidths. The output at w = 100 μm is basically diffraction-limited, and even the 350-μm and 600-μm stripes produce single-lobed beams with FWHM divergence angles of only 0.8° and 1.2°, respectively. These profiles show no evidence for parasitic Fabry-Perot oscillation.

In order to quantify the improvement, the experimental far-field data were correlated with theoretical near-field profiles from the TDFT simulations [83,84]. We then used the definition in [81] to calculate etendues as a function of stripewidth. Figure 14 plots the results, in units of the diffraction limit (DL), for a variety of mid-IR α-DFB and PCDFB devices, all of which were pumped at $I = 10 \times I_{th}$ [80,87,88]. The PCDFB lasers substantially improved on the α-DFB beams (open triangles), which in turn exhibited far smaller etendue than typical unpatterned Fabry-Perot lasers (including one fabricated from the same wafer). Unfortunately, beam qualities for the early non-optimized PCDFB structures (open boxes and circles) progressively degraded with increasing stripewidth, until at w = 400 μm the best result was 33 × DL. However, by eliminating parasitic Fabry-Perot emission, the non-optimized device with laser-ablation confinement (filled triangles) remained

essentially diffraction-limited to w = 100 μm, and was no worse than 5.7 × DL at w = 400 μm. Etendues for the most optimized device (filled circles) are even better, with the etendue not exceeding 8 × DL at w = 600 μm. These dramatically-improved beam qualities, which are far superior to any previously reported for wide-stripe mid-IR lasers, are accompanied by relatively modest penalty in the differential power conversion efficiencies. For example, at w = 350 μm the efficiency was 45% of that for an unpatterned Fabry-Perot laser from the same bar. Theory projects maximum efficiency ratios of 50-60%.

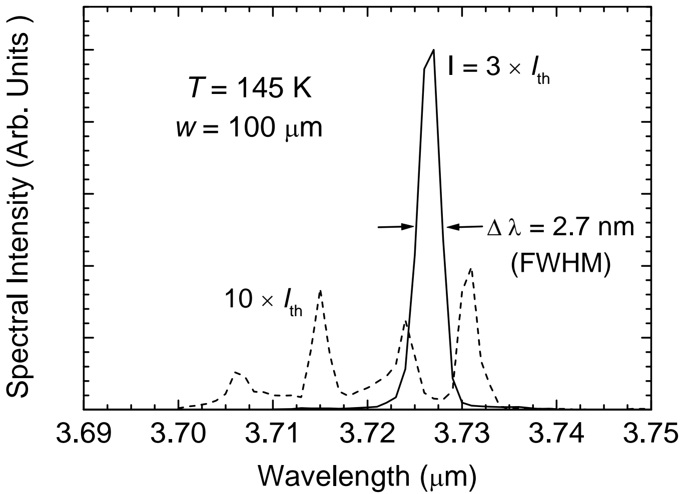


Figure 12. Emission spectra in pulsed mode for a second-order PCDFB laser at optical pumping intensities of three and ten times the lasing threshold. The operating temperature is 145 K, and the Gaussian pump stripe width is w = 100 μm ($1/e^2$ half-width).

A recent theoretical study considered the potential performance of *surface-emitting* (SE) PCDFB lasers [90]. In contrast to edge emitters, SE PCDFB lasers can produce a nearly circular output beam and do not require any facets for light extraction. Highly-efficient (up to 60% quantum efficiency if reflectors are built into the waveguide structure) emission into a near-DL beam should be possible if the hexagonal (TE polarization) or square (TM polarization) photonic-crystal lattice is optimized. The optimization requires a relatively weak SE diffractive coupling and self-pumping to suppress undesirable modes. These lasers profit from a moderate (rather than low) linewidth enhancement factor, in contrast to edge emitters. Mid-IR “W” SE PCDFB laser experiments are currently underway to verify these projections.

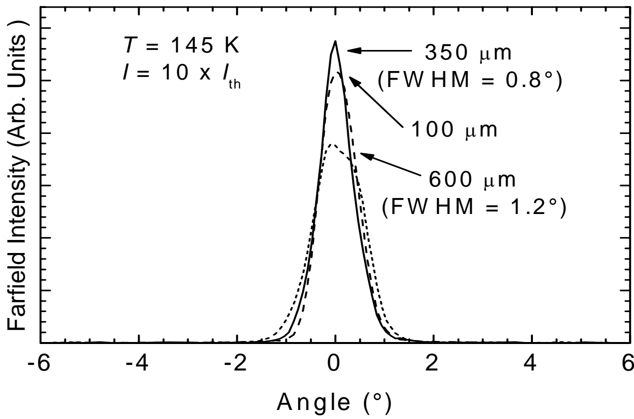


Figure 13. Far-field intensity profiles for three stripe widths of the same second-order PCDFB laser (T = 145 K). Optical pumping was at ten times the lasing threshold.

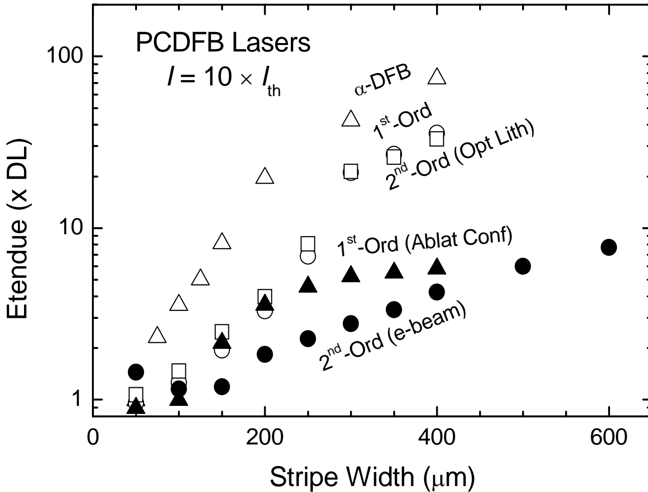


Figure 14. Etendue (in units of the diffraction limit) *vs* stripe width for the α-DFB laser from [80] (open triangles), the second-order PCDFB laser from [85] (boxes), the first-order PCDFB laser from [87] (*open circles*), the first-order PCDFB laser with laser-ablation confinement from [88] (*filled triangles*), and the optimized second-order PCDFB laser from [89] (*filled circles*). All lasers were operated at ten times the lasing threshold.

7 Critical Issues in Improving the Performance of "W" Lasers

It should be clear from the preceding discussions that "W" lasers have not yet progressed beyond an intermediate stage of development, with their performance still far from optimized. Three broad classes of issues continue to require special attention: (1) reduction of the lasing threshold and increase of the wall-plug efficiency at high temperatures; (2) enhancement of the thermal dissipation; and (3) optimization of the electrical transport in "W" diodes. While a significant breakthrough in any of these areas would relax the requirements for the others, ambient-temperature CW operation will likely occur only after the series resistance, internal loss, and thermal resistance are all reduced. In this section, we consider these issues in greater detail.

The high threshold current densities observed at thermoelectric-cooler temperatures or above are attributable to non-radiative Auger recombination. Which of the two main classes of Auger process dominates under given conditions has, however, still not been determined unambiguously. In CCCH events, the energy gained from electron-hole recombination promotes another electron to a position roughly one bandgap higher in the conduction band. Alternatively, the recombination energy can excite a second hole to a lower valence band (most often the split-off band) or subband (H2, H3, L1, L2, *etc.*), whose energy separation from the valence band maximum is nearly resonant with the bandgap. A recent study [91] considered how disorder and phonon scattering modify these mechanisms.

When CCCH Auger processes dominate the threshold current, it should be advantageous to increase the In fraction x in the active region's $Ga_{1-x}In_xSb$ hole QWs, so as to increase the compressive strain. This reduces the hole mass near the top of the valence band and thereby maximizes the CCCH activation energy [92,93]. However, many of the current "W" active regions already approach the strain limit, since the critical thickness becomes thinner than the hole QW when x exceeds ≈ 35-40%.

On the other hand, if multi-hole Auger processes dominate there may still be considerable room for improvement. A Type-II antimonide active region effectively removes the possibility of coincidence between the bandgap and the split-off gap, although resonances with other valence subbands may play an important role [94,95]. While a compilation of experimental Type-II InAs/GaInSb Auger coefficients at T = 300 K suggests a significant reduction (by a factor of 4-6) compared to Type-I materials with the same energy gaps [22], the improvement is relatively uniform over a wide range of wavelengths and material structures. This would seem to argue for dominance by the CCCH mechanism, since intervalence resonances should have varying influence from sample to sample. On the other hand, multi-hole processes may lose some of their sensitivity to the valence subband alignments at room temperature, since thermal broadening can cause several classes of transitions to contribute simultaneously [96]. Recent pump-probe measurements on four different Type-II antimonide samples with the same energy gap found that the Auger lifetimes at fixed carrier density (close to the gain threshold of 10^{18} cm^{-3}) differed by more than a factor of 2 [97]. The structure with the lowest theoretical

Auger rate displayed the largest discrepancy (nearly an order of magnitude) between theory and experiment.

While the path toward substantially further suppressing the Auger coefficients in "W" active regions is not yet clear, the Auger lifetime also depends strongly on carrier density. In the non-degenerate limit: $\tau_A \propto 1/N^2$, while the variation saturates to a $1/N$ dependence in the opposite limit of highly-degenerate populations [95,98]. In turn, the carrier density required to reach threshold is sensitive to the internal loss α_{int}, whose experimental dependence on temperature was already discussed in connection with Figure 5. Although the loss is favorable at low T, we must ask whether there is any way to suppress its drastic increase with temperature.

We first consider potential loss mechanisms in optically pumped "W" lasers. Photon scattering from random imperfections in the waveguide can contribute if the roughness length scale is comparable to the wavelength divided by the modal index. Since this contribution should be nearly independent of T and the threshold carrier density, its magnitude can be no greater than a few cm^{-1}. A similar limit can be applied to free carrier absorption and other loss mechanisms in the optical cladding layers, whose properties should not depend significantly on pump intensity. The remainder of the loss may therefore be attributed to free-carrier absorption in the active region.

While the Drude and defect/phonon-assisted contributions can be particularly troublesome in the mid-IR, since they scale approximately as λ^2, an even larger loss most likely results from valence intersubband transitions [94,99,100]. According to the usual TE-polarization selection rules [101,102], transitions between H1 and any lower-lying heavy-hole subband vanish at zero in-plane wavevector. However, interband mixing relaxes this rule at finite k, so that the matrix elements for such processes as H1→H3 have substantial peaks at larger k (*e.g.,* ≈ 0.03 Å^{-1}). Therefore, the cross sections for intra-heavy-hole transitions tend to increase substantially with temperature. Transitions from H1 to L1, L3, L5,…, which do not involve a change of parity, similarly become allowed only at finite k. On the other hand, transitions from H1 to L2, L4, L6, …, which *do* involve a change of parity, are strong at all temperatures since they are allowed even at $k = 0$. We must be especially careful to avoid the latter class of transitions.

Unfortunately, at present the band structure computation techniques, knowledge of the various band parameters, and properties of the grown structures (*e.g.,* the degree of interdiffusion) are all too uncertain to allow the lower valence-subband positions in "W" QWs to be predicted with sufficient accuracy and reliability. For example, modifying the $k \bullet p$ formalism to correct for IBT symmetry produces enough shift of the valence subbands to completely alter the theoretical loss projections for a given structure [103,104]. Other computation techniques, such as the pseudopotential method, yield further distinct sets of theoretical projections [105-107]. Factoring in also the rather imprecise knowledge of valence-band effective masses in some of the antimonide materials, it is clear that detailed α_{int} predictions are only possible if the transition broadening is assumed to exceed the existing uncertainty. The broadening results reported above do not support this statement.

Nevertheless, analysis of the experimental data allows us to draw several conclusions concerning internal losses in "W" structures. First, some but not all of

the temperature dependence in Figure 5 may be attributed to the increasing threshold carrier density with T. Second, results to date for a variety of "W" designs suggest somewhat weaker sensitivity to intervalence absorption than is expected theoretically. Third, the internal loss at both low and high T appears to correlate with growth quality (as for the two samples in Figure 5), suggesting that the Drude contribution may play a role after all. And finally, no matter which mechanism is responsible for the strong temperature increase of α_{int} in Figure 5, reducing the number of QWs in the active region should counteract it so long as there is sufficient optical gain to reach threshold at higher T.

For "W" diode lasers, minimizing the optical-mode overlap with the heavily doped cladding layers must be added to the considerations presented above. Since the absorption cross section for holes is believed to be much larger than that for electrons (due to the much lower hole mobility), it is particularly important to limit the mode overlap with the p cladding to no more than a few percent. Fortunately, these guidelines can generally be realized by incorporating separate-confinement layers into the laser waveguide [44,45,47,108]. Because of the large number of active periods, loss minimization is especially important for interband cascade lasers [109].

Section 6 briefly discussed the influence of the linewidth enhancement factor on the optical coherence and beam quality attainable for a given cavity configuration. Several theoretical [1,110] and experimental [111,112] determinations of this parameter have been reported for optically-pumped, diode, and interband cascade structures.

Since the tendency toward high lasing threshold is a serious concern for "W" diodes operating at high T, the heat dissipation requirements are stringent. One approach to reducing the thermal resistance was discussed above in connection with ICLs: gold electroplating of narrow ($\leq$ 10 μm) epi-side-up mounted devices [69]. However, some applications will require higher powers than are likely to be attainable from narrow current stripes. An alternative is to employ epi-side-down mounting with a broad stripe, for which the heat flow is almost entirely one-dimensional. If a PCDFB grating is fabricated to maintain high beam quality and/or spectral purity, the active region is separated from the heat sink by a GaSb grating layer and a low-index top cladding layer. While the ideal top clad would be an overgrown semiconductor such as Al(As)Sb or InP, in the following simulations we assume the less favorable case of a deposited layer of alumina. The net 1D specific thermal resistance is then estimated to be $\approx$ 2.5 K-cm^2/kW.

It was mentioned above that ICL active regions should outperform single-stage "W" diodes whenever the series resistance becomes appreciable. Recent experiments on "W" diodes have yielded moderate resistivities in the $\approx 10^{-4}$ Ωcm^2 range, which are large enough roughly to double the required bias (and hence the thermal load) when the device is operated at high current density (*e.g.,*mid-10^3 A/cm^2). Thus some form of ICL configuration is likely to be advantageous. On the other hand, standard ICL designs with 20-30 periods may be difficult to operate at high temperatures because of the excessive amount of heat that must be generated just to reach the lasing threshold. Since the thermal generation rate scales as the current-voltage product, it may be advantageous to reduce the voltage and also reduce the internal loss by decreasing the number of stages (as long as that can be accomplished without substantially increasing j_{th}).

We now simulate the performance of a five-stage ICL with cavity length 2 mm, stripe width of 300 μm, a PCDFB grating with alumina top clad, and epitaxial-side-down mounting. One facet is taken to have a high-reflectivity coating (R = 0.95), while the output facet is uncoated. To be consistent with typical experimental results for "W" active regions (see Section 2), we assume a Shockley-Read lifetime of 20 ns and a 2D Auger coefficient of $4 \times 10^{-16} \exp[(T_e - 300)/144]$ cm^4/s, where T_e is the electron temperature. The gain is derived from an eight-band $k{\bullet}p$ calculation with a Gaussian broadening of 5 meV (a little more than the upper limit extracted from the data in Section 2 when band filling effects were included). We take the cladding loss to be 10 cm^{-1} and the free-carrier absorption cross section in the active region to be 10^{-17} cm^2, values which are slightly optimistic but not unreasonable. Diffraction losses induced by the PCDFB grating are taken from TDFT simulations for the given stripe width.

Figure 15 shows the simulated L-I results for CW operation at several temperatures. The five-period PCDFB ICL is projected to produce 1.2 W at room temperature, which goes up to 6.3 W at 250 K and 9.5 W at 200 K. Since the "W" structures studied to date have exhibited active-region free-carrier absorption cross sections of at least several times 10^{-17} cm^2, whereas theory projects that the smaller value should be achievable in optimized structures, Figure 16 plots the projected maximum CW output power as a function of cross section. The simulation predicts that the device will not operate at 300 K if the cross section exceeds 5×10^{-17} cm^2. Nevertheless, even for cross sections as high as 10^{-18} cm^2 several Watts may still be attainable at 250 K.

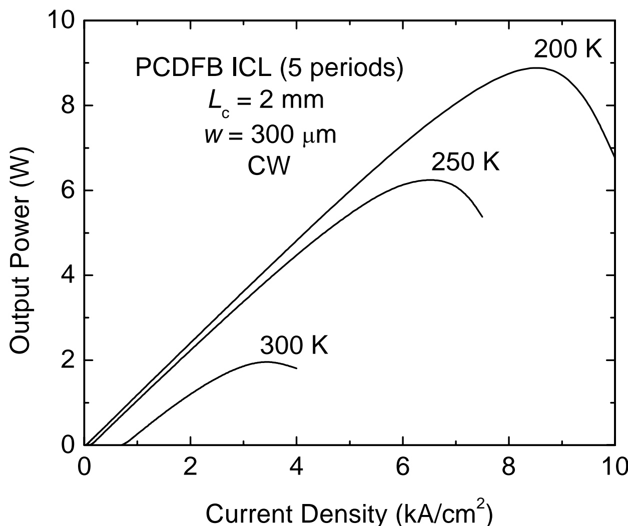


Figure 15. Projected CW L-I characteristics of a five-period PCDFB ICL with stripe width 300 μm and cavity length 2 mm, at several operating temperatures. Parameters employed in the simulations are specified in the text.

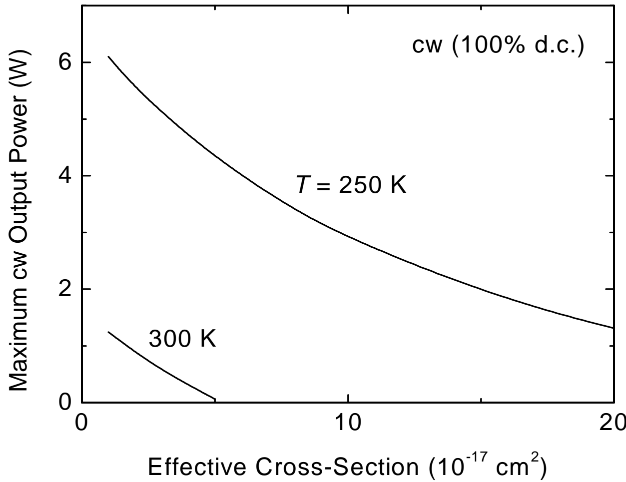


Figure 16. Projected maximum output power of the same CW PCDFB ICL as a function of effective absorption cross section in the active region, at two operating temperatures

8 Conclusions

This chapter has reviewed the various classes of antimonide Type-II "W" lasers developed to date and the status of their performance, as well as factors limiting their operating temperature, output power, spectral purity, and beam quality. Since the "W" concept was first introduced in 1995, dramatic progress has been made by both optically-pumped devices and their electrically-injected counterparts. Recent results from the latest embodiments of the "W"-ICL configuration have been especially encouraging. PCDFB designs that maintain coherent output from very wide stripes open a promising avenue to "W" devices exhibiting both much higher powers and good beam qualities.

Nonetheless, further issues must be resolved before "W" diodes can operate efficiently in CW mode at higher temperatures. Carrier transport should benefit from the incorporation of smoothing transition regions and appropriate doping profiles. Less straightforward to address are the dominance of Auger non-radiative decay and greater-than-expected internal losses at higher operating temperatures, which induce high lasing thresholds and reduced differential efficiencies. Systematic variation of the design parameters may elucidate the fundamental limitations and suggest solutions. Optimization of the heat sinking, especially using epitaxial-side-down mounting, should close the gap between pulsed and CW performance levels. While there is still substantial room for improving the current generation of antimonide Type-II "W' devices, we envision that they will ultimately play a valuable role in addressing the need for convenient and low-cost mid-IR sources.

References

1. J. R. Meyer, C. A. Hoffman, F. J. Bartoli, and L. R. Ram-Mohan, Appl. Phys. Lett. **67**, 757 (1995).
2. J. I. Malin, J. R. Meyer, C. L. Felix, J. R. Lindle, L. Goldberg, C. A. Hoffman, F. J. Bartoli, C.-H. Lin, P. C. Chang, S. J. Murry, R. Q. Yang, and S.-S. Pei, Appl. Phys. Lett. **68**, 2976 (1996).
3. W. Braun, P. Dowd, C.-Z. Guo, S.-L. Chen, C. M. Ryu, U. Koelle, S. R. Johnson, Y.-H. Zhang, J. W. Tomm, T. Elsasser, and D. J. Smith, *J. Appl. Phys.* **88**, 3004 (2000).
4. W. W. Chow, O. B. Spahn, H. C. Schneider, and J. F. Klem, *IEEE J. Quant. Electron.* **37**, 1178 (2001).
5. S.-W. Ryu and P. D. Dapkus, *Electron. Lett.* **38**, 564 (2002).
6. X. H. Zheng, D. S. Jiang, S. Johnson, and Y. H. Zhang, *Appl. Phys. Lett.* **83**, 4149 (2003).
7. I. Vurgaftman, J. R. Meyer, N. Tansu, and L. J. Mawst, *Appl. Phys. Lett.* **83**, 2742 (2003).
8. I. Vurgaftman, J. R. Meyer, N. Tansu, and L. J. Mawst, *J. Appl. Phys.* (in press).
9. R. H. Miles, D. H. Chow, Y.-H. Zhang, P. D. Brewer, and R. G. Wilson, *Appl. Phys. Lett.* **66**, 1921 (1995).
10. M. W. Sigrist, (ed.) *Air Monitoring by Spectroscopic Techniques*, (New York, Wiley, 1994).
11. M. J. Yang, W. J. Moore, B. R. Bennett, and B. V. Shanabrook, Electron. Lett. **34**, 1525 (1998).
12. M. J. Yang, W. J. Moore, B. R. Bennett, B. V. Shanabrook, J. O. Cross, W. W. Bewley, C. L. Felix, I. Vurgaftman, and J. R. Meyer, J. Appl. Phys. **86**, 1796 (1999).
13. G. W. Turner and H. K. Choi, in *Antimonide-Related Strained-Layer Heterostructures*, edited by M. O. Manasreh, (Gordon and Breach, Singapore, 1997), Chap. 8.
14. C. Mourad, D. Gianardi, K. J. Malloy, and R. Kaspi, J. Appl. Phys. **88**, 5543 (2000).
15. T. Borca Tasciuc, D. W. Song, J. R. Meyer, I. Vurgaftman, M. J. Yang, B. Z. Nosho, L. J. Whitman, H. Lee, R. U. Martinelli, G. W. Turner, M. J. Manfra, and G. Chen, J. Appl. Phys. **92**, 4994 (2002).
16. J. Steinshnider, M. Weimer, R. Kaspi, and G. W. Turner, Phys. Rev. Lett. **85**, 2593 (2000).
17. B. Z. Nosho, W. Barvosa-Carter, M. J. Yang, B. R. Bennett, and L. J. Whitman, Surf. Sci. **465**, 361 (2000).
18. R. Kaspi, J. Steinshnider, M. Weimer, C. Moeller, and A. Ongstad, J. Cryst. Growth **225**, 544 (2001).
19. C. L. Canedy, W. W. Bewley, C. S. Kim, M. Kim, I. Vurgaftman, and J. R. Meyer, J. Appl. Phys. **94**, 1347 (2003).
20. R. Kaspi, A. Ongstad, G. C. Dente, J. Chavez, M. L. Tilton, and D. Gianardi, Appl. Phys. Lett. **81**, 406 (2002).
21. L. R. Ram-Mohan and J. R. Meyer, J. Nonlinear Opt. Phys. Mater. **4**, 191 (1995).
22. J. R. Meyer, C. L. Felix, W. W. Bewley, I. Vurgaftman, E. H. Aifer, L. J. Olafsen, J. R. Lindle, C. A. Hoffman, M. J. Yang, B. R. Bennett, B. V. Shanabrook, H. Lee, C.-H. Lin, S. S. Pei, and R. H. Miles, Appl. Phys. Lett. **73**, 2857 (1998).
23. C. L. Canedy, G. I. Boishin, W. W. Bewley, C. S. Kim, I. Vurgaftman, M. Kim, J. R. Lindle, J.R. Meyer, and L. J. Whitman, *J. Vac. Sci. Technol. B* **22**, 1575 (2004).
24. C. L. Felix, J. R. Meyer, I. Vurgaftman, C.-H. Lin, S. J. Murry, D. Zhang, and S.-S. Pei, IEEE Photon. Technol. Lett. **9**, 734 (1997).
25. W. W. Bewley, C. L. Felix, E. H. Aifer, I. Vurgaftman, L. J. Olafsen, J. R. Meyer, H. Lee, R. U. Martinelli, J. C. Connolly, A. R. Sugg, G. H. Olsen, M. J. Yang, B. R. Bennett, and B. V. Shanabrook, Appl. Phys. Lett. **73**, 3833 (1998).

26. C.-H. Lin, R. Q. Yang, S. J. Murry, S.-S. Pei, C. Yan, D. L. McDaniel, Jr., and M. Falcon, IEEE Photon. Technol. Lett. **9**, 1573 (1997).
27. M. E. Flatte, T. C. Hasenberg, J. T. Olesberg, S. A. Anson, T. F. Boggess, C. Yan, and D. L. McDaniel, Jr., Appl. Phys. Lett. **71**, 3764 (1997).
28. H. Q. Le, C. H. Lin, and S. S. Pei, *Appl. Phys. Lett.* **72**, 3434 (1998).
29. D. W. Stokes, L. J. Olafsen, W. W. Bewley, I. Vurgaftman, C. L. Felix, E. H. Aifer, J. R. Meyer, and M. J. Yang, J. Appl. Phys. **86**, 4729 (1999).
30. G. C. Dente, R. Kaspi, A. P. Ongstad, M. L. Tilton, J. C. Chavez, and D. M. Gianardi, Jr., *6th Int. Conf. Mid-IR Optoelectronics Materials and Devices*, (St. Petersburg, June 2004).
31. C. L. Felix, W. W. Bewley, I. Vurgaftman, L. J. Olafsen, D. W. Stokes, J. R. Meyer, and M. J. Yang, Appl. Phys. Lett. **75**, 2876 (1999).
32. A. K. Goyal, G. W. Turner, H. K. Choi, P. J. Foti, M. J. Manfra, T. Y. Fan, and A. Sanchez, Conference Proceedings of the IEEE Lasers and Electro-Optics Society (LEOS) Annual Meeting, Puerto Rico, 2000, p. 249.
33. W. W. Bewley, C. L. Felix, I. Vurgaftman, D. W. Stokes, E. H. Aifer, L. J. Olafsen, J. R. Meyer, M. J. Yang, B. V. Shanabrook, H. Lee, R. U. Martinelli, and A. R. Sugg, Appl. Phys. Lett. **74**, 1075 (1999).
34. W. W. Bewley, C. L. Felix, E. H. Aifer, D. W. Stokes, I. Vurgaftman, L. J. Olafsen, J. R. Meyer, M. J. Yang, and H. Lee, IEEE J. Quantum Electron. **35**, 1597 (1999).
35. W. W. Bewley, C. L. Felix, I. Vurgaftman, D. W. Stokes, J. R. Meyer, H. Lee, and R. U. Martinelli, IEEE Photon. Technol. Lett. **12**, 477 (2000).
36. T. C. McAlpine, K. R. Greene, M. R. Santilli, L. J. Olafsen, W. W. Bewley, C. L. Felix, I. Vurgaftman, J. R. Meyer, H. Lee, and R. U. Martinelli, *J. Appl. Phys.* (in press).
37. R. Kaspi, A. Ongstad, C. Moeller, G. C. Dente, J. Chavez, M. L. Tilton, and D. Gianardi, Appl. Phys. Lett. **79**, 302 (2001).
38. A. K. Goyal, G. W. Turner, M. J. Manfra, P. J. Foti, P. O'Brien, and A. Sanchez, Conference Proceedings of the IEEE Lasers and Electro-Optics Society (LEOS) Annual Meeting, San Diego, CA, 2001, p. 200.
39. A. P. Ongstad, R. Kaspi, C. E. Moeller, M. L. Tilton, J. R. Chavez, and G. C. Dente, J. Appl. Phys. **95**, 1619 (2004).
40. E. V. Arzhanov, A. P. Bogatov, V. P. Konyaev, O. M. Nikitina, and V. I. Shveikin, Quantum Electron. **24**, 581 (1994).
41. W. W. Bewley, C. L. Canedy, C. S. Kim, I. Vurgaftman, M. Kim, and J. R. Meyer, Physica E **20**, 466 (2004).
42. G. Hoffmann, C. Schwender, B. Vogelgesang, H. J. Schimper, J. O. Drumm, N. Herhammer, G. F. West, H. Fouckhardt, and M. Scheib, *IEE Proc.-Optoelectron.* **150**, 403 (2003).
43. L. Song, I. Vurgaftman, S. Degroote, W. W. Bewley, C. S. Kim, G. Borghs, J. R. Meyer, and P. Heremans, IEE Proc.-Optoelectron. **150**, 327 (2003).
44. W. W. Bewley, H. Lee, I. Vurgaftman, R. J. Menna, C. L. Felix, R. U. Martinelli, D. W. Stokes, D. Z. Garbuzov, J. R. Meyer, M. Maiorov, J. C. Connolly, A. R. Sugg, and G. H. Olsen, Appl. Phys. Lett. **76**, 256 (2000).
45. H. Lee, L. J. Olafsen, R. J. Menna, W. W. Bewley, R. U. Martinelli, I. Vurgaftman, D. Z. Garbuzov, C. L. Felix, M. Maiorov, J. R. Meyer, J. C. Connolly, A. R. Sugg, and G. H. Olsen, Electron. Lett. **35**, 1743 (1999).
46. S. Suchalkin, D. Donetski, D. Westerfeld, R. Martinelli, I. Vurgaftman, J. R. Meyer, S. Luryi, and G. Belenky, *Appl. Phys. Lett.* **80**, 2833 (2002).
47. M. Kim, W. W. Bewley, J. R. Lindle, C. S. Kim, I. Vurgaftman, J. R. Meyer, J. G. Kim, and R. U. Martinelli, Appl. Phys. Lett. **83**, 5374 (2003).
48. W. W. Bewley, I. Vurgaftman, C. S. Kim, M. Kim, C. L. Canedy, J. R. Meyer, J. D. Bruno and F. J. Towner, submitted to *Appl. Phys. Lett.*

49. W. W. Bewley, C. L. Canedy, I. Vurgaftman, C. S. Kim, M. Kim, J. R. Lindle, and J. R. Meyer, unpublished data.
50. A. Joullie, E. M. Skouri, M. Garcia, P. Grech, A. Wilk, P. Christol, A. N. Baranov, A. Behres, J. Kluth, A. Stein, K. Heime, M. Heuken, S. Rushworth, E. Hulicius, and T. Simacek, *Appl. Phys. Lett.* **76**, 2499 (2000).
51. P. Christol, M. El Gazouli, P. Bigenwald, and A. Joullie, *Physica E* **14**, 375 (2002).
52. F. Capasso, R. Paiella, R. Martini, R. Colombelli, C. Gmachl, T. L. Myers, M. S. Taubman, R. M. Williams, C. G. Bethea, K. Unterrainer, H. Y. Hwang, D. L. Sivco, A. Y. Cho, A. M. Sergent, H. C. Liu, and E. A. Whittaker, *IEEE J. Quantum Electron.* **38**, 511 (2002).
53. R. Q. Yang, *Superlatt. Microstruct.* **17**, 77 (1995).
54. I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, *IEEE Phot. Tech. Lett.* **9**, 170 (1997).
55. J. R. Meyer, I. Vurgaftman, R. Q. Yang, and L. R. Ram-Mohan, Electron. Lett. **32**, 45 (1996).
56. C. L. Felix, W. W. Bewley, I. Vurgaftman, J. R. Meyer, D. Zhang, C.-H. Lin, R. Q. Yang, and S.-S. Pei, *IEEE Phot. Tech. Lett.* **9**, 1433 (1997).
57. R. Q. Yang, B. H. Yang, D. Zhang, C.-H. Lin, S. J. Murry, H. Wu, and S. S. Pei, *Appl. Phys. Lett.* **71**, 2409 (1997).
58. B. H. Yang, D. Zhang, R. Q. Yang, C.-H. Lin, S. J. Murry, and S. S. Pei, Appl. Phys. Lett. **72**, 2220 (1998).
59. J. L. Bradshaw, R. Q. Yang, J. D. Bruno, J. T. Pham, and D. E. Wortman, Appl. Phys. Lett. **75**, 2362 (1999).
60. J. D. Bruno, J. L. Bradshaw, R. Q. Yang, J. T. Pham, and D. W. Wortman, Appl. Phys. Lett. **76**, 3167 (2000).
61. R. Q. Yang, C. J. Hill, B. Yang, and J. K. Liu, Appl. Phys. Lett. **83**, 2109 (2003).
62. I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, IEEE J. Quantum Electron. **34**, 147 (1998).
63. [1] C.-H. Lin, R. Q. Yang, D. Zhang, S. J. Murry, S. S. Pei, A. A. Allerman, and S. R. Kurtz, Electron. Lett. **33**, 598 (1997).
64. L. J. Olafsen, E. H. Aifer, I. Vurgaftman, W. W. Bewley, C. L. Felix, J. R. Meyer, D. Zhang, C.-H. Lin, and S. S. Pei, Appl. Phys. Lett. **72**, 2370 (1998).
65. R. Q. Yang, J. L. Bradshaw, J. D. Bruno, J. T. Pham, D. E. Wortman, and R. L. Tober, Appl. Phys. Lett. **81**, 397 (2002).
66. R. Q. Yang, J. L. Bradshaw, J. D. Bruno, J. T. Pham, and D. E. Wortman, IEEE J. Quantum Electron. **38**, 559 (2002).
67. J. L. Bradshaw, J. D. Bruno, J. T. Pham, D. E. Wortman, and R. Q. Yang, *J. Vac. Sci. Technol. B* **18**, 1628 (2000).
68. J. L. Bradshaw, J. T. Pham, R. Q. Yang, J. D. Bruno, and D. E. Wortman, *IEEE J. Sel. Topics Quant. Electron.* **7**, 102 (2001).
69. J. L. Bradshaw, N. P. Breznay, J. D. Bruno, J. M. Gomes, J. T. Pham, F. J. Towner, D. E. Wortman, R. L. Tober, C. J. Monroy, and K. A. Olver, Physica E **20**, 479 (2004).
70. C. J. Hill, B. Yang, and R. Q. Yang, Physica E **20**, 486 (2004).
71. R. Q. Yang, *Proc. 27th Int. Conf. Phys. Semicond.* (Flagstaff, 2004).
72. Yang, C. J. Hill, B. H. Yang, and C. M. Wong, *IEEE Photon. Technol. Lett.* **16**, 987 (2004).
73. C. J. Hill, C. M. Wong, B. Yang, and R. Q. Yang, *Electron. Lett.* **40**, 878 (2004).
74. J. L. Bradshaw, J. D. Bruno, J. T. Pham, D. E. Wortman, S. Zhang, and S. R. J. Brueck, IEE Proc.-Optoelectron. **150**, 288 (2003).
75. R. Q. Yang, C. J. Hill, B. H. Yang, C. M. Wong, R. E. Muller, and P. M. Echternach, Appl. Phys. Lett. **84**, 3699 (2004).
76. C. L. Felix, W. W. Bewley, I. Vurgaftman, J. R. Meyer, L. Goldberg, D. H. Chow, and E. Selvig, Appl. Phys. Lett. **71**, 3483 (1997).

77. W. W. Bewley, C. L. Felix, I. Vurgaftman, E. H. Aifer, J. R. Meyer, L. Goldberg, J. R. Lindle, D. H. Chow, and E. Selvig, IEEE Photon. Technol. Lett. **10**, 660 (1998).
78. R. J. Lang, K. Dzurko, A. A. Hardy, S. DeMars, A. Schoenfelder, and D. F. Welch, IEEE J. Quantum Electron. **34**, 2196 (1998).
79. A. M. Sarangan, M. W. Wright, J. R. Marciante, and D. J. Bossert, IEEE J. Quantum Electron. **35**, 1220 (1999).
80. R. E. Bartolo, W. W. Bewley, I. Vurgaftman, C. L. Felix, J. R. Meyer, and M. J. Yang, *Appl. Phys. Lett.* **76**, 3164 (2000).
81. I. Vurgaftman, W. W. Bewley, R. E. Bartolo, C. L. Felix, M. J. Jurkovic, J. R. Meyer, M. J. Yang, H. Lee, and R. U. Martinelli, *J. Appl. Phys.* **88**, 6997 (2000).
82. S. Kalluri, T. Vang, R. Lodenkamper, M. Nesnidal, M. Wickham, D. Forbes, J. Lacey, L. Lembo, and J. Brock, Proceedings, OSA Topical Meeting on Advanced Semiconductor Lasers and Applications, 1999, paper AWA6.
83. I. Vurgaftman and J. R. Meyer, *Appl. Phys. Lett.* **78**, 1475 (2001).
84. I. Vurgaftman and J. R. Meyer, *IEEE J. Quantum Electron.* **38**, 592 (2002).
85. W. W. Bewley, C. L. Felix, I. Vurgaftman, R. E. Bartolo, J. R. Lindle, J. R. Meyer, H. Lee, and R. U. Martinelli, Appl. Phys. Lett. **79**, 3221 (2001).
86. W. W. Bewley, C. L. Felix, I. Vurgaftman, R. E. Bartolo, J. R. Lindle, J. R. Meyer, H. Lee, and R. U. Martinelli, Solid State Electron. **46**, 1557 (2002).
87. I. Vurgaftman, W. W. Bewley, C. L. Canedy, J. R. Lindle, C. S. Kim, J. R. Meyer, S. J. Spector, D. M. Lennon, G. W. Turner, and M. J. Manfra, IEE Proc.-Optoelectron. **150**, 322 (2003).
88. W. W. Bewley, C. S. Kim, M. Kim, C. L. Canedy, J.R. Lindle, I. Vurgaftman, J. R. Meyer, R. E. Muller, P. M. Echternach, and R. Kaspi, Appl. Phys. Lett. **83**, 5383 (2003).
89. C. S. Kim, W. W. Bewley, C. L. Canedy, I. Vurgaftman, M. Kim, and J. R. Meyer, IEEE Photon. Technol. Lett. **16**, 1250 (2004).
90. I. Vurgaftman and J. R. Meyer, IEEE J. Quantum Electron. **39**, 689 (2003).
91. C. H. Grein and H. Ehrenreich, J. Appl. Phys. **93**, 1075 (2003).
92. C. H. Grein, P. M. Young, and H. Ehrenreich, J. Appl. Phys. **76**, 1940 (1994).
93. H. P. Hjalmarson and S. R. Kurtz, Appl. Phys. Lett. **69**, 949 (1996).
94. M. E. Flatte, C. H. Grein, H. Ehrenreich, Appl. Phys. Lett. **72**, 1424 (1998).
95. M. E. Flatte, C. H. Grein, T. C. Hasenberg, S. A. Anson, D. J. Jang, J. T. Olesberg, and T. F. Boggess, Phys. Rev. B **59**, 5745 (1999).
96. M. E. Flatte and C. H. Grein, Opt. Express **2**, 131 (1998).
97. C. H. Grein, M. E. Flatte, J. T. Olesberg, S. A. Anson, L. Zhang, and T. F. Boggess, J. Appl. Phys. **92**, 7311 (2002).
98. J. O. Drumm, B. Vogelgesang, G. Hoffmann, C. Schwender, N. Herhammer, and H. Fouckhardt, *Semicond. Sci. Technol.* **17**, 1115 (2002).
99. M. E. Flatte, J. T. Olesberg, S. A. Anson, T. F. Boggess, T. C. Hasenberg, R. H. Miles, and C. H. Grein, Appl. Phys. Lett. **70**, 3212 (1997).
100. W. W. Bewley, I. Vurgaftman, C. L. Felix, J. R. Meyer, C.-H. Lin, D. Zhang, S. J. Murry, S. S. Pei, and L. R. Ram-Mohan, J. Appl. Phys. **83**, 2384 (1998).
101. H. H. Chen, M. P. Houng, Y. H. Wang, and Y.-C. Chang, Appl. Phys. Lett. **61**, 509 (1992).
102. F. Szmulowicz and G. J. Brown, Phys. Rev. B **51**, 13203 (1995).
103. W. H. Lau and M. E. Flatte, Appl. Phys. Lett. **80**, 1683 (2002).
104. F. Szmulowicz, H. Haugan, and G. J. Brown, Phys. Rev. B **69**, 155321 (2004).
105. G. C. Dente and M. L. Tilton, Phys. Rev. B **66**, 165307 (2002).
106. R. Magri and A. Zunger, Phys. Rev. B **65**, 165302 (2002).
107. R. Magri and A. Zunger, Phys. Rev. B **68**, 155329 (2003).
108. J. T. Olesberg, M. E. Flatté, T. C. Hasenberg, and C. H. Grein, *J. Appl. Phys.* **89**, 3283 (2001).

109. S. Suchalkin, J. Bruno, R. Tober, D. Westerfeld, M. Kisin, and G. Belenky, *Appl. Phys. Lett.* **83**, 1500 (2003).
110. J. T. Olesberg, M. E. Flatté, and T. F. Boggess, *J. Appl. Phys.* **87**, 7164 (2000).
111. S. A. Anson, J. T. Olesberg, M. E. Flatté, T. C. Hasenberg, and T. F. Boggess, *J. Appl. Phys.* **86**, 713 (1999).
112. M. Lerttamrab, S. L. Chuang, R. Q. Yang, and C. J. Hill, *J. Appl. Phys.* **96**, 3568 (2004).

Interface Lasers with Asymmetric Band Offset Confinement

K.D. Moiseev and Y.P. Yakovlev
A. F. Ioffe Physico-Technical Institute, Russian Academy of Science,
St. Petersburg, Russia

1 Introduction

A laser structure with indirect radiative transitions across the interface of Type II staggered heterojunction was proposed by Kroemer and Griffiths in 1983 [1]. Such a Type II GaInAsSb/GaSb heterostructure with self-consistent quantum wells at the interface was experimentally realised for the first time in 1986 [2]. The basic feature of the Type II heterojunction is that the conduction band and the valence band offsets at the heterointerface have the same sign. Then, under applied bias, electrons and holes which are localised in self-consistent potential wells at the interface are spatially separated, contrary to the situation in Type I heterojunctions, and therefore the recombination between them is only possible by carrier tunnelling through the potential barrier.

Lasers incorporating a Type II heterojunction in the active region have a number of advantages: these lasers employ majority carriers which move towards the heterointerface, tunnel through the interface barrier and radiatively recombine with carriers of the opposite sign. Furthermore, it is possible to produce photons of energy less than that of the forbidden gap of the alloys forming the heterojunction. The usual processes of diffusion and carrier thermalisation, which are typical for conventional lasers, can be completely excluded.

Rogachev of the Ioffe Institute proposed a mid-infrared laser based on tunnelling injection in Type II broken-gap InAs/GaSb heterostructures in 1989 [3]. Such a laser structure was realised in 1997 when the observation of room temperature lasing at 2 μm from Type II broken-gap InAs/GaSb multiple quantum wells was reported [4]. An improvement in laser performance was expected due to a decrease in threshold current and a weakening of its temperature dependence in connection with the decrease in carrier leakage and the suppression of the non-radiative Auger-recombination rate at the Type II interface. The intense interface electroluminescence, observed from a single broken-gap p-GaInAsSb/p-InAs heterojunction, allowed the design of a mid-infrared (3-4 μm) laser with a Type II broken-gap heterojunction in the active region [5]. The lasing occurs due to the radiative transitions of electrons and holes from spatially quantised energy subbands across the heterointerface. This is a new Type of laser with tunnel-assisted interface recombination [6]. Quantum cascading of tunnel-injection lasers was proposed by Kazarinov and Laikhtman in 1994 [7].

2 The 2D-Electron Channel in a Type II Broken-gap p-GaInAsSb/p-InAs Heterointerface

Recently, an electron channel with a high Hall mobility (up to 70,000 $cm^2V^{-1}s^{-1}$ at 77 K) was observed in lattice-matched p-$Ga_{1-x}In_xAs_{1-y}Sb_y$/p-InAs single heterostructures. The undoped and slightly doped quaternary solid solutions, enriched with GaSb, in the composition range $x<0.22$ were grown on InAs substrate by LPE [8]. Existence of the electron channel at the p-GaInAsSb/p-InAs heteroboundary was visually demonstrated by an STM study of a cleaved surface of the heterostructure at room temperature [9]. The 2D-electron channel is a unique feature of the abrupt heteroboundary with a transition layer four monolayers (MLs) wide formed during the planar growth of the quaternary solid solution under thermodynamic equilibrium conditions [10]. The presence of at least two electron energy subbands (E_1 and E_2) was detected from Hall measurements and cyclotron resonance studies in high magnetic fields [11,12]. The most impressive result was the observation of integer quantum Hall effect (QHE) plateaux in the Hall conductivity with filling factors of 2, 3 and 6 when the quantum limit for the lowest E_1 subband was realized [13]. This was the first demonstration of the QHE in a Type II single GaInAsSb/InAs heterostructure with self-consistent quantum wells grown by LPE. Using a high level of Zn doping ($>10^{-2}$ at %) in the GaInAsSb solid solution produced a severe drop in the Hall mobility. This arises from the electron channel becoming shallower and depletion of the mobile carriers as a result of localisation by random potential fluctuations at the heterointerface and the increasing role of interface roughness scattering [14]. An abrupt variation of the Hall mobility, $\mu_H \sim d^2$ with the electron channel width (d) was found on increasing the acceptor doping level [15]. These facts confirm that electrons produce the major contribution to the total conductivity.

3 Interface Luminescence Properties of the Type II Broken-gap Single p-GaInAsSb/p-InAs Heterojunction

3.1 Interface EL in a Single p-GaInAsSb/p-InAs Heterojunction

The first observation of interface electroluminescence (EL) from a Type II broken-gap p-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/p-InAs heterojunction was reported by Mikhailova *et al* [16]. Intense EL with a maximum photon energy less than that of the InAs energy gap was observed in a wide temperature range from 4.2 to 300 K, under both quasi steady-state conditions and in pulsed mode [17,18]. The heterostructure under investigation was based on wide-gap quaternary epilayer heavily doped with Zn. Current-voltage (I-V) characteristics for the Type II broken-gap single p-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/p-InAs heterostructure manifested rectifying behaviour at 77 K. Assuming it to behave as a Schottky diode, it would be expected to exhibit electroluminescence in the narrow-gap part of the p-$GaIn_{0.16}As_{0.22}Sb$/p-InAs heterojunction under forward bias (when the p-type InAs was biased positively).

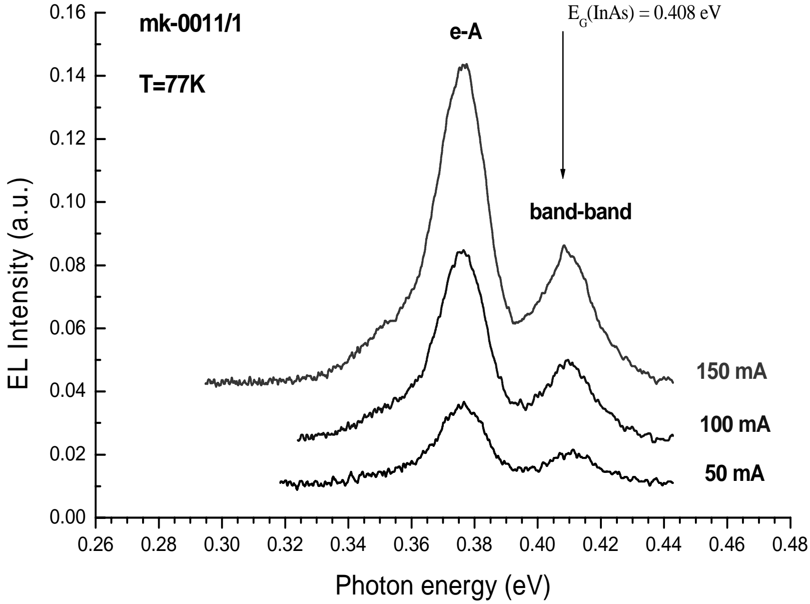


Figure 1. Electroluminescence emission spectra for the p-p Type II broken-gap single heterostructure in forward bias at 77K

However, it is striking that intense luminescence for this structure was also observed under "reverse" bias (*i.e.* when the p-InAs was biased negatively).

Under forward bias, EL was obtained in the spectral range of 0.32-0.44 eV (Figure 1). The EL spectra at 77 K contained two pronounced symmetric, Gaussian-like, emission bands with peak photon energies of 0.383eV ($h\nu_1$) and 0.408 eV ($h\nu_2$). The emission band $h\nu_2$ can be ascribed to interband radiative transitions in the InAs substrate while the emission band $h\nu_1$ can be associated with transitions into the Zn impurity level with an activation energy, E_{Zn}, of 25 meV [16]. On the other hand, three pronounced emission bands were detected in the spectral range 0.28-0.68 eV under reverse bias at 77 K (Figure 2). Two low-energy emission bands at 0.310 eV ($h\nu_A$) and 0.378 eV ($h\nu_B$) were narrow (with a FWHM of ~10-12 meV) while a third, high-energy band at 0.633eV ($h\nu_C$) was weaker and broadened, with a FWHM of 60meV. The separation between the peaks A and B was found to be 68meV. As can be seen in Figure 2, a redistribution in the relative intensity and a change in spectral positions was observed with increasing injection current [5]. The emission band B appears at a high injection level and then dominates the EL spectra.

To explain the experimental results it is necessary to consider the energy band diagram of the p-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/p-InAs heterostructure under applied reverse bias (Figure 3). A 2D-electron gas is produced in the triangular quantum well at the interface due to electron transfer from the valence band of the wide-gap semiconductor to the conduction band of InAs. By applying an external electric field

with a negative potential to the narrow-gap semiconductor, the conduction band of p-InAs and the valence band of p-GaInAsSb at the interface are shifted with respect to each other in such a way that the corresponding bands move up and down, respectively. Due to strong band bending, there are no free holes near the interface in the p-InAs to recombine with the electrons localised in the quantum well. Under applied bias, electrons can resonantly tunnel from acceptor levels in the bulk p-InAs to the electron quantum well at the interface where the quasi Fermi-level crosses the electron level in the well. This leads to an increase in the electron population in the well. As a result, non-equilibrium electrons and holes localised at different sides of the interface can effectively recombine. We suppose that the narrow emission peaks at 0.310 eV ($h\nu_A$) and 0.378 eV ($h\nu_B$) are due to radiative tunnelling of the electrons in discrete levels E_1 and E_2 and recombination with holes on the GaInAsSb side. Possible transitions are labelled as A and B (see Figure 2 and 3). The radiative transition labelled C can be ascribed to the recombination of Auger-electrons ejected from the quantum well with bulk holes in the GaInAsSb layer.

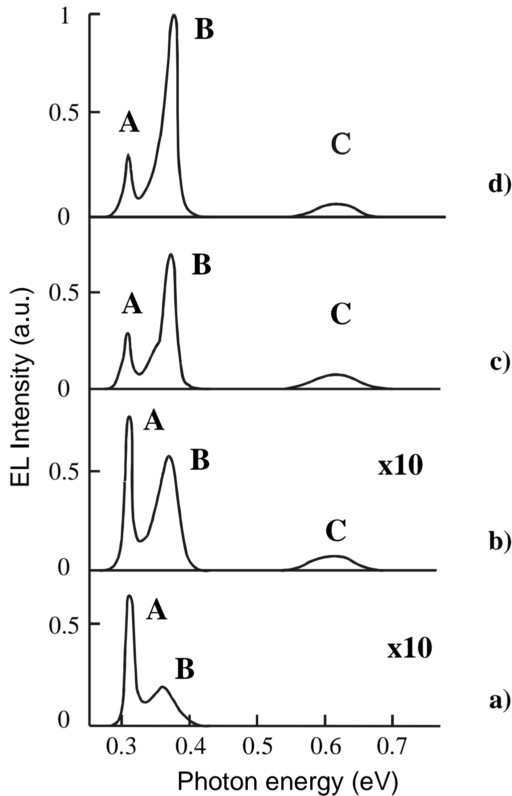


Figure 2. EL spectra of the p-$GaIn_{0.16}As_{0.22}Sb$/p-InAs single heterostructure under applied "reverse" bias at T=77 K for different drive current: (**a**) - 20 mA; (**b**) - 50 mA; (**c**) - 100 mA; (**d**) - 120 mA

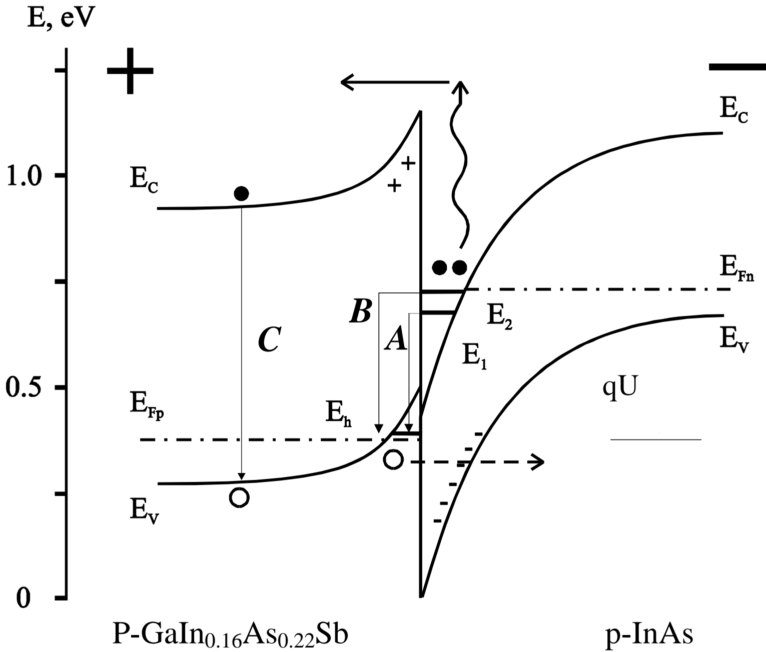


Figure 3. Energy band diagram of p-$GaIn_{0.16}As_{0.22}Sb$/p-InAs heterojunction under reverse bias at T=77 K. E_1, E_2 and E_h are electron and hole levels in the quantum wells at the heterointerface. E_{Fn} and E_{Fp} are quasi Fermi-levels for electrons and holes, respectively.

The energy spectrum of the quantum well was theoretically calculated using a quasi-classical approach. The energies of the electron and hole levels were found to be 247 meV, 315 meV and 50 meV, for E_1, E_2, and E_h respectively. A radiative recombination rate as high as 2×10^{18} cm^2s^{-1} was computed, with recombination lifetimes for electrons in the first level, $E_{1,}$ and the second level, $E_{2,}$ as long as 2.1×10^{-7} s and 1.8×10^{-7} s respectively.

We also examined the two main emission bands at 0.310 eV ($h\nu_A$) and 0.378 eV ($h\nu_B$) in the low-energy region, under pulsed bias conditions at low temperatures and observed the splitting of the B line into two sub-bands with peak photon energies of 0.374 eV ($h\nu_{B1}$) and 0.402 eV ($h\nu_{B2}$) (Figure 4). The diamagnetic shifts of the spectral position for each EL line were found to be $\Delta h\nu_{B2}/\Delta B$=1.74 meV/T, $\Delta h\nu_{B1}/\Delta B$=2.17 meV/T and $\Delta h\nu_A/\Delta B$=0.87 meV/T, respectively, in low fields up to B=5 T [17]. These dependencies reflect a difference in effective mass for the electron energy sub-bands in the 2D-electron channel at the heterointerface and confirm that the lower electron sub-band, E_1, is associated with a heavier effective mass value (m_{e1}=0.042 m_0) than that calculated for the upper sub-band E_2. The origin of the emission bands B_1 and B_2 (m~0.026-0.03m_0) is not yet clearly known and is currently the subject of further study. It is natural to suppose that transitions B_1 and B_2 can be ascribed to two higher electron sub-bands with an energy

separation, ΔE_{2-3}, of 28 meV at low temperature. Due to their energy positions close to the Fermi level in p-InAs, they are very sensitive to an increase in temperature. As can be seen in Figure 4, line A exhibits superior radiative performance compared to line B in the higher temperature range (T>77 K). This can be associated with the redistribution of the quantum well occupation due to thermalisation of the upper lying electron sub-bands with increasing temperature.

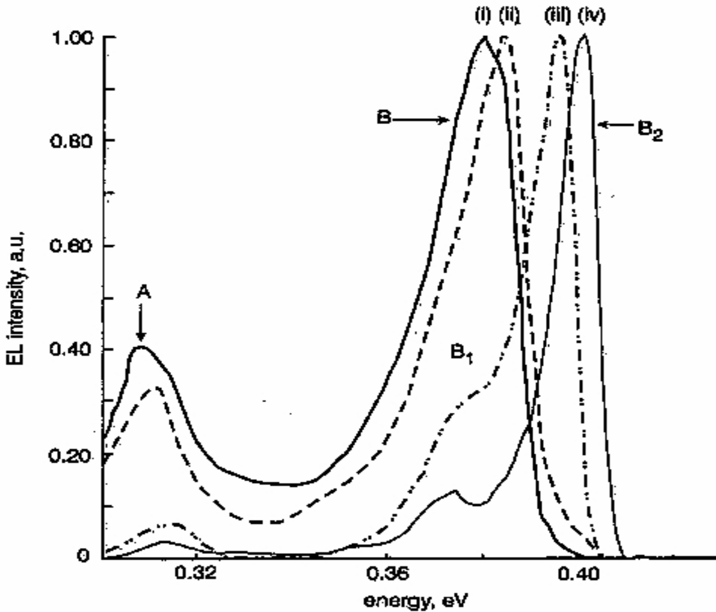


Figure 4. EL spectra of the p-GaInAsSb/p-InAs heterojunction under pulsed electric field at low temperature (1) 4.2 K, (2) 15 K, (3) 77 K and (4) 100 K. (Reproduced by permission from Moiseev *et al.*, IEE Proc.-Optoelectron., v.145, IEE, 1998).

3.2 Tunnelling-injection Laser Based on the Type II p-GaIn$_{0.17}$AsSb/p-In$_{0.83}$GaAsSb Heterojunction

A laser structure with separate electron and optical confinement was grown on p-InAs (100) substrate by LPE. A Type II broken-gap p-p heterojunction in the active layer was formed by two quaternary p-$Ga_{1-x}In_xAs_ySb_{1-y}$ solid solutions with different indium contents (x=0.17 for wide-gap and x=0.83 for narrow-gap) where the energy band gaps were 0.635 eV and 0.393 eV at 77 K, respectively (Figure 5).

To achieve optical confinement a cladding layer of p-InAsSbP (E_g=0.524 eV) was used because of the large difference in refractive index, (Δn~0.2), between this layer and the GaInAsSb epilayer. The wide-gap GaInAsSb p-n junction acted as a capping layer to facilitate electron injection [19]. The contact layer was formed using a heavily doped narrow-gap n^+-InGaAsSb solid solution.

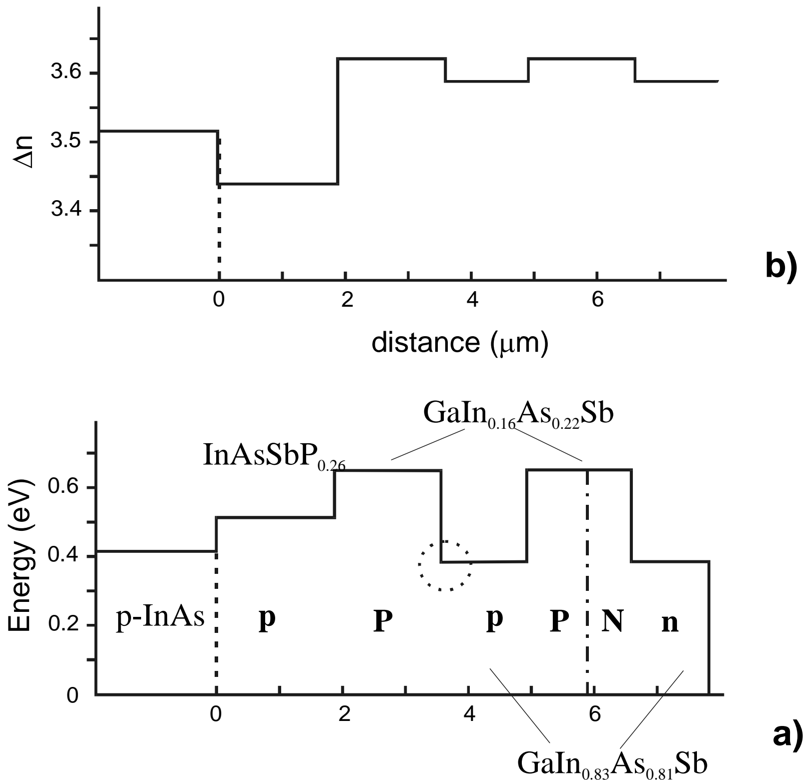


Figure 5. Laser structure based on Type II broken-gap p-p heterojunction: **a**) Layered energy band gap diagram; **b**) Layered refractive index profile

As shown in Figure 6, the resulting 77 K spontaneous emission spectra contained two lines at 0.318 meV ($h\nu_A$) and 0.378 meV ($h\nu_B$) which are similar to the low-energy ones observed for the single p-$GaIn_{0.16}As_{0.22}Sb$/p-InAs heterostructure. This data allows us to suppose that p-p heteroboundary marked by the dotted circle in Figure 5a is responsible for the observed interface electroluminescence.

Coherent emission appeared at the high-energy edge of the spontaneous emission band ($h\nu_B$) associated with tunnelling-assisted radiative transitions between the electron sub-band E_2 and the hole sub-band (see Figure 3). Single-mode lasing peaking at 3.26 μm was obtained in pulsed mode at 77 K. It was found that the threshold current depended weakly on temperature in the range 77-110 K, and a high value of 60 K was obtained for the characteristic temperature, T_0. However, the threshold current density observed in this laser structure was high (~2 kA/cm^2) mainly due to electron leakage at the p-p interface. Using a number of wide-gap epilayers resulted in a significant contribution to the series resistance for this laser structure. Nevertheless, the attainment of a high characteristic temperature indicates the potential benefits of the interface laser for mid-infrared operation. To decrease the threshold current, the design of the laser structure will have to be further refined.

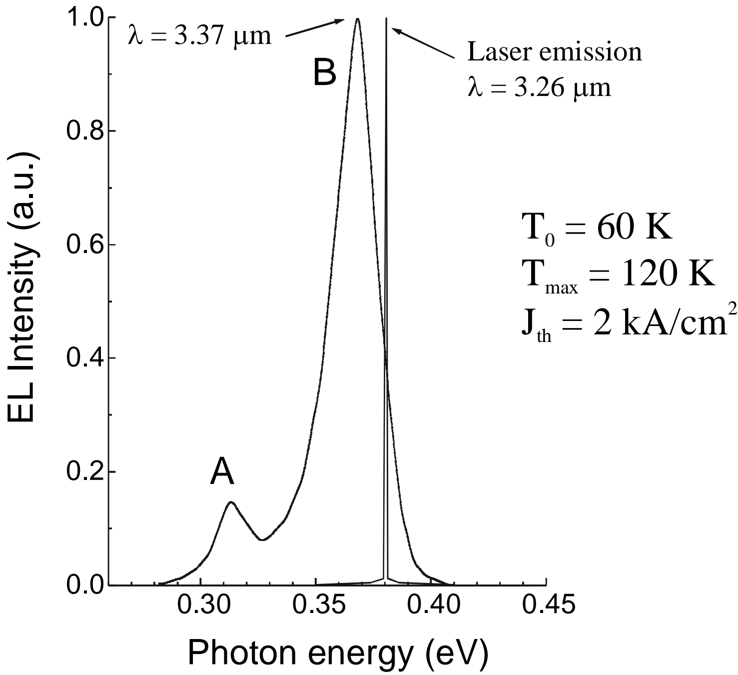


Figure 6. EL spectra for the laser structure based on a Type II broken-gap p-p heterojunction at 77K: spontaneous emission at I=100 mA and lasing at $I=1.05*I_{th}$

3.3 Suppression of Auger-recombination Processes at the Type II Heterointerface

An improved temperature characteristic and reduced threshold current can be achieved as a result of Auger suppression at the Type II heterointerface leading to weaker temperature dependence of the threshold current and higher quantum efficiency in mid-infrared laser structures [20,21]. In a Type II heterojunction, two Auger-recombination processes can be identified as making an important contribution to the total threshold current (I_{th}) for III-V mid-infrared lasers: (i) the CHCC process with a hot electron that goes up into the conduction band and (ii) the CHHS process with participation of an electron and a hole resulting in the transition of another hole to the spin-orbit split valence band, if the condition $(E_g-\Delta)/E_g>m_c/m_{so}$ is satisfied, where Δ is the spin-orbit splitting energy, E_g is the effective band gap, m_c and m_{so} are the electron effective mass and hole effective mass in the spin-orbit split valence band, respectively (Figure 7).

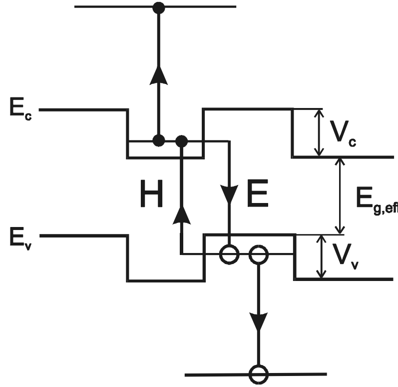


Figure 7. Schematic band diagram of Type II quantum well heterostructure. E and H - two possible recombination channels of electron and hole, respectively, across the heterointerface

As shown by Zegrya *et al.* [21] contributions to the Auger transition matrix element (M) from the channels mentioned above are of the same order due to mutual transformation of light to heavy holes and otherwise when they interact with the heteroboundary. These contributions to M have opposite signs in the Type II heterojunction and compensate each other when being summed. Hence, the two channels of electron-hole recombination interfere destructively, and could result in a decrease of the matrix element and the Auger-recombination rate

$$\frac{G_{II}}{G_I} \sim \left(\frac{T}{V_C}\frac{m_h}{m_c}\right)^3 \frac{V_C}{E_g} << 1, \tag{1}$$

but the radiative recombination rates in Type I and II heterojunctions are comparable:

$$\frac{R_{II}}{R_I} \cong \left(\frac{T}{V_C}\frac{m_h}{m_c}\right)^2 \leq 1, \tag{2}$$

where R_I (R_{II}) and G_I (G_{II}) are rates for radiative and non-radiative Auger recombination, respectively, ascribed to the heterojunction with corresponding Type I and Type II band alignment, V_C is the conduction band offset; m_c and m_h are effective masses for electrons and holes, respectively; T is the temperature and E_g is the energy of the forbidden gap of the active region.

The total threshold current density for long wavelength lasers mainly consists of the contribution of radiative and non-radiative Auger recombination currents:

$$J_{th} \sim J_R + J_A \sim e(R(n_{th})^2 + C_A(n_{th})^3) \tag{3}$$

where C_A is the Auger coefficient and n_{th} is the threshold carrier density. Using an interface with Type II alignment instead of Type I, it is possible to meet the condition for the suppression of the CHHS process. The calculated value of Auger coefficients for Type I and Type II heterojunctions are $C_A^I \sim 2\times10^{-15}$ cm^{4}s^{-1} and $C_A^{II} \sim 4.6\times10^{-16}$ cm^{4}s^{-1}, respectively, and, the ratio was therefore found to be $C_A^{II}/C_A^I \sim 0.22$.

4 Interface Laser with Improved Temperature Dependence

4.1 Electroluminescence in a Single p-Ga$In_{0.17}$AsSb/n-$In_{0.83}$GaAsSb Heterostructure

In order to increase the potential barriers for electrons at the Type II heterointerface, a P-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/n-$In_{0.83}Ga_{0.17}As_{0.80}Sb_{0.20}$ heterojunction was employed. Potential barriers, formed by the band-offsets at the heteroboundary, were obtained for both the conduction (ΔE_C=0.6 eV) and valence bands (ΔE_V=0.35 eV), significantly inhibiting the CHCC process (see Figure 8). As illustrated in Figure 8, Auger-electrons are stopped at the p-n heteroboundary, where they accumulate on the narrow-gap epilayer side of the heterojunction.

Electroluminescent properties of a Type II P-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/n-$In_{0.83}Ga_{0.17}As_{0.80}Sb_{0.20}$ single heterostructure grown by LPE and lattice-matched to the InAs substrate have been studied [22,23]. The quaternary solid solutions forming the heterojunction were unintentionally doped. The P-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/n-$In_{0.83}Ga_{0.17}As_{0.82}Sb_{0.18}$ heterostructure was grown on an n-InAs substrate which, in turn, forms a Type II broken-gap n-P heterojunction with the wide-gap P-GaInAsSb. Using such a layer sequence the n-InAs/P-GaInAsSb heterojunction can be used as an effective contact to the wide-gap layer since it exhibits ohmic behaviour over a wide temperature range from liquid helium to room temperature [8]. EL was observed under forward bias (positive voltage was applied to the P-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$ layer and negative to the n-$In_{0.83}Ga_{0.17}As_{0.80}Sb_{0.20}$ layer). Under applied bias, electrons and holes move towards the heteroboundary where they accumulate. The holes are localised on the P-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$ side and the electrons on the n-$In_{0.83}Ga_{0.17}As_{0.80}Sb_{0.20}$ side. The injected carriers fill the potential wells on both sides of the heterointerface due to band bending at the heteroboundary.

As in the case of the p-GaInAsSb/p-InAs heterojunction under "reverse" bias described above, the EL spectra at 77 K also contained two pronounced emission bands, A and B, in the spectral range between 0.25 and 0.45 eV. The low-energy emission band (A) could be detected at a very low current (3mA), with a peak photon energy of 0.260 eV (see insert in Figure 9). This broad band shifted towards higher energy with increasing drive current and reached a value of 0.345 eV at a drive current in excess of 150 mA. Band A was easily resolved up to 30 mA, but thereafter emission band B began to dominate the EL spectrum. Band B ($h\nu_B$=0.383 eV) appeared at higher injection levels than A and remained at constant energy over the whole range of drive currents. The intensity of band B exhibited a superlinear dependence on drive current with slight saturation at high injection levels while the dependence of the EL intensity on drive current for band A could be divided into two regions: ~I^2 at low current but ~I at higher current.

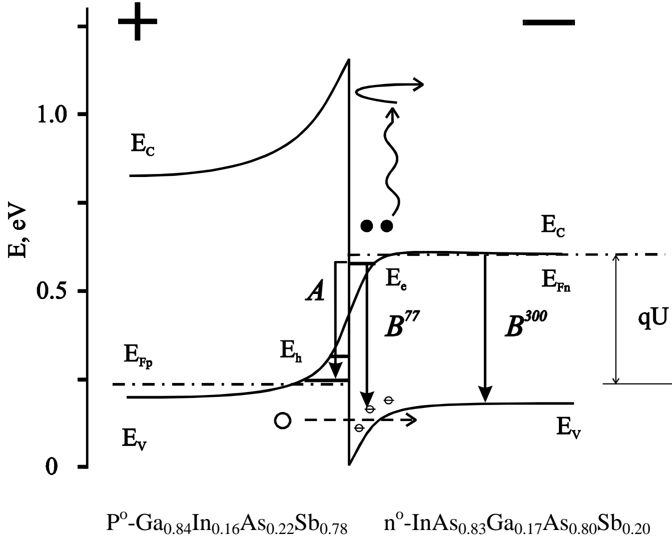


Figure 8. Energy band diagram of Type II P-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/n-$In_{0.83}Ga_{0.17}As_{0.80}Sb_{0.20}$ heterojunction under applied external bias at T=77 K

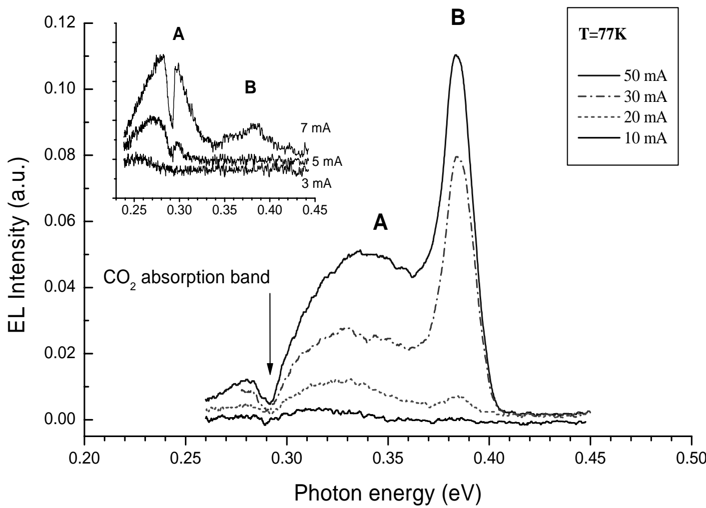


Figure 9. EL spectra of the Type II P^0-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/$n^0$-$In_{0.83}Ga_{0.17}As_{0.82}Sb_{0.18}$ heterostructure measured at various drive currents at T=77 K. *Insert*: low injection level. *Arrows* indicate the atmospheric CO_2 absorption region (4.25-4.27 μm).

4.2 Advanced Tunnel-injection Heterostructure Laser

Further improvements in the laser performance and operating temperature were achieved using a Type II broken-gap p-n heterojunction as the active layer, which provided larger asymmetric potential barriers at the interface [24,25]. This inhibited carrier leakage across the interface significantly and decreased the threshold current density to 400 A/cm^2: a factor of five improvement on the previous Type II lasers based on a broken-gap isoType p-GaInAsSb/p-InAs heterojunction.

The active region of the laser structure consisted of a Type II heterojunction formed with an undoped narrow-gap n-$In_{0.83}Ga_{0.17}As_{0.82}Sb_{0.18}$ (Eg=0.393 eV at 77K) layer and a wide-gap p-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$ (Eg=0.635 eV at 77K) layer lattice-matched to an InAs substrate, sandwiched between p- and n- doped $InAsSb_{0.12}P_{0.26}$ (E_g=0.524eV at 77K) cladding layers. The schematic energy band diagram and refractive index profile for the laser structure are presented in Figure 10. The asymmetric design of the active region provides potential barriers of 0.6 eV for electrons and 0.35 eV for holes at the p-GaInAsSb/n-InGaAsSb heterointerface. [25].

Double-channel mesa-stripe diodes have been fabricated by wet chemical etching using a standard photo-lithographic technique (Figure 11). The stripe width and cavity length of the samples under study were 30 and 500µm, respectively. Ohmic contacts were formed to the p-InAs side using Cr/Au and Zn/Au metal deposition and to the n$^+$-InAsSbP epilayer by Au and Te/Au deposition. Laser structure chips were mounted on a copper heatsink and bonded using Au wire.

The single-mode coherent emission peaked at 3.14 µm in pulsed mode at 77 K. The lasing arose on the high-energy edge of the spontaneous emission spectra and was similar to the p-p laser structure considered above (Figure 12). The difference in wavelength can be explained by the character of the energy bands in the active region of the laser structure.

Light polarization was measured in the p-n laser structure both in spontaneous and coherent emission mode. As can be seen in Figure 13, TM-polarization prevails over TE-mode as the current is increased. It was found that a high contribution of TM-polarisation occurs for both Types of laser structure under study, *i.e.* E is perpendicular to the p-n junction plane. The polarization extent, α, calculated using the equation: $\alpha=(P_{TM}-P_{TE})/(P_{TM}-P_{TE})$, where P_{TM}, P_{TE} are the emission intensities for TM- and TE-polarised light, yields maximum values of α as high as 80 and 73% (at a current density of ~$1.5J_{th}$) for Type II and Type I heterostructures, respectively. As shown by Averkiev *et al.* [26], the presence of TM-polarisation for indirect radiative transitions can be associated with the hybridisation of carrier wave functions at the Type II interface. In a tunnel-injection laser based on a Type II broken-gap p-n heterojunction, radiative transitions have been shown to take place near the interface at low temperature [23]. Figure 14 shows the threshold current temperature dependencies of two Type II tunnel-injection laser structures with either a p-p or p-n broken-gap heterojunction in the active region. In this novel laser structure, the weak temperature dependence of the threshold current density was extended up to 140-150 K, with a characteristic temperature T_0 of 53 K. A maximum operation temperature of 195 K (in pulsed mode) was achieved in this novel laser design owing to Auger suppression.

Figure 10. Laser structure based on Type II broken-gap p-n heterojunction: **a**) Layered energy band gap diagram; **b**) Layered refractive index profile

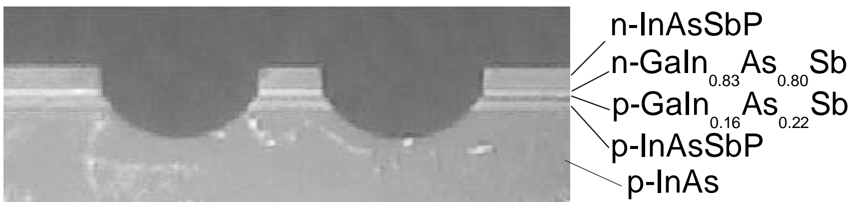


Figure 11. Etched profile of double-channel stripe laser structure

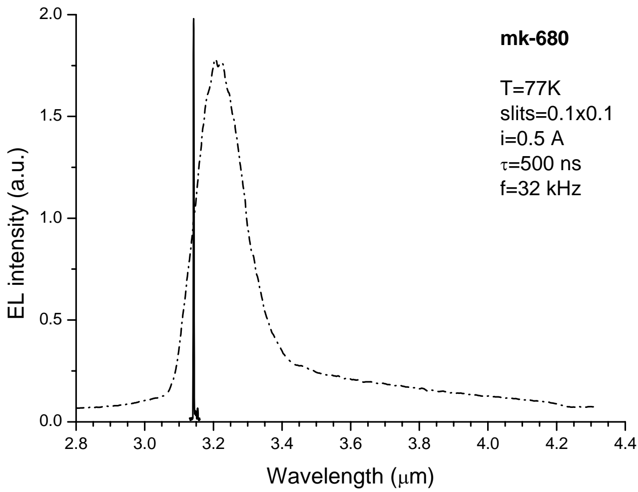


Figure 12. EL spectra from a laser structure based on a Type II broken-gap p-n heterojunction at T=77K: spontaneous emission at I=150 mA and lasing at $I=1.05*I_{th}$

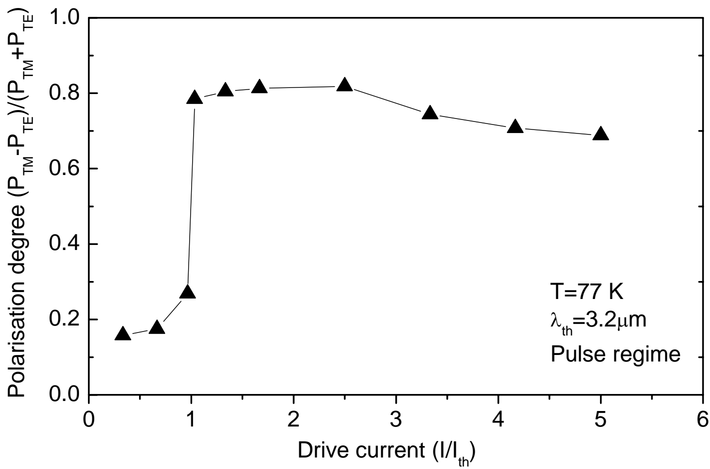


Figure 13. Degree of emission polarisation for the p-n laser structure as a function of the driving current divided by the threshold current

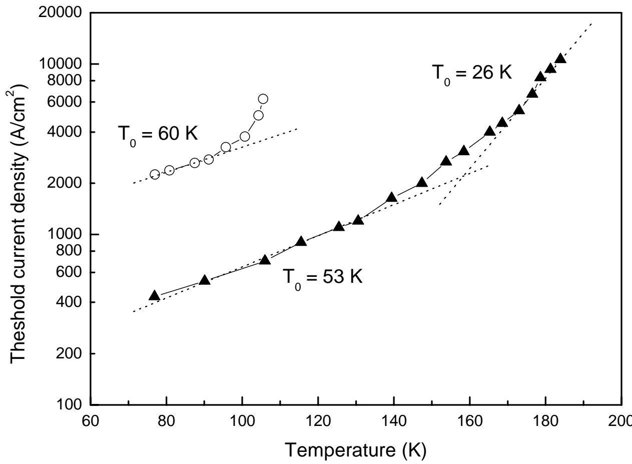


Figure 14. Temperature dependence of threshold current density for two kinds of Type II broken-gap lasers with p-p *(open points)* and p-n (*solid points*) active regions (see Figures 5 and 8)

4.3 Prospects for High-temperature Luminescence in Type II Broken-gap Heterojunctions

The Type II P-$Ga_{0.84}In_{0.16}As_{0.22}Sb_{0.78}$/n-$In_{0.83}Ga_{0.17}As_{0.80}Sb_{0.20}$ single heterostructure has been shown to manifest intense luminescence at room temperature. The EL spectra for the structure contained two emission bands (A and B) in the spectral range 0.25-0.45 eV (Figure 15). The most noticeable difference compared with the 77 K spectra is that the relative intensity of the two bands is reversed at 300 K. At room temperature, band A, with a peak photon energy of 0.320 eV and a FWHM of 26 meV, was narrower and more intense than band B ($h\nu_B$=0.355 eV and FWHM=68 meV).

The intensity of both emission bands depends linearly on drive current at 300 K. The intensity of band B decreased by a factor of seven, while the intensity of the emission band A decreased by only a factor of two between 77 and 300 K. Such a small decrement in intensity with increasing temperature (over the same temperature range) was recently observed for radiative transitions between occupied electron and hole sub-bands in a single AlSb/InAsSb/AlSb quantum well [27]. We suggest that the emission band A can be attributed to the interface-induced radiative transitions at the heteroboundary while the high-energy band B can be attributed to bulk radiative recombination in the n-$In_{0.83}Ga_{0.17}As_{0.80}Sb_{0.20}$ solid solution. In contrast to the low-temperature EL spectra, emission band A dominates at room temperature. Thus, we believe that the Type II heterojunction with a slight broken-gap alignment and large asymmetric band-offsets at the heteroboundary could be a promising structure for the active region in light-emitting diodes and lasers operating in the mid-infrared at high temperatures.

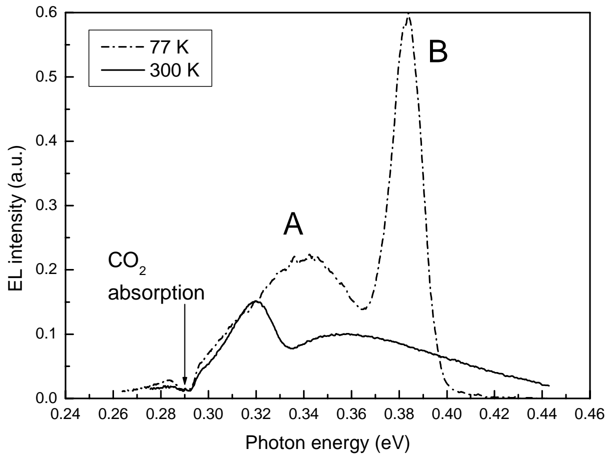


Figure 15. Comparison of the EL spectra (150 mA drive current) from the Type II heterostructure at 77 and 300 K. *Arrows* indicate the atmospheric CO_2 absorption region (4.25-4.27 μm).

It is important to note that the proposed physical approach simplifies laser design in comparison with Type II superlattice lasers or quantum cascade lasers, although this laser structure can be cascaded as well. To improve laser performance and extend the lasing up to room temperature, further optimisation of the laser structure design is needed using multi-quantum well structures with asymmetric band offset confinement. Theoretical modelling of multiple quantum well lasers with tunnelling injection and tunnelling carrier transport between quantum wells has been considered by Kucharczuk *et al.* [28].

5 Conclusions

The fundamental feature of the Type II broken-gap GaInAsSb/InAs(GaSb) heterojunction is the partial overlapping of the conduction band of the n-InAs like alloy with the valence band of the p-GaSb like epilayer. In such a heterostructure electrons and holes are spatially separated and localised in self-consistent quantum wells formed on both sides of the heterointerface. This leads to unusual tunnelling-assisted radiative recombination and novel transport properties.

Direct evidence for suppression of Auger-recombination was obtained for a novel tunneling-injection laser based on this concept and containing a Type II p-GaInAsSb/n-InGaAsSb heterojunction within the active region. It has been shown that lasers with improved temperature characteristics and a lower threshold current can be achieved using the p-n heterojunction rather than the p-p heterojunction in the active region. One can conclude that suppression of Auger-recombination leads to a weaker temperature dependence of the threshold current and a higher quantum

efficiency in the long-wavelength laser structures based on these Type II heterojunctions.

To improve laser performance and extend the lasing up to room temperature further optimisation of laser design is needed using multi-quantum well interface structures with asymmetric band offset confinement. Theoretical modelling of multiple quantum well lasers with tunnelling injection and tunnelling carrier transport between quantum wells was considered in [28]. It is important to note that the physical approach proposed in the current chapter simplifies the laser design compared to Type II superlattice lasers or quantum cascade lasers, and also that this laser structure could be cascaded as well.

References

1. Kroemer H and Griffiths G, 1983 *Electr. Dev. Lett.* **4** 20
2. Baranov A N, Dzhurtanov B E, Imenkov A N, Rogachev A A, Shernyakov Yu M and Yakovlev Yu P 1986 *Sov. Phys. Semicond.* **20** 2217
3. Kop'ev P S, Ledentsov N N, Mohakhov A M and Rogachev A A 1991 Russian patent No 2019895
4. Baranov A N, Betru N, Cuminal Y, Boissier G Alibert C and Joullie A 1997 *Appl. Phys. Lett.* **71** 735
5. Mikhailova M P, Zegrya G G, Moiseev K D and Yakovlev Yu P 1996 *Solid-State Electron.* **40** 673
6. Mikhailova M P, Moiseeev K D, Ershov O G, Zegrya G G and Yakovlev Yu P 1996 *Semicond.* **30** 223
7. Kazarinov R and Lakhtman B 1994 *Private Communication*
8. Mikhailova M P, Andreev I A, Voronina T I, Lagunova T S, Moiseev K D and Yakovlev Yu P 1995 *Semicond.* **29** 678
9. Mikhailova M P, Voronina T I, Lagunova T S, Moiseev K D, Ankoudinov V V, Obukhov S A, Yakovlev Yu P and Titkov A N 1995 *Proc. 3rd Int. Symp. Nanostructures: Physics and Technology 2003 St.Petersburg, Russia* 82
10. Moiseev K D, Faleev N N, Sitnikova A A and Yakovlev Yu P 2000 *Semicond.* **34** 1381
11. Berezovets V A, Mikhailova M P , Moiseev K D, Parfeniev R V, Yakovlev Yu P and Nizhankovski V I 2003 *phys.stat.solidi (a)* **195** 194
12. Moiseev K D, Zeman J, Sadowski M L, Martinez G, Berezovets V A, Brunkov P N, Falko V I, Mikhailova M P, Parfeniev R V and Yakovlev Yu P *Proc. 11th Nanostructures: Physics and Technology*, Saint Petersburg, June 23-28 2003, 216
13. Moiseev K D, Berezovetz V A, Mikhailova M P, Nizhankovski V I, Parfeniev R V and Yakovlev Yu P 2001 *Surf. Sci.* **482** 1083
14. Voronina T S, Lagunova T S, Mikhailova M P, Moiseev K D, Rozov A E and Yakovlev Yu P 1998 *Semicond.* **32** 215
15. Hoffman C A, Meyer J R, Youngdall E R, Bartoli T J, Miles R H and Ram-Mohan L R 1994 So*lid-St. Electr.* **37** 1203
16. Mikhailova M P, Moiseev K D, Timchenko I N, Sherstnev V V and Yakovlev Yu P 1995 *Semicond.* **29** 358
17. Bazhenov N L, Zegrya G G, Ivanov-Omskii V I, Mikhailova M P, Moiseev K D, Smirnov V I, and Yakovlev Yu P *Proc. of 23rd Int Symp. on Compound Semicond. ISCS-96, St.Petersburg, Russia 1996* 1997 *IOP Publish.* 1025
18. Bazhenov N L, Zegrya G G, Mikhailova M P, Moiseev K D, Smirnov V A, Mikhailov M Yu, Solovjeva O Yu and Yakovlev Yu P 1997 *Semicond.* **31** 560

19. Mikhailova M P, Moiseeev K D, Ershov O G, Zegrya G G and Yakovlev Yu P 1996 *Semicond.* **30** 223
20. Zegrya G G and Andreev A D 1995 *Zh. Expr. Teor. Fiz.* **109** 615
21. Zegrya G G, Mikhailova M P, Danilova T N, Imenkov A N, Moiseeev K D, Sherstnev V V and Yakovlev Yu P 1999 *Semicond.* **33** 97
22. Moiseev K D, Krier A, Mikhailova M P and Yakovlev Yu P, 2002 *J. Phys. D Appl. Phys.* **35** 631
23. Moiseev K D, Mikhailova M P and Yakovlev Yu P, 2003 *Semicond.* **37** 1010
24. Moiseev K D, Mikhailova M P, Ershov O G, Zegrya G G and Yakovlev Yu P 1997 *Phys. Tech. Lett.* **23** 151
25. Moiseev K D, Mikhailova M P and Yakovlev Yu P 2003 *Physica E* **20** 491
26. Averkiev N S, Baranov A N, Rogachev A A and Yakovlev Yu P 1992 *Sov. Tech. Phys. Lett.* **26** 1971
27. Moiseev K D, Astakhova A P, Zegrya G G, Mikhailova M P, Yakovlev Yu P, Hulicius E, Hospodkova A, Pangrac J and Simecek T *Proc. 12th Nanostructures: Physics and Technology*, St. Petersburg, June 21-25 2004, 58
28. Kucharczuk M, Wartak M S Weetman and Lau P K 1999 *J. Appl. Phys.* **86** 3218

IV-VI Semiconductors for Mid-infrared Optoelectronic Devices

P. J. McCann
School of Electrical and Computer Engineering
University of Oklahoma, Norman, Oklahoma 7301

1 Introduction

Use of IV-VI semiconductors for solid state electronic device fabrication began in 1874 with Ferdinand Braun's report on electrical rectification with natural galena (PbS) crystals [1]. This was followed in the early twentieth century by use of PbS rectifiers in crystal radio sets. Today, IV-VI semiconductors are used primarily in optoelectronic devices designed for detection and emission of mid-infrared electromagnetic radiation. Technologically important IV-VI semiconductors include binary PbS, PbSe, and PbTe and pseudobinary ternary and quaternary alloys such as PbSnSe, PbSnTe, PbSrSe, PbEuTe, and PbSnSeTe. Depending on alloy composition and temperature their bandgaps range from 0 to over 500 meV. They are direct gap materials with conduction and valence band minima at the L-point in *k*-space [2]. Such properties make IV-VI semiconductor materials useful for fabrication of optoelectronic devices covering the 3 μm to 30 μm spectral range. For example, Figure 1 shows mid-IR laser emission data for IV-VI semiconductor double heterostructure (DH) lasers grown on (100)-oriented PbSe and PbTe substrates [3]. Each laser can cover as much as 200 cm^{-1} of the mid-IR spectrum while exhibiting single-mode tuning regions as wide as 3 cm^{-1} as shown in Figure 1(b). With continuous wave emission powers as high as 2 mW, these widely tunable lasers are useful light sources for high spectral resolution mid-IR spectrometers where the primary application is measurement of specific trace gas molecules.

2 Spectroscopy with IV-VI Semiconductor Lasers

A potentially significant application for mid-IR laser spectrometers is in the area of medical diagnostics where sensitive measurement of specific biomarker molecules can help diagnose diseases and monitor therapy. Such an application was recently demonstrated with IV-VI semiconductor lasers where real-time measurements of nitric oxide (NO) at parts-per-billion (ppb) levels are measured in exhaled breath samples [4]. Exhaled nitric oxide (eNO) measurements have clinical utility because they can be used to assess airway inflammation [5]. Such a clinical capability, if widely used to ensure appropriate anti-inflammatory therapy, is expected to improve the health of asthmatics since it is well established that untreated airway inflammation results in more life-threatening acute asthma attacks, hospitalizations, and long term lung tissue damage [6].

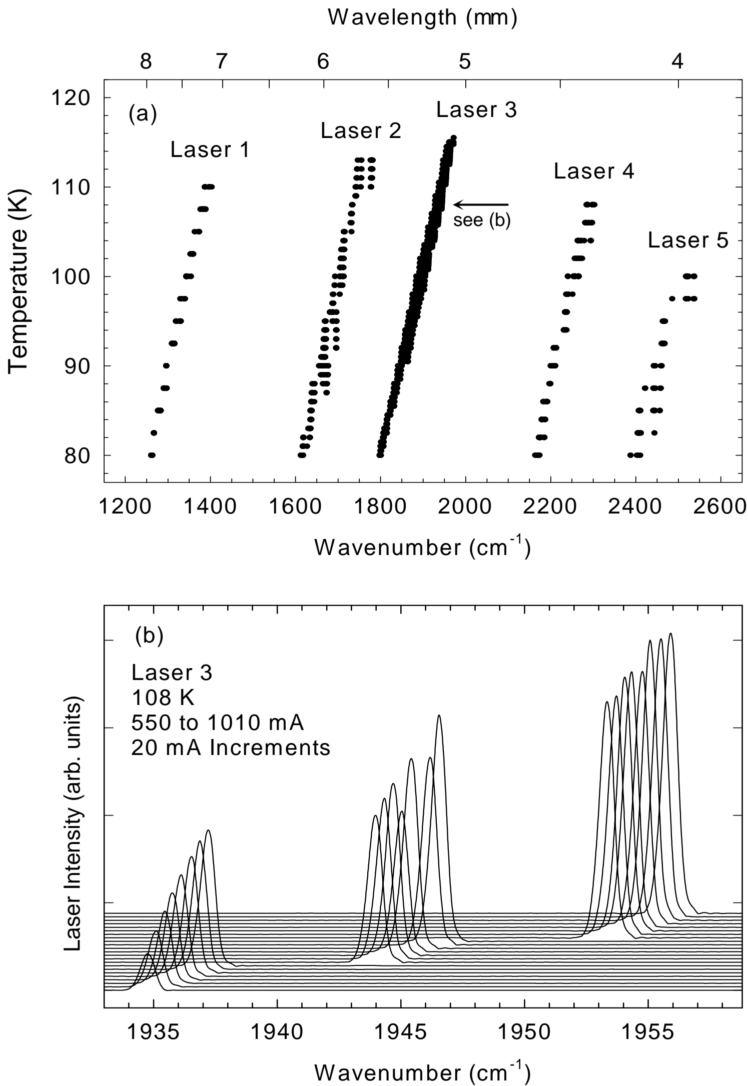


Figure 1. (**a**) Laser emission data for five different double heterostructure (DH) IV-VI semiconductor lasers. Each dot represents single-mode emission at a specific heat sink temperature and injection current. (**b**) Laser emission intensity for Laser 3 at a heat sink temperature of 108 K (see arrow in (**a**)) for a range of injection currents. Mode hops between regions of single mode emission are typical for these edge-emitting Fabry-Perot lasers. The spectral linewidth of each single-mode emission is typically less than 0.001 cm^{-1}. Note that the spectral width appears much wider due to the 0.5 cm^{-1} resolution of the Fourier transform infrared (FTIR) spectrometer that was used for the measurements.

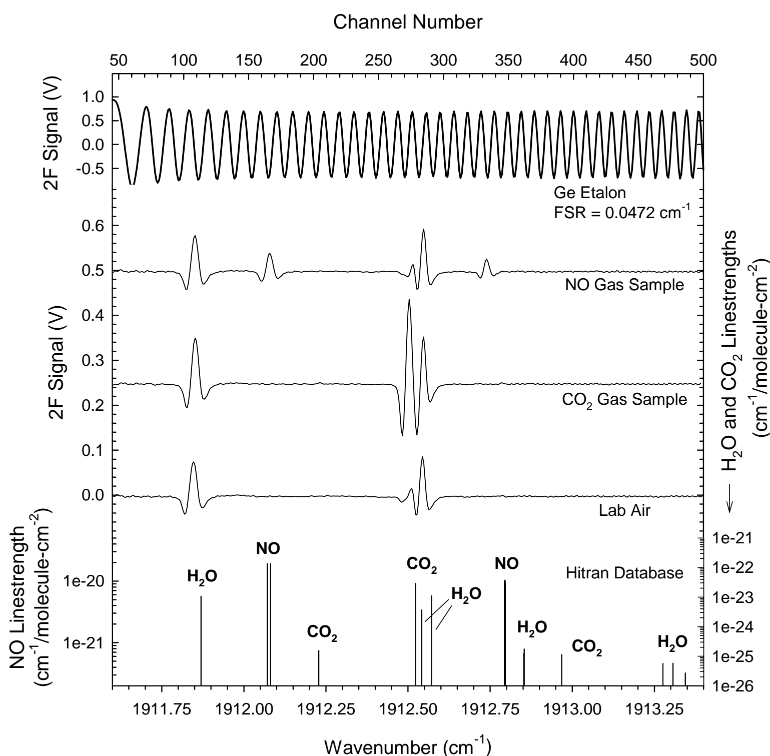


Figure 2. Second harmonic laser absorption spectra in the 5.2 μm spectral range and absorption line strengths from the Hitran database for NO, CO_2 and H_2O. *Top spectrum* is for a Ge etalon showing mode-hop-free single-mode tuning of over 1.5 cm^{-1}. Spectra for three different gas samples, dilute NO (ppm level), CO_2 (% level), and lab air, show the simultaneous and selective detection capabilities of a widely tunable mid-IR laser.

Figure 2 shows measured laser absorption spectra in the 5.2 μm spectral range using Laser 3 (shown in Figure 1). At the top is a spectrum for a germanium etalon, which has a free spectral range of 0.047 cm^{-1}. Its uniformity and large number of fringes confirm a single mode tuning range of more than 1.5 cm^{-1}. Absorption spectra for three gas samples that contain (1) dilute NO (ppm level) in N_2, (2) CO_2 (% level), and (3) lab air are displayed. Identification of individual absorption features are made with the aid of rotational-vibrational line positions for NO, CO_2, and H_2O found in the Hitran database, shown at the bottom. (Due to the non-linear tuning characteristics of the laser, the Hitran line positions do not line up exactly with the measured absorption peaks). Ability to detect independently and simultaneously NO and CO_2 without interference from water and other molecules is an important feature since the signal for CO_2 can serve as an internal reference for calculating eNO concentrations in exhaled breath samples. Earlier eNO research employing chemiluminescence instrumentation does not offer this self-calibrating capability. More reliable laser-based eNO measurements using the patient's own exhaled CO_2 (eCO_2) as the calibration gas is a potentially important advance that can

facilitiate the adoption of exhaled breath testing in clinical settings where routine instrument calibrations are not desired.

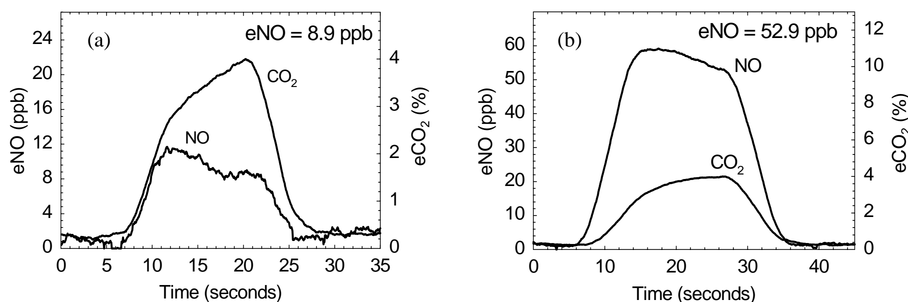


Figure 3. (**a**) Exhaled breath trends for NO and CO_2 from an asthmatic patient treated with anti-inflammatory medication. (**b**) Exhaled breath trends for NO and CO_2 from an untreated asthmatic patient. Note the much smaller eNO concentration of 8.9 ppb for the treated asthmatic *vs* the 52.9 ppb value for the untreated asthmatic. Reprinted with permission of the Electrochemical Society.

Figure 3 shows measured eNO and eCO_2 concentrations *vs* time for exhalations from two different subjects who participated in a clinical research project approved by the University of Oklahoma Health Sciences Center Institutional Review Board (IRB). The trends in Figure 3(a) were from a five-year-old mild-persistent asthmatic female treated with inhaled corticosteroid anti-inflammatory therapy. The eNO concentration at end-tidal, which is determined by the peak in the eCO_2 trend and represents breath sample from the lower airway, is 8.9 ppb, a level that is in the normal range for non-asthmatics or properly-treated asthmatics. The trends in Figure 3(b) were from a 14-year-old mild-persistent asthmatic who was not treated with anti-inflammatory medication. There is a clear difference observed in that the untreated asthmatic has a much larger eNO concentration of 52.9 ppb. In both cases a larger eNO concentration is observed initially in the trend, and this is due to contribution from the nasal cavity, which is known to have a larger NO concentration than the lower airway.

The field of exhaled biomarker molecule breath testing is still in its infancy. The United States Food and Drug Administration has only recently (May 1, 2003) cleared eNO measurements as a valid way to monitor anti-inflammatory therapy for asthmatic patients. Clearly, further improvements in mid-IR laser technologies, including IV-VI semiconductor materials and devices, will help make such measurements and associated clinical research easier to perform than with what is presently possible using laborious and complex gas chromatography mass spectroscopy (GC/MS) or more limited chemiluminescence techniques. Although prior GC/MS research has provided some guidance on which biomarkers might be linked to certain diseases or health conditions, results from GC/MS measurements should not be considered conclusive since there are many potentially important exhaled biomarker molecules that are not detected using standard GC/MS techniques. For example, nitric oxide, due to its high reactivity, does not survive the

high temperature GC column, so its link to airway inflammation would never have been discovered using GC/MS. Exhaled biomarkers that can be readily measured using mid-IR laser absorption spectroscopy methods include ammonia (NH_3) for monitoring renal function [7], carbonyl sulfide (COS) for monitoring liver function and organ rejection [8,9], carbon disulfide (CS_2) for diagnosing schizophrenia [10], and formaldehyde (CH_2O) and ethane (C_2H_6) for detecting cancer [11,12]. Continued work on the development of mid-IR laser spectroscopy instruments for breath analysis offers the prospects of clinically useful breath tests that are as versatile as blood tests but much less invasive to perform.

3 IV-VI Semiconductor Growth and Characterization

Presently, the primary problem that needs to be solved to improve chemical sensors based on IV-VI semiconductor lasers is elimination of the requirement to cool the laser to cryogenic temperatures. Continuous wave (CW) operation, which is required for high sensitivity chemical sensing, typically requires a heat sink temperature of 120 K or less. Such temperatures can be achieved using single-stage closed-cycle refrigerators or Stirling cycle coolers, but they are bulky and costly. An increase of more than 100 degrees in maximum operating temperature would allow IV-VI lasers to be cooled with multi-stage thermoelectric cooling modules. Two approaches can be taken to achieve this increase: (1) reduce laser threshold current and (2) enhance active region heat dissipation. The demonstration of CW operation of a buried heterostructure IV-VI semiconductor laser at a heat sink temperature of 223 K illustrates the kind of improvements that are possible when attention is devoted to such issues [13]. Further reductions in laser thresholds and significant increases in active region heat dissipation are possible with IV-VI semiconductor laser structures, and specific ways to accomplish these objectives will be discussed in Sections 5 and 6. In each case the improvements are made possible by the growth of IV-VI heterostructures using novel combinations of materials, which is discussed in this section.

Figure 4 is a band gap *vs* lattice parameter plot for IV-VI semiconductors and related materials. Silicon, CaF_2, and BaF_2 are shown along with IV-VI binary, ternary, and quaternary materials. Silicon with fluoride buffer layers as well as bulk BaF_2 can serve as substrates for growth of epitaxial layers, a capability that has facilitated the development of novel IV-VI semiconductor epitaxial layer structures. General procedures involve growth of the binary compound PbSe, which has a bandgap of 175 meV at 77 K (278 meV at 300 K), along with either larger band gap PbSrSe [14] or smaller band gap PbSnSe [15]. In addition, it is also possible to grow pseudobinary ternary and quaternary alloys that are lattice matched with BaF_2 such as PbSeTe [16] and PbSnSeTe [17] as indicated in Figure 4. A summary of the issues involved with growth of IV-VI semiconductors on BaF_2 and silicon substrates is provided below.

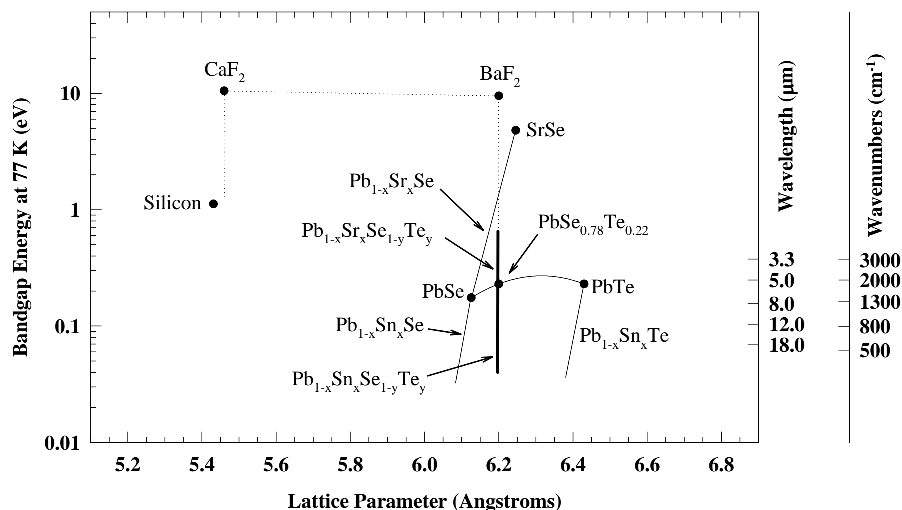


Figure 4. Band gap (and corresponding wavelength/frequency) *vs* lattice parameter plot for IV-VI semiconductors and related materials. Binary and pseudobinary ternary and quaternary alloys can be grown that cover the 3 μm (~3000 cm^{-1}) to 20 μm (500 cm^{-1}) spectral range where many molecules have strong vibrational absorption bands.

3.1 IV-VI Layers on BaF_2

Hot wall epitaxy (HWE), molecular beam epitaxy (MBE), and liquid phase epitaxy (LPE) have been used to grow IV-VI compounds and alloys on both (111)- and (100)-oriented BaF_2 substrates. Most work, however, has been done with growth on cleaved (111) BaF_2 surfaces. Review articles by Lopez-Otero [18] and Holloway and Walpole [19] summarize the early work on HWE and MBE growth of IV-VI layers on BaF_2. Good thermal expansion and lattice parameter matches allow growth of high quality epitaxial layers, but fabrication of edge emitting lasers is complicated by the inability to obtain perpendicular cleaved facets due to the {111} preferential cleavage habit of the BaF_2 substrate. Recent work on developing vertical cavity surface emission lasers (VCSELs) is motivated in part because of the problems with fabricating in-plane lasers using IV-VI/fluoride structures grown on (111) BaF_2 [20,21]. Another approach to laser fabrication that works around the cleaving problem and can enable development of edge-emitting lasers will be described in Section 6 of this chapter.

A number of fundamental materials studies have been performed with IV-VI layers grown on BaF_2, a substrate that facilitates electrical and optical characterization because of its insulating and infrared transparent properties. Examples include doping studies of lattice-matched $PbSe_{0.78}Te_{0.22}$ ternary layers grown by LPE on (100) BaF_2 showing that undoped layers were n-type, as expected due to group VI vacancies caused by growth from Pb-rich liquid solutions, and that

thallium was an effective acceptor impurity [22]. In addition, infrared transmission spectra collected using FTIR spectroscopy showed that absorption edges of lattice-matched three-layer DH laser structures on (100) BaF_2 at 130 K shifted from 195 meV (6.4 μm) to 143 meV (8.7 μm) as tin content in the growth solution for the $Pb_{1-x}Sn_xSe_{1-y}Te_y$ quaternary middle layer was increased from 0 to 12% [23]. This absorption edge energy dependence on Sn content is similar to what has been observed for laser emission energy measurements from DH lasers made with LPE-grown $Pb_{1-x}Sn_xSe_{1-y}Te_y$ quaternary alloys lattice matched with PbSe substrates [24]. Room-temperature FTIR transmission spectra for $Pb_{1-x}Sr_xSe$ layers grown by MBE on (111) BaF_2 substrates showed an absorption edge shift from 277 meV to 716 meV as the Sr/PbSe flux ratio was increased from 0 to 10% [25]. The spacings of below band-gap Fabry-Perot interference fringes were also used to determine refractive indices, which decreased from 4.8 to 3.8 as the Sr/PbSe MBE flux ratio was increased from 0 to 10%.

More recent work has focused on MBE growth and characterization of PbSrSe/PbSe multiple quantum well (MQW) structures. As will be discussed in Section 5, these structures have unique properties that make them particularly well suited as active region materials for mid-infrared lasers. Detailed procedures for MBE growth of PbSrSe/PbSe MQWs on (111)-oriented BaF_2 are described elsewhere [25]. Important results from this work include differential transmission spectroscopy (DTS) measurement of quantized energy levels in square and parabolic MQWs [26,27]. In both cases removal of L-valley degeneracy is clearly observed, and it is shown that the quantum size effect (*i.e.* differences in normal and oblique valley effective masses) rather than epitaxial layer strain is responsible for L-valley degeneracy splitting for QWs thinner than 30 nm. Figure 5 shows DTS spectra at 4 K and 70 K for a 15 period MQW with a well thickness of 20.6 nm, (a), and depiction of n=1 L-valley energy levels in (111)-oriented QWs, (b). The lowest energy interband transition, $(1\text{-}1)^N$, is between heavier effective mass electron and hole states in L-valleys that are normal to the (111) plane in reciprocal space, while the next lowest transition, $(1\text{-}1)^O$, is between lighter effective mass electron and hole states in the other three oblique valleys that lie along the $[\bar{1}11]$, $[1\bar{1}1]$, and $[11\bar{1}]$ directions. The larger total density of states for the three-fold degenerate (neglecting spin) oblique valleys is reflected in the relatively larger absorption peaks for the $(1\text{-}1)^O$ and $(2\text{-}2)^O$ transitions as compared to the $(1\text{-}1)^N$ and $(2\text{-}2)^N$ transitions, respectively, in the 4 K DTS spectrum (see Figure 5(a)). Interband transitions as high as n=6 have been observed in DTS spectra of MQW samples with wide QWs, and this has allowed determination of the band non-parabolicity parameter for PbSe [27].

A number of photoluminescence (PL) measurements of IV-VI layer structures grown on BaF_2 substrates have also been performed. Good PL results have been obtained both at cryogenic temperatures and above room temperature. Low temperature (~5 K) PL emission energies from LPE-grown $PbSe_{0.78}Te_{0.22}$ and $Pb_{0.95}Sn_{0.05}Se_{0.80}Te_{0.20}$ layers on (100) BaF_2 were 174 meV (7 μm) and 100 meV (12 μm), respectively [28]. Above room temperature CW PL between 310 meV (4 μm) and 400 meV (3 μm) has also been observed from PbSe/PbSrSe MQW structures grown on (111) BaF_2 substrates by MBE [29]. It should be noted that such room

temperature PL has not been observed from (100)-oriented IV-VI materials. In fact, MQW structures grown at the same time on (111)-oriented BaF_2 and (100)-oriented PbSe substrates exhibited very different PL behavior. The (100)-oriented material on PbSe exhibited measurable CW PL only up to about 120 K, which is about the limit for CW operation of IV-VI lasers on PbSe substrates. By contrast, the (111)-oriented material on BaF_2 exhibited strong CW PL at 328 K, a temperature limited only by the calibrated range of the resistance thermal detector (RTD) used in the experimental setup. Possible explanations for this observed difference include lower PbSe thermal conductivity, which leads to much more epilayer heating that quenches PL, and/or band structure effects associated with L-valley degeneracy removal, which occurs in (111)-oriented QWs but not in (100)-oriented QWs. This latter effect will be discussed in more detail in Section 5.

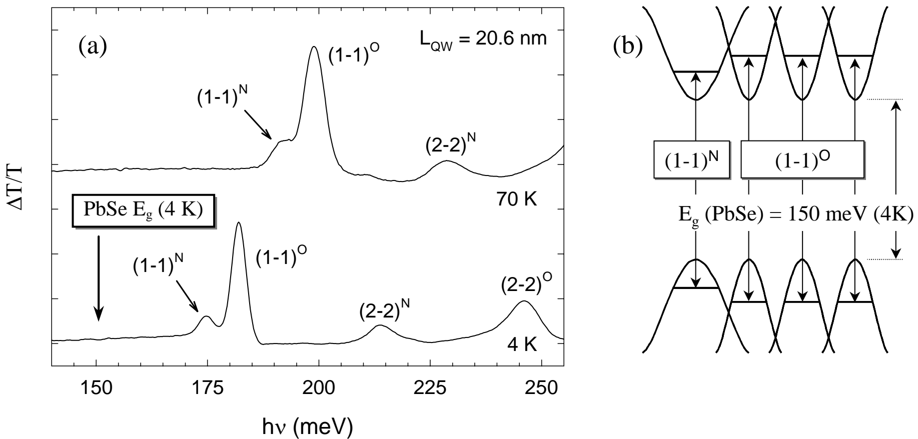


Figure 5. (**a**) Differential transmission spectra at 4 K and 70 K for a 15 period PbSrSe/PbSe MQW structure with 20.6 nm thick PbSe wells. Quantum confinement in the [111] direction results in two different valley types, normal and oblique, and two different effective masses for electrons and holes. Two different sets of quantized energy levels, depicted in (**b**), are thus observed in the DTS spectra.

Figure 6 is a plot of PL peak emission energy, as determined from Gaussian fits to the measured spectra (Fabry-Perot interference fringes dominated the spectra, see Figure 14), as a function of PbSe QW width. The solid lines are fits to data obtained at heat sink temperatures of 5 and 55°C, and the different symbols indicate measured PL energies at temperatures in between these values. The quantum size effect along with the large IV-VI semiconductor bandgap temperature-tuning coefficient of about 0.4 meV/K are clearly observed in these data. Note that this large change in band gap with temperature is partially responsible for the wide tuning ranges possible with IV-VI lasers as shown in Figures 1 and 2. Interband radiative recombination between n=1 normal valley states is believed to be the PL mechanism in these materials. It should also be pointed out that these PL data were obtained with very low opitcal pumping powers of less than 2 W/cm^2 and the FTIR spectrometer system used to measure the spectra did not employ any signal enhancing techniques such as

step-scan detection or double modulation indicating that the mid-IR light producing mechanism in these MQW materials is very efficient.

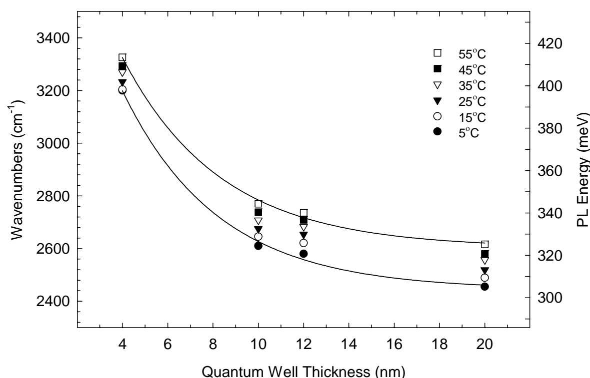


Figure 6. PL energy *vs* QW width for four different PbSrSe/PbSe MQW samples at six different temperatures

3.2 IV-VI Layers on Silicon

High quality IV-VI semiconductor layers can be grown on (111)-oriented silicon substrates using MBE despite the large thermal expansion and lattice parameter mismatches between IV-VI materials and silicon, about 12 and 700%, respectively. Initial work in this area focused on using a stacked BaF_2/CaF_2 buffer layer [30], but more recent growth experiments have shown that just a thin CaF_2 layer produces good PbSe epitaxy [31]. This fluoride buffer layer, even when as thin as 50 Å, appears to be necessary for good PbSe epitaxy since attempts to grow PbSe directly on Si(111) have resulted in poor layer quality [32]. PbSe/CaF_2 interface chemistry has been studied using *in situ* X-ray photoelectron spectroscopy (XPS), and the interface was found to be abrupt and dominated by Pb-F and Ca-Se bonds [33]. In addition, observation of reflection high energy electron diffraction (RHEED) intensity oscillations indicate layer-by-layer growth modes for both CaF_2 and PbSe [31]. After growth the layers are entirely free of cracks and crystalline quality is reasonably good as evidenced by high resolution X-ray diffraction (HRXRD) full-width half maximum (FWHM) values of typically less than 200 arcsec [34].

It is remarkable that crack-free crystalline layers can be obtained when there are large in-plane tensile strains due to the thermal expansion coefficient mismatch with the silicon substrate and the approximately 300°C of cooling from the MBE growth temperature. An explanation for this phenomenon was offered by Müller *et al.* [35] where plastic deformation of the IV-VI layer can occur by movement of dislocations along {100} planes in the <110> directions. Figure 7 shows an optical microscope image of an exposed PbSe growth interface revealed by removal of the silicon substrate through aqueous dissolution of a BaF_2 buffer layer [34]. (Use of substrate removal techniques in laser fabrication will be discussed in Section 6.) The bottom of the PbSe layer is where the in-plane tensile strain is highest. Plastic deformation

of the layer via dislocation glide in the {100}<110> slip system will result in three-fold symmetric step edges on a (111) surface, as is observed in Figure 7. This primary slip system will allow strain relief in IV-VI layers grown on (111)-oriented substrates, but not in (100)-oriented layers because there are no resolved shear stresses along the {100} planes in the <110> directions. Indeed, attempts to grow PbSe on (100)-oriented silicon by MBE at temperatures around 280°C always resulted in high crack densities when samples were cooled to room temperature (LPE growth of PbSe on (100) silicon, however, did result in crack-free layers [36]).

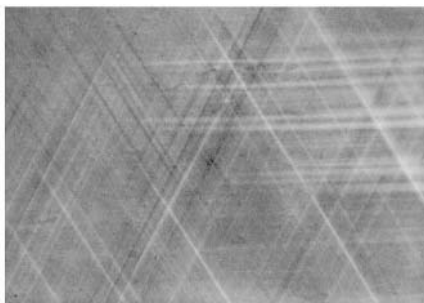


Figure 7. Nomarski optical microscopic image of an exposed MBE-grown PbSe growth interface after lift-off from a $BaF_2/CaF_2/Si$ (111) substrate. The three-fold symmetry straight lines run along the intersections of the {100} glide planes with the (111) growth interface. These features are associated with the plastic deformation strain relaxation mechanism proposed by Müller *et al.* [35]. Reprinted with permission of the Electrochemical Society.

Structural softness, *i.e.* ease of plastic deformation, is a well known property of IV-VI semiconductors and has complicated processing procedures for fabrication of devices made from these materials. But this softness is an advantageous property that enables high crystalline quality epitaxial growth on (111)-oriented silicon substrates. Besides allowing MBE experimentation on large area and low cost substrates, this growth ability creates many opportunities for development of novel device concepts. Monolithic integration of mid-IR optoelectronic devices with silicon circuitry, for example, is made possible by this materials compatibility. In fact, large area focal plan arrays for IR imaging systems have been fabricated from IV-VI semiconductor layers grown on silicon substrates [37]. With increasing interest in using IV-VI materials for fabrication of thermoelectric devices [38] as well as ongoing interest in their use for mid-IR laser fabrication, the ability to grow on silicon is an important capability.

Smaller band-gap $Pb_{1-x}Sn_xSe$ ternary layers were also grown by LPE on Si (100) substrates, but unlike binary PbSe layers [36] they were not entirely crack-free [15]. Solid solution hardening effects are believed to be responsible for this result since HRXRD characterization showed that PbSnSe layers had a smaller elasticity as evidenced by residual in-plane tensile strains of about 0.06% following growth, much lower than the 0.21% obtained for PbSe layers [39]. Crack densities in PbSnSe layers were observed to decrease when LPE growth temperatures were increased. In addition, HRXRD FWHM data showed that PbSnSe crystalline quality

improved as growth temperature was increased [40]. These results show that layers grown at higher temperatures are more able to plastically deform and that slip on one or more secondary glide systems, such as the <110>{110} system, increasingly occurs as growth temperature is raised from 420 to 480°C. High Se vacancy concentration, which increases with LPE growth temperature, can enhance layer plasticity if vacancies agglomerate and form edge dislocations along one or more higher order slip planes. This mechanism may explain why it is possible to obtain crack-free PbSe layers on (100) Si by LPE and not by MBE, where growth occurs at much lower temperatures and vacancy concentrations are smaller. FTIR transmission spectra for $Pb_{1-x}Sn_xSe$ layers grown on (100) Si substrates, nevertheless, showed expected band gap dependence on temperature and alloy composition where $Pb_{1-x}Sn_xSe$ alloys with Sn contents between 0 and 10% could cover the 5 μm (2000 cm^{-1}) to 16 μm (625 cm^{-1}) spectral range. In addition, below band-gap Fabry-Perot interference fringe spacings showed an increase in refractive index from 5.3 to 6.0 as the Sn content in the LPE growth solution was increased from 3 to 10%.

As with IV-VI MQW materials grown on BaF_2 substrates, above room temperature CW PL has also been observed from IV-VI MQWs grown on (111) Si substrates with CaF_2 and BaF_2/CaF_2 buffer layers [41,42]. These results confirm that high quality optically active IV-VI epitaxial layer structures can be grown on silicon. The measured temperature-tuning coefficient of 0.38 meV/K is similar to what was observed for PL emission from MQW layers grown on BaF_2 substrates (see Figure 6). In addtion, blue shifts in the PL energy as the CW pump laser power was increased were observed to be smaller than those observed for MQW strucrtures grown on BaF_2. Gaussian fitted PL peak energies shifted from 337 to 340 meV as the pump laser current was increased from 300 to 800 mA. This corresponds to a localized epilayer heating effect of only 7.8 degrees as compared to an approximate 30 degree heating effect in IV-VI layers on BaF_2, a difference that is due to the higher thermal conductivity of the Si substrate (1.41 W/cm K) compared to that of BaF_2 (0.12 W/cm K).

Work has also been performed on removing Si growth substrates from MBE- and LPE-grown PbSe layers by dissolving fluoride buffer layers in water. This technique was successfully used to transfer 3-μm-thick MBE-grown PbSe epilayers from Si to copper by using PbSn [43] and AuIn [34] bonding metallurgy. Transferred PbSe layers had mirror-like surfaces, and profilometer characterization showed a typical peak roughness of about 22 nm for the growth interface, not much higher than the roughness of the as-grown PbSe surface, about 12 nm. Nomarski microscopy did reveal three-fold slip lines indicating dislocations at the growth interface (see Figure 7), whereas no slip lines were observed on the as-grown PbSe surface. The FWHM of the transferred PbSe layer's HRXRD peak was 160 arcsec, only slightly larger than the 153 arcsec FWHM for the as-grown layer, demonstrating that the transfer process preserves PbSe crystalline quality. Surface chemical analysis of transferred layers by X-ray photoelectron spectroscopy (XPS) also indicated complete removal of the BaF_2 buffer layer. PbSe layers grown by combination MBE/LPE on (100) Si have also been removed from growth substrates and then cleaved along {100} planes using a novel cleaving tool [44]. Recent work has shown that it is also

possible to cleave MBE-grown (111)-oriented IV-VI layers along {110} planes after substrate removal. This ability, which will be described in more detail in Section 6, creates promising prospects for fabrication of edge-emitting lasers from (111)-oriented MQW materials.

4 Self-Heating Effects in IV-VI Mid-IR Lasers

The highest reported operating temperatures for electrically pumped IV-VI mid-IR lasers are 223 K in CW mode for a PbEuSeTe/PbTe buried heterostructure (BH) laser [13] and 333 K in pulsed mode for a PbSrSe/PbSe DH laser [45]. Recently developed optically pumped vertical cavity surface emitting lasers (VCSELs) made from IV-VI materials have exhibited similar performance characteristics with CW laser emission at 230 K [21] and pulsed laser emission at 317 [46] and 310 K [47]. The large difference between CW and pulsed operating temperatures shows that there is a significant self-heating effect in the active region of these devices. Poor active region heat dissipation is particularly severe in IV-VI mid-IR lasers because of the low thermal conductivities of the IV-VI semiconductor or barium fluoride substrate materials used in device fabrication (see Table 1). It should be noted that self-heating is a universal problem in all semiconductor lasers. The effect limits maximum CW operating temperatures in mid-IR lasers and maximum output powers in visible and UV lasers [48].

Table 1. Thermal expansion coefficients and thermal conductivities for various IV-VI mid-IR laser fabrication materials

	Thermal expansion coefficient (10^{-6}/K at 300K)	Thermal conductivity (watt/cm K)
PbSe	19.0	0.018
PbTe	19.8	0.023
BaF_2	19.8	0.117
Silicon	2.6	1.412
Copper	17	3.98
Silver	18	4.27
Diamond	2.0	>13.0

Two approaches to solving the self-heating problem involve (1) reducing laser thresholds (injection current or optical pumping level) and (2) improving active region heat dissipation. Higher quality materials along with appropriate bandgap engineering designs can enable lower laser thresholds. In addition, novel quantum well designs that reduce electron lifetime in the lower laser transition state and/or increase electron lifetime in the upper laser transition state can reduce thresholds for population inversion. Such electron lifetime engineering has been demonstrated for Type I quantum cascade (QC) lasers where electrophonon resonance involving longitudinal optical (LO) phonons is used to depopulate rapidly the ground state laser transition [49]. Similar electron lifetime engineering is also possible with IV-

VI semiconductor QW materials and will be discussed in more detail in the next section. Regarding the latter approach, more efficient heat sinking would enable higher CW operation temperatures for IV-VI mid-IR lasers. A new laser fabrication method that involves growth substrate removal will be described in Section 6. Experimental results showing improved active region heat dissipation for epilayer structures transferred to copper mounts will also be presented.

5 Electrophonon Resonance in PbSe QWs

Unlike III-V materials, IV-VI materials have nearly symmetric conduction and valence bands at the band gap minimum, which is at the four equivalent L-points in the Brillouin zone. In addition, there is no degenerate heavy hole band. Such differences create distinctly different quantization effects in IV-VI QW materials as compared to III-V QW materials. The L-point band gap, for example, produces a situation where the four-fold L-valley degeneracy can be removed, which has been shown to occur when IV-VI QW structures are grown on (111)-oriented substrates [26,27] (see Figure 5). This new knowledge allows development of an accurate model for quantized energy levels in IV-VI QW materials based solely upon application of Schrödinger's equation.

Electrons and holes confined in PbSe quantum wells grown on (111) surfaces can exist in two different types of L-valleys. One L-valley is normal to the (111) plane in reciprocal space, while the other three lie along the $<11\bar{1}>$ directions at oblique angles to the (111) plane in reciprocal space. Charge carriers with two different effective masses are thus possible,

$$m_{111}^{\mathrm{N}} = m_l \qquad \text{and} \qquad m_{11\bar{1}}^{\mathrm{O}} = \frac{9 m_l m_t}{8 m_l + m_t} \tag{1}$$

for the normal and oblique valleys, respectively, where m_l and m_t are the longitudinal and transverse effective masses associated with the prolate ellipsoids of revolution that define the constant energy surfaces for electrons and holes in IV-VI materials [50]. The four different effective masses for normal and oblique valley electrons and holes for (111)-oriented PbSe quantum wells are $m^{N}_{e} = 0.0788\ m_o$, $m^{O}_{e} = 0.0475\ m_o$, $m^{N}_{h} = 0.0764\ m_o$, and $m^{O}_{h} = 0.0408\ m_o$. There will thus be two degeneracy-split energy levels in each band for each principle quantum number, and the amount of splitting will depend on QW width and depth.

Figure 8 shows conduction band L-valley degeneracy splitting energies for n=1 energy levels as a function of QW width for various well depths. The range of widths from 1 to 30 nm and depths from 25 to 175 meV are practical values for MBE-grown PbSrSe/PbSe/PbSrSe QW structures where the corresponding strontium content in the $Pb_{1-x}Sr_xSe$ barrier material varies from about 1 to about 8% [14], a sufficiently low concentration that allows growth of high crystalline quality material and permits efficient doping with donor and acceptor impurities. In addition, keeping well depths below 175 meV, where deviations from band parabolicity are less than 10% (see Figure 6 in [27]), makes the use of constant effective masses for each L-valley a reasonable assumption. The largest conduction

band splitting energy is 19 meV for a 4 nm wide and 175 meV deep QW. Valence band splitting energies are slightly larger due to a greater difference in normal and oblique valley effective masses. The largest valence band splitting energy, for example, is 24 meV for a 4.0 nm wide and 175 meV deep QW. The inset in Figure 8 shows an energy level diagram for a 6.5 nm wide and 65 meV deep PbSe quantum well with calculated energy levels for n=1 electrons showing an L-valley degeneracy splitting value of $\Delta E_{1c} = E^{O}_{1c} - E^{N}_{1c} = 6.3$ meV. It was found that splitting energies exhibit a maximum value when the quantized energy levels are equal to about one half of the well depth.

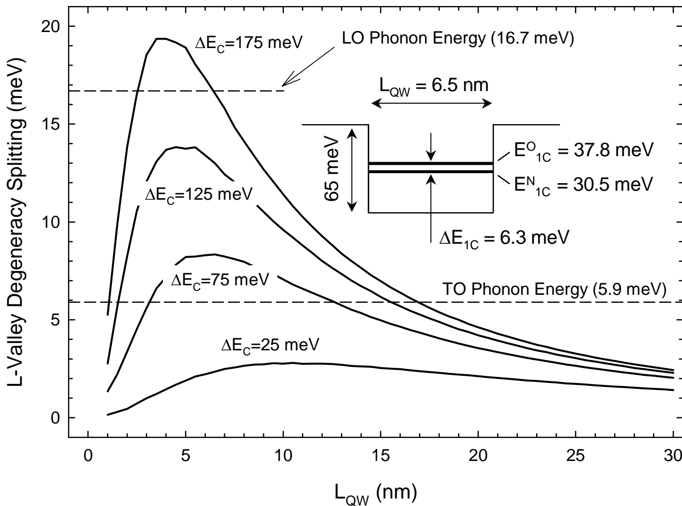


Figure 8. Conduction band L-Valley degeneracy splitting energies for n=1 electrons as a function of quantum well width for different quantum well depths. *Inset* shows example of degeneracy splitting for a 6.5 nm wide and 65 meV deep well. PbSe LO and TO phonon energies are also indicated. Reprinted with permission of the Electrochemical Society.

An important outcome of this energy level analysis is the fact that there are many combinations of well widths and depths at which the L-valley degeneracy splitting energy is equal to either the PbSe longitudinal optical ($\hbar\omega_{LO}$=16.7 meV) or transverse optical ($\hbar\omega_{TO}$=5.9 meV) phonon energy [51]. This is significant because it makes it possible to engineer the QW such that electrophonon resonance takes place between electron energy levels in the conduction and valence bands. Electrophonon resonance dramatically reduces the electron lifetime in the upper state and can be used to design lasers with reduced thresholds for population inversion and laser emission. Such designs are now an integral part of mid-IR and THz QC laser fabrication where a three-level energy structure is created with the bottom two energies spaced by 36 meV, the LO phonon energy of the GaAs QW material [49,52]. In the case of (111)-oriented PbSe QWs, due to the nearly mirror images of the conduction and valence bands, it is possible to create a four-level energy structure. This allows rapid filling of the upper laser transition state along with rapid emptying of the bottom laser transition state thus helping to reduce

further the threshold for population inversion. Examples of efficient four-level laser systems include HeNe gas and Nd:YAG solid state lasers.

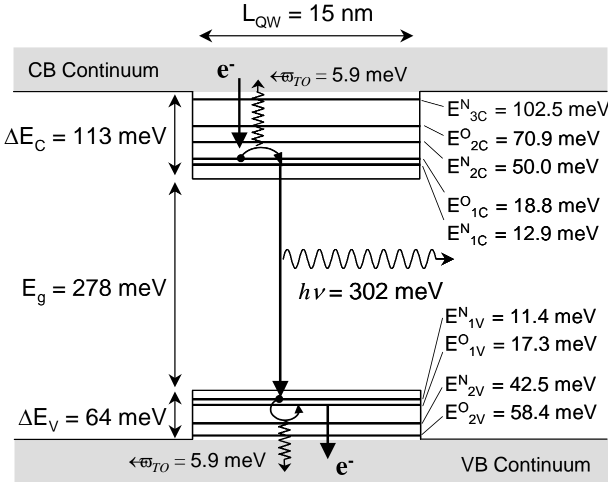


Figure 9. Energy level structure for a 15 nm wide PbSe quantum well with conduction and valence band depths of 113 and 64 meV. Energy levels for degeneracy split n=1 and n=2 states are indicated where the energy spacing between the n=1 states in the conduction and valence bands are equal to the PbSe TO phonon energy. Electrophonon resonance facilitates rapid filling of the E^N_{1C} state and rapid emptying of the E^N_{1V} state when excited electrons relax back to the valence band continuum. Such an energy level structure can reduce the threshold for population inversion between the E^N_{1C} and E^N_{1V} states and allow efficient emission of photons. Reprinted with permission of the Electrochemical Society.

Figure 9 shows an energy level diagram for a PbSe QW with a width of 15 nm and conduction and valence band well depths of 113 and 64 meV, respectively. The total band gap difference for this QW is 177 meV, which is close to the room temperature band gap difference of 173 meV between the PbSe (E_g=278 meV) well and $Pb_{0.93}Sr_{0.07}Se$ (E_g=451 meV) barrier materials used in prior experimental studies [25-27,29,41,42]. The band edge discontinuity ratio ($\Delta E_C/\Delta E_g$) of 63% is also consistent with prior analysis of PbSrSe/PbSe [14] and PbEuTe/PbTe heterostructures [53]. Electrons excited either optically (as in PL) or electrically (as in a diode laser) will decay from the conduction band continuum into quantized 2D electron gas states in the PbSe QW. Electron transitions between the E^O_{1C} and E^N_{1C} conduction band states are expected to be fast due to resonance with the lattice and emission of 5.9 meV TO phonons. Direct interband transitions between the two normal valley states, E^N_{1C} and E^N_{1V}, will be radiative as indicated by the emission of 302 meV photons, while transitions between the E^N_{1V} and E^O_{1V} valence band states are also expected to be fast due to the electrophonon resonance effect. Overall, this energy level structure constitutes a four-level system in which rapid filling of the E^N_{1C} state and rapid emptying of the E^N_{1V} state facilitate population inversion between these two levels. Such a quantum well should exhibit efficient light emission at low excitation levels.

Differential transmission and PL spectra obtained from PbSe/PbSrSe MQW samples [26,29] confirm the energy level calculations shown in Figure 8 and provide evidence in support of the light-emission-enhancing electrophonon resonance effect model shown in Figure 9. Figure 10(a) shows calculated values for the total amount of L-valley degeneracy splitting, $\Delta E_{1C} + \Delta E_{1V}$, and measured values, from differential infrared transmission spectroscopy measurements performed at 4 K [26], which also yield total L-valley degeneracy splitting, as a function of QW width. The good agreement between the theoretical and experimental L-valley degeneracy splitting energies validates the overall accuracy of the energy level model for (111)-oriented PbSe QWs. Note that degeneracy splitting is relatively insensitive to temperature increase as indicated by the constant spacings of the normal and oblique valley transition peaks at 4 and 70 K in Figure 5(a). Splitting energies are not observed in transmission spectra at higher temperatures due to phonon scattering induced peak broadening. Nevertheless, if effective mass is assumed not to change significantly with temperature, or at least the relative difference in masses for normal and oblique valley electrons and holes do not change, then electron and hole states can be expected to have the same degeneracy splitting energies at higher temperatures.

Figure 10(b) shows 5°C PL intensity data from [29] for four MBE-grown samples with different QW widths, which were all grown within three weeks of each other using the same source materials on substrates cleaved from the same BaF_2 ingot, so effects due to material quality differences are not expected to be significant. A log normal fit to the data reveals a maximum PL intensity at a QW thickness of 14.2 nm, nearly the same as the 15 nm thickness at which the population inversion enhancing electrophonon resonance effect is expected to be optimum. This good agreement supports the light-emission-enhancing electrophonon resonance effect model that is proposed here. The fact that higher PL emission intensity is observed for QW thicknesses that do not correspond to electrophonon resonance with a single optical phonon is likely due to multi-phonon resonance processes. For example, according to Figure 8 the electron energy level spacings are on the order of 9 meV for a 113 meV deep and 10 nm wide well suggesting that electrophonon resonance processes involving a TO phonon and two or more acoustic phonons for momentum and energy conserving transitions can be involved in this light emission enhancing mechanism. Another issue to be considered is the fact that phonon, electron, and hole populations will be more spread out over energy at higher temperatures, and this will reduce electron and hole populations in the lower energy normal valley states. At room temperature, where kT is greater than the 5.9 eV TO phonon energy, electrophonon resonance will also facilitate electron and hole excitation from normal valley to higher energy oblique valley states. It is remarkable that even a thermally obscured light emission enhancing electrophonon resonance effect may be observable at room temperature. Clearly, more work is needed in this area to confirm and possibly exploit this potentially useful phenomenon.

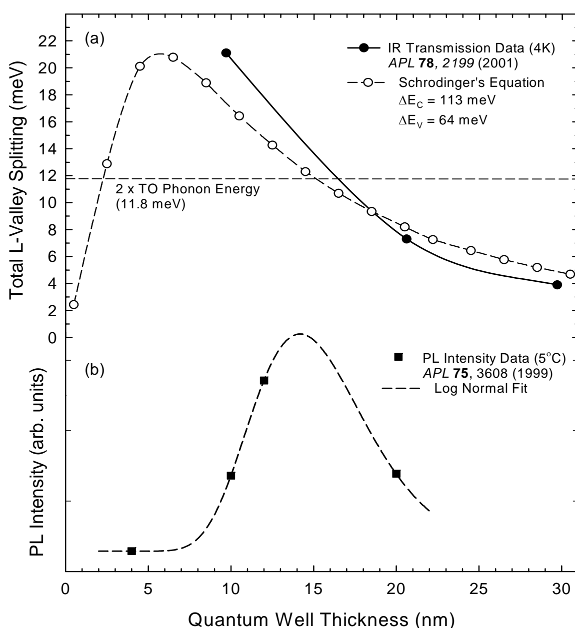


Figure 10. (**a**) Total L-valley degeneracy splitting ($\Delta E_{1V} + \Delta E_{1C}$) as a function of well width calculated using Schrödinger's equation for ΔE_C = 113 meV and ΔE_V = 64 meV (*dashed line*) and measured using differential transmission infrared spectroscopy (*solid line*) [26]. (**b**) PL intensity data at 5°C from [29] and a log normal fit showing a maximum intensity near the same 15 nm thickness at which electrophonon resonance occurs in both the valence and conduction bands as indicated by the crossing in (**a**) of the total splitting energy and the *horizontal dashed line* representing the energy of two PbSe TO phonons. Reprinted with permission of the Electrochemical Society.

In addition to the electrophonon resonance effect, the approximately four-fold reduction in the density of states involved in the interband transitions will also help to reduce population inversion thresholds. These effects can help explain why optically pumped IV-VI VCSELs with MQW active regions have much lower thresholds (10.5 kW/cm^2) [54] than those with bulk PbSe active regions (69 kW/cm^2) [20] and why room temperature PL has been observed from (111)-oriented IV-VI semiconductor QWs while it has not been observed from (100)-oriented IV-VI semiconductor QWs. L-valley degeneracy removal, and thus the possibility for small electron energy level spacings that are resonant with lattice vibrations, does not occur in (100)-oriented QWs because all four valleys are at the same oblique angle with respect to the (100) plane in reciprocal space. Moreover, the low optical pumping powers of less than 2 W/cm^2 used in these PL experiments, when compared to the pump powers of more than 1000 W/cm^2 that are typically used in PL measurements of III-V semiconductor materials, provide further evidence of an efficient light emission mechanism taking place in IV-VI QW materials. These studies show that (111)-oriented IV-VI semiconductor QWs are promising materials for fabrication of low threshold lasers.

6 Progress in Laser Fabrication Using Substrate Removal

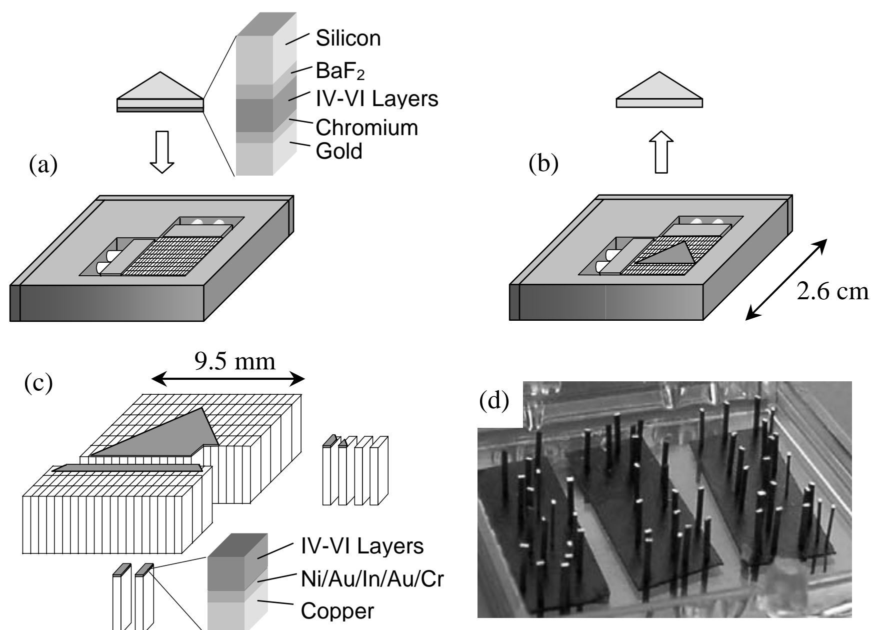


Figure 11. Process flow for bonding and transfer of a IV-VI semiconductor epilayer structure from silicon to the tips of copper bars: (**a**) epilayer and copper metallization and low-temperature bonding to copper bar tips, (**b**) silicon substrate removal by immersing sample into deionized water to dissolve the fluoride buffer layer; (**c**) copper bar separation and epilayer cleaving, and (**d**) photograph of processed samples consisting of cleaved epilayer material on the 500 μm x 1040 μm rectangular tips of 4 mm long copper bars.

Although (111)-oriented MQW materials grown on BaF_2 or Si substrates offer unique features for obtaining low threshold current lasers, it is not possible to use standard cleaving techniques to fabricate in-plane lasers from these structures because silicon and BaF_2 both cleave preferentially along the {111} planes. Such cleaving would produce facets that are not perpendicular to the epilayer surface thus preventing formation of Fabry-Perot resonant cavities. A solution to this problem involves removing the growth substrate and using an alternative "substrate" to cleave the (111)-oriented layer along $(1\bar{1}0)$ planes, which are perpendicular to the (111)-oriented epilayer surface. Figure 11 outlines the steps for a substrate-removal laser fabrication procedure that produces cleaved epitaxial layer structures on the tips of copper bars [55]. Epilayer mounting to the copper is accomplished with a low temperture metallic bonding medium composed of thermally evaporated layers of Au and In-Sn eutectic, which, due to interdiffusion and intermetallic AuIn and AuSn compound formation, results in a high melting point bond. In addition to allowing the cleaving of the epilayer without influence of the growth substrate, this method

also allows placement of two copper heat sinks within microns of the laser active region thus enabling nearly optimal active region heat dissipation since copper has a much higher thermal conductivity (see Table 1) than presently used IV-VI substrate materials. A dual copper heat sink package effectively doubles active region heat dissipation, and thermal modeling predicts that such packaging will increase maximum CW laser operation temperatures by more than 90 degrees [56]. This is sufficient improvement to allow IV-VI laser operation with thermoelectric cooling modules.

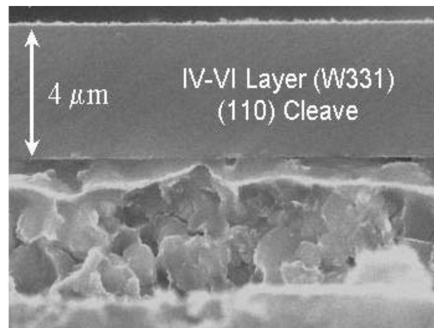


Figure 12. SEM image showing a cleaved IV-VI semiconductor epilayer structure bonded to the tip of a copper bar. Note the high quality of the cleaved $(1\bar{1}0)$ facet in the (111)-oriented epilayer.

Figure 12 shows a scanning electron micrograph (SEM) cross-sectional image of a IV-VI epilayer transferred to copper using the method outlined in Figure 11. This image clearly shows that it is possible to obtain $(1\bar{1}0)$ cleaved facets in (111)-oriented epitaxial layer material. This means that it will be possible to fabricate in-plane lasers that contain (111)-oriented QWs in the active region to take advantage of the light emission enhancing electrophonon resonance effect discussed above. Key to this success was the development of a metallurgical bonding medium. The basic bonding system consisted of Ni (50 nm), Au (300 nm), and In-Sn eutectic (52% In, 48% Sn) (1 μm) deposited on both the gold-coated MBE-grown IV-VI semiconductor structure and the tips of the copper bar assembly. Bonding was accomplished by placing the MBE sample, metallized side down, onto the metallized copper bar assembly and heating to about 200°C. Care was taken to align one of the cleaved silicon {111} edges to be parallel with the copper bar edges. This ensured close alignment of the $(1\bar{1}0)$ planes in the epilayer to be parallel with the copper bar edges. Upon heating, the In-Sn eutectic layers melt first at 120°C. Adhesion between the copper and PbSe is thus promoted by indium, tin, and gold interdiffusion through the liquid layers and formation of high melting point In_2Au, Sn_2Au, and other intermetallic compounds [57]. The silicon growth substrate was removed by immersing the assembly into magnetically stirred deionized water at room temperature. Removal took about 24 h and was successful with either BaF_2 or CaF_2 buffer layers incorporated in the MBE-grown structure.

Current-*vs*-voltage (I-V) data for a transferred and cleaved IV-VI semiconductor *p-n* junction at room and liquid nitrogen temperatures are shown in Figure 13. The non-linear I-V relationship clearly shows successful packaging of a *p-n* junction device, and the onset of forward bias current is consistent with the 451 meV band gap of the $Pb_{0.93}Sr_{0.07}Se$ material (416 meV at 77 K). Low temperature data show a smaller reverse bias saturation current and a larger forward bias series resistance as compared to the data for the same sample at room temperature. Both results are consistent with a lower density of thermally generated carriers in the IV-VI material. The values of the series resistances are 20 ohms at room temperature and 42 ohms at 77 K. The 20 ohm room temperature value represents a significant improvement over the 184 ohm value obtained for a similar IV-VI *p-n* junction structure that was fabricated with different bonding metallurgy and without ohmic contact metallization before cleaving [55]. Clearly, improved bonding metallization along with ohmic contact metallization helped to reduce series resistance.

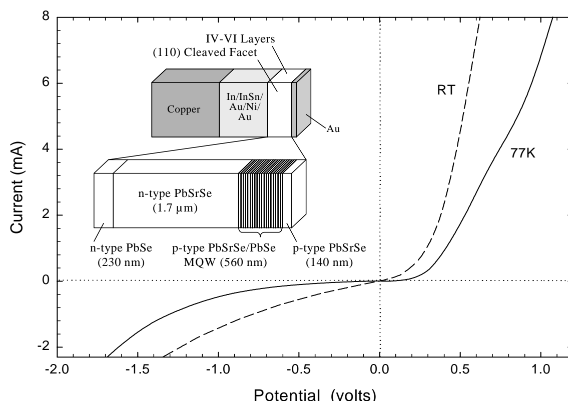


Figure 13. Current *vs* voltage characteristics at room and liquid nitrogen temperatures for a IV-VI semiconductor *p-n* junction bonded to a copper bar. The *inset* shows the MBE-grown structure that was tested. Reprinted with permission of the Electrochemical Society.

PL measurements from as-grown samples and cleaved IV-VI MQW structures bonded onto the tips of the copper bars have been performed using diode laser pumping. This work employed a diode laser with a peak emission wavelength of λ=910 nm, which illuminated the samples from a 45° angle and produced a spot size of about 2 mm in diameter. The luminescence from the sample surface was collected by a 2 inch diameter gold-coated off-axis parabolic mirror and passed through a modular FTIR spectrometer (Oriel, MIR8000). The experimental setup was the same as described in [29] except that a liquid nitrogen cooled photoconductive HgCdTe detector with a 13 μm cut-off was used. The diode laser injection current was varied from 0.5 to 1.8 A, which approximately corresponded to a linear variation in pump power (density) from 200 mW (6.3 W/cm^2) to 880 mW (28.7 W/cm^2). Heat sink temperature was stabilized with a single-stage thermoelectric cooler (TEC) module and measured with an RTD sensor placed near the sample on a copper plate sub-

mount attached to the TEC module. Figure 14 shows mid-IR surface emission spectra from an (a) as-grown and (b) transferred IV-VI MQW layer (W336, L_{QW}=20 nm) for a range of diode laser pumping powers at a constant heat sink temperature of 25ºC. Gaussian fits to the measured spectra, which are dominated by Fabry-Perot interference fringes caused by resonance in the optical cavity formed by the layer, are also shown as dashed lines. Bonding, substrate removal, and cleaving did not significantly degrade the emission intensity of this and other MQW structures tested.

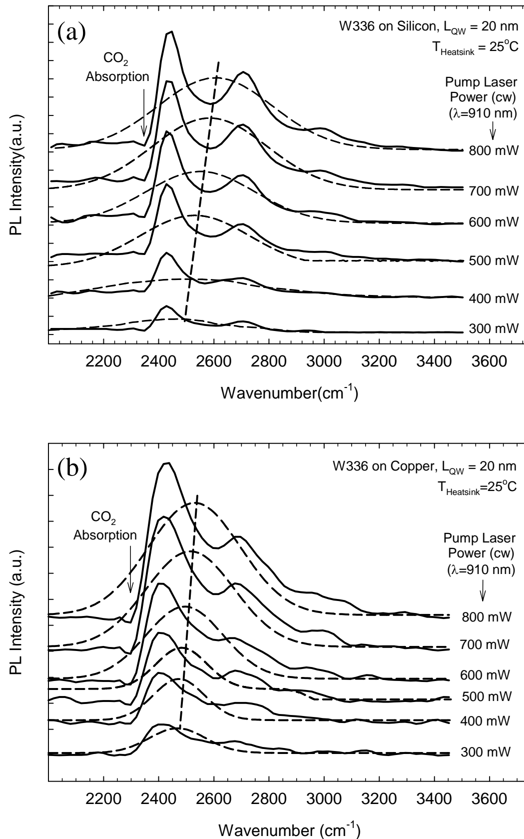


Figure 14. PL spectra from as-grown and transferred PbSe/PbSrSe MQW structures for a constant heat sink temperature of 25ºC and different CW diode laser pump powers. Gaussian fits are also shown. (**a**) MQW on Si growth substrate and (**b**) MQW transferred to Cu. Note the smaller blue shift with increasing pump power for the MQW-on-Cu sample.

With increasing diode laser pumping power, blue shifts in the PL peak energies are observed with both as-grown and transferred layers. PL peak energy increase is due to localized heating of the IV-VI material (a significant amount of electron-phonon scattering is expected due to the large 1.36 eV photon energy relative to the 0.45 eV band gap of the PbSrSe barrier layers) and the associated increase in the band gap energy caused by higher photon flux from the near-IR laser. Table 2 lists

measured interband transition energies for different optical pumping levels along with calculated IV-VI epilayer temperatures (using the 0.41 meV/K coefficient for PbSe) for two different MQW samples on silicon and copper. Epilayer structures bonded to copper exhibit a smaller blue shift at higher pumping levels and thus less heating. Results show that epilayers bonded to copper are more than 20ºC cooler than the same epilayers on silicon when exposed to a continuous wave (λ=910 nm) laser at a pump power (density) of 800 mW (25 W/cm^2).

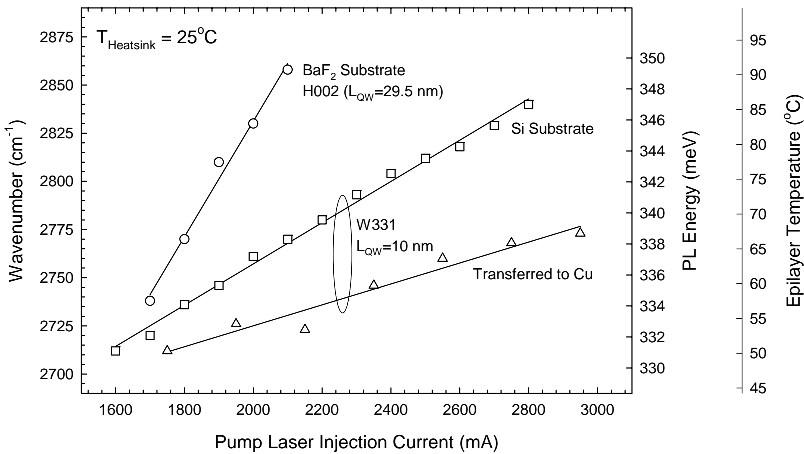


Figure 15. PL energies and calculated epilayer temperatures as a function of diode laser optical pumping level for a MQW structure on BaF_2 and another MQW structure on Si and Cu. The diode pumping laser had a peak emission wavelength of λ=972 nm, and the injection current ranged between 1.5 and 2.8 A corresponding to a power density range between 1.7 and 23.3 W/cm^2, which are similar levels to those obtained with the λ=910 nm laser used in the other set of PL experiments summarized in Table 2.

It should be pointed out that the 22 to 24ºC improvements shown in Table 2 are for comparison with epilayers on silicon, which has a fairly good thermal conductivity. If compared to layers on BaF_2 or IV-VI substrates the improvement would be much greater. To illustrate this, Figure 15 shows measured PL energies for MQW structures on BaF_2, Si, and Cu substrates as a function of the diode laser injection currents used for optical pumping. Note that much higher pumping power densities are possible with epilayers on Si and Cu substrates without causing excessive heating, which was found to degrade PL emission intensity over time, especially for layers on BaF_2 substrates.

Successful transfer of MBE-grown MQW structures from growth substrates to higher thermal conductivity copper without degradation of the optical emission properties shows that it should be possible to fabricate mid-IR lasers using these new fabrication techniques. The improved active region heat dissipation made possible by such transfer can be an important part of a development strategy that will lead to the fabrication of IV-VI mid-IR lasers with CW emission at room temperature.

Table 2. Measured interband transition energies between normal L-valley states as determined by differential transmission spectroscopy (DTS) [26], where there is non-significant epilayer heating, and photoluminescence (PL) using a λ=910 nm diode laser pump, where there is significant epilayer heating. Localized epilayer temperatures at high PL pumping powers are also listed. Note the much lower temperatures for the epilayers-on-copper samples. The heat sink temperature for all PL measurements was 25ºC.

		DTS $(1\text{-}1)^N$	PL peak energy T_H =25°C (meV)		Blue shift	Epilayer temperature
Sample	Substrate	meV	9.4	25 W/cm^2	meV	at 25 W/cm^2 (°C)
W331 (L_{QW}=10 nm)	Si	329	337	353	15.9	83.5
	Cu	329	336	344	7.6	61.6
W336 (L_{QW}=20 nm)	Si	296	307	324	16.7	93.3
	Cu	296	307	314	7.4	68.9

7 Summary

IV-VI semiconductor materials and device development continues to be a vibrant area of research. Recent experimental and theoretical results show that there is great promise for the realization of much improved mid-IR lasers made from IV-VI materials. Electrophonon resonance effects due to L-valley degeneracy removal in IV-VI QWs and new laser structure packaging techniques are just two specific ways that improvements can be achieved. Together, these new materials and fabrication methods can allow fabrication of low threshold mid-IR lasers with widely tunable CW emission at room temperature. Although much of the recent device fabrication progress was accomplished using IV-VI materials grown on silicon substrates, future work should focus on using materials grown on BaF_2 substrates since they have higher crystalline quality due to less thermally-induced strain. In addition, optical pumping, which offers viable prospects for realization of room temperature mid-IR lasing, should be performed with longer wavelength diode lasers. Such lasers, which have recently been demonstrated using GaSb-based materials [58], will allow much higher CW pumping levels without excessive active region heating. Further work in reducing contact resistances with *p-n* junction devices designed for electrical pumping should also be performed.

Interestingly, IV-VI semiconductor mid-IR lasers were first demonstrated in 1964 [59], only two years after the first demonstration of III-V semiconductor near-IR lasers [60]. Subsequent development of IV-VI materials and device technology, however, lagged considerably behind III-V technology. Slower development of IV-VI technology was due in part to lack of compelling commercial applications for mid-IR laser devices. It is likely that this situation is about to change with the recent demonstration of mid-IR laser spectrometers that can be used to assess airway inflammation for asthma diagnosis and therapy monitoring. This clinical application along with other potential health care applications described in Section 2 are helping to create demand for further advances in the mid-IR laser field. Moreover, such practical real-world clinical applications are also helping to define performance criteria for mid-IR lasers. For example, reliable exhaled nitric oxide measurements

require lasers with only moderate CW emission powers (1 mW is more than sufficient) but should have wide tunability to enable simultaneous NO and CO_2 detection. In spite of the long history of IV-VI semiconductor materials, it is clear that these materials still have much to offer in the future.

Acknowledgements

Many people over the last few years have made significant contributions to the work summarized in this chapter. They include Yuefa Li, Paulo Rappl, Huizhen Wu, Xiao-Ming Fang, Khosrow Namjou, Arnaud Sow, James Jeffers, Chad Roller, and Robert Nicholas. Much of the funding for this work was provided by the National Science Foundation under grant numbers DMR-9802396, DMR-0080054, and EPS-0132354 and by the Oklahoma Center for the Advancement of Science and Technology under grant numbers AR021-043 and AR03-057.

References

1. F. Braun, “Ueber die Stromleitung durch Schwefelmetalle (“On Current Flow through Metallic Sulfides”), *Annalen der Physik und Chemie* **153**, 556, (1874).
2. Many references such as C. Kittel, Introduction to Solid State Physics, 5th Edition, Wiley (1976) mistakenly state that the IV-VI compounds PbS, PbSe, PbTe are indirect gap semiconductors.
3. U. Schiessl, J. John, P. J. McCann, “Lead-Chalcogenide-based Mid-Infrared Diode Lasers”, Long-Wavelength Infrared Semiconductor Lasers, p. 145, H.K. Choi (Editor), Wiley, New York (2004).
4. C. B. Roller, K. Namjou, J. Jeffers, W. Potter, P. J. McCann, and J. Grego, "Simultaneous NO and CO_2 Measurements in Human Breath Using a Single IV-VI Mid-Infrared Laser", *Optics Letters* **27**, 107 (2002).
5. C. B. Roller, K. Namjou, J. Jeffers, M. Camp, P. J. McCann, and J. Grego, "Nitric Oxide Breath Testing Using Tunable Diode Laser Absorption Spectroscopy: Application in Respiratory Inflammation Monitoring", *Applied Optics* **41**, 6018 (2002).
6. O. Selroos, A. Pietinalho, AB. Löfroos, H. Riska. "Effect of Early vs. Late Intervention with Inhaled Corticosteroids in Asthma", *Chest* **108**, 1228-1234 (1995).
7. L. R. Narasimhan, W. Goodman, and C. K. N. Patel, “Correlation of breath ammonia with blood urea nitrogen and creatinine during hemodialysis”, *Proc. National Academy of Sciences* **98**, 4617 (2001).
8. S. S. Sehnert, L. Jiang, J. F. Burdick, T. H. Risby, "Breath biomarkers for detection of human liver diseases: preliminary study", *Biomarkers* **7**, 174 (2002).
9. S. M. Studer, J. B. Orens, I. Rosas, J. A. Krishnan, K. A. Cope, S. Yang, J. V. Conte, P. B. Becker, and T. H. Risby, "Patterns and significance of exhaled-breath biomarkers in lung transplant recipients with acute allograft rejection", *Journal of Heart and Lung Transplantation* **20**, 1158-1166 (2001).
10. M. Phillips, G. A. Erickson, M. Sabas, J. P. Smith, and J. Greenberg, "Volatile organic compounds in the breath of patients with schizophrenia", *Journal of Clinical Pathology* **48**, 466-469, (1995).

11. S. E. Ebeler, A. J. Clifford, and T. Shibamoto, "Quantitative analysis by gas chromatography of volatile carbonyl compounds in expired air from mice and human", *Journal of Chromatography B* **702**, 211 (1997).

12. M. Phillips, R. N. Cataneo, B. A. Ditkoff, P. Fisher, J. Greenberg, R. Gunawardena, C. S. Kwon, F. Rahbari, C. Wong, "Volatile Markers of Breast Cancer in the Breath", *The Breast Journal* **9**, 184 (2003).

13. Z. Feit, M. McDonald, R. J. Woods, V. Archambault, and P. Mak, "Low Threshold PbEuSeTe/PbTe Separate Confinement Buried Heterostructure Diode Lasers", *Applied Physics Letters* **68**, 738 (1996).

14. W. Z. Shen, H. F. Yang, L. F. Jiang, K. Wang, G. Yu, H. Z. Wu and P. J. McCann, "Band gaps, effective masses and refractive indices of PbSrSe thin films: Key properties for mid-infrared optoelectronic device applications", *Journal of Applied Physics* **91**, 192 (2002).

15. H. K. Sachar, I. Chao, P. J. McCann, and X. M. Fang, "Growth and Characterization of PbSe and $Pb_{1-x}Sn_xSe$ on Si(100)", *Journal of Applied Physics* **85**, 7398 (1999).

16. P. J. McCann and C. G. Fonstad, "Growth of $PbSe_{0.78}Te_{0.22}$ Lattice Matched with BaF_2", *Thin Solid Films* **227**, 185 (1993).

17. P. J. McCann and D. Zhong, "Liquid Phase Epitaxy Growth of $Pb_{1-x}Sn_xSe_{1-y}Te_y$ Alloys Lattice Matched with BaF_2", *Journal of Applied Physics* **75**, 1145 (1994).

18. A. Lopez-Otero, "Hot Wall Epitaxy", *Thin Solid Films* **49**, 3 (1978).

19. H. Holloway and J. N. Walpole, "MBE Techniques for IV-VI Optoelectronic Devices", Prog. Crystal Growth Character. Vol. 2, Pergamon Press, Great Britain (1979).

20. Z. Shi, G. Xu, P. J. McCann, X. M. Fang, N. Dai, C. L. Felix, W. W. Bewley, I. Vurgaftman, and J. R. Meyer, "IV-VI Compound Mid-Infrared High-Reflectivity Mirrors and Vertical-Cavity Surface-Emitting Lasers Grown by Molecular Beam Epitaxy", *Applied Physics Letters* **76**, 3688 (2000).

21. F. Zhao, H. Wu, A. Majumdar, and Z. Shi, "Continuous Wave Optically Pumped Lead-Salt Mid-Infrared Quantum-Well Vertical-Cavity Surface-Eemitting Lasers", *Applied Physics Letters* **83**, 5133 (2003).

22. P. J. McCann, S. Aanegola, and J. E. Furneaux, "Growth and Characterization of Thallium and Gold Doped $PbSe_{0.78}Te_{0.22}$ Layers Lattice Matched with BaF_2 Substrates", *Applied Physics Letters* **65**, 2185 (1994).

23. I. Chao, P. J. McCann, W. Yuan, E. A. O'Rear, and S. Yuan, "Growth and Characterization of IV-VI Semiconductor Heterostructures on (100) BaF_2", *Thin Solid Films* **323**, 126 (1998).

24. Z. Feit, J. Fuchs, D. Kostyk, and W. Jalenak, "Liquid Phase Epitaxy Grown PbSnSeTe/PbSe Double Heterostructure Diode Lasers", *Infrared Physics & Technology* **37**, 439 (1996).

25. X. M. Fang, K. Namjou, I. Chao, P. J. McCann, N. Dai, and G. Tor, "Molecular Beam Epitaxy of PbSrSe and PbSe/PbSrSe Multiple Quantum Well Structures for use in Mid-Infrared Light Emitting Devices", *Journal of Vacuum Science and Technology* **18**, 1720 (2000).

26. H. Z. Wu, N. Dai, M. B. Johnson, P. J. McCann, Z. S. Shi, "Unambiguous Observation of Subband Transitions from Longitudinal Valley and Oblique Valleys in IV–VI multiple Quantum Wells", *Applied Physics Letters* **78**, 2199 (2001).

27. H. Z. Wu, N. Dai, and P. J. McCann, "Experimental determination of deformation potentials and band nonparabolicity parameters for PbSe", *Physical Review B* **66**, 045303 (2002).

28. P. J. McCann, L. Li, J. Furneaux, and R. Wright, “Optical Properties of Ternary and Quaternary IV-VI Semiconductor Layers on (100) BaF_2 Substrates”, *Applied Physics Letters* **66**, 1355 (1995).

29. P. J. McCann, K. Namjou, and X. M. Fang, “Above-Room-Temperature Continuous Wave Mid-Infrared Photoluminescence from PbSe/PbSrSe Quantum Wells”, *Applied Physics Letters* **75**, 3608 (1999).

30. H. Zogg, A. Fach, J. John, J. Masek, P. Müller, and C. Paglino, "Epitaxy of IV-VI Materials on Si with Fluoride Buffers and Fabrication of IR-Sensor Arrays", Narrow Gap Semiconductors 1995, p. 160, Institute of Physics Publishing Ltd., London (1995).

31. P. J. McCann, X. M. Fang, W. K. Liu, B. N Strecker, and M. B. Santos, "MBE Growth of PbSe/CaF_2/Si(111) Heterostructures", *Journal of Crystal Growth* **175/176**, 1057 (1997).

32. P. Müller, A. Fach, J. John, A. N. Tiwari, H. Zogg, and G. Kostorz, "Structure of Epitaxial PbSe Grown on Si(111) and Si(100) Without a Fluoride Buffer Layer", *Journal of Applied Physics* **79**, 1911 (1996).

33. X. M. Fang, W. K. Liu, P. J. McCann, B. N. Strecker, and M. B. Santos, "XPS Study of the PbSe/CaF_2 (111) Interface Grown on Si by MBE", Infrared Applications of Semiconductors – Materials, Processing and Devices, p. 457, Editors: M. O. Manarsreh, T. H. Meyers, F. H. Julien, and J. E. Colon, Materials Research Society, Pittsburgh, PA (1997). (*Mat. Res. Soc. Symp. Proc.* **450**, 457 (1997)).

34. H. Z Wu, X. M. Fang, D. McAlister, R. Salas, Jr., and P. J. McCann, "Molecular Beam Epitaxy Growth of PbSe on BaF_2-Coated Si(111) and Observation of the PbSe Growth Interface", *Journal of Vacuum Science and Technology B* **17**, 1263 (1999).

35. P. Müller, H. Zogg, A. Fach, J. John, C. Paglino, A. N. Tiwari, M. Krejci, and G. Kostorz, "Reduction of Threading Dislocation Densities in Heavily Lattice Mismatched PbSe on Si(111) by Glide", *Physical Review Letters* **78**, 3007 (1997).

36. B. N. Strecker, P. J. McCann, X. M. Fang, R. J. Hauenstein, M. O’Steen, and M. B. Johnson, "LPE Growth of Crack-Free PbSe Layers on (100)-Oriented Silicon Using MBE-Grown PbSe/BaF_2/CaF_2 Buffer Layers", *Journal of Electronic Materials* **26**, 444 (1997).

37. K. Alchalabi, D. Zimin, H. Zogg, and W. Buttler, "Monolithic Heteroepitaxial PbTe-on-Si Infrared Focal Plane Array with 96 X 128 Pixels" *IEEE Electron Device Letters* **22**, 110 (2001).

38. H. Beyer, J. Nurnus, H. Böttner, A. Lambrecht, T. Roch, and G. Bauer, "PbTe based superlattice structures with high thermoelectric efficiency", *Applied Physics Letters* **80,** 1216 (2002).

39. H. K. Sachar, P. J. McCann, and X. M. Fang, “Strain Relaxation in IV-VI Semiconductor Layers Grown on Silicon (100) Substrates”, Thin-Films – Stresses and Mechanical Properties VII, p. 185, Editors: R. C. Cammarata, M. Nastasi, E. P. Busso, and W. C, Oliver, Materials Research Society, Pittsburgh, PA (1998). (*Mat. Res. Soc. Symp. Proc.* **505**, 185 (1998)).

40. C. P. Li, P. J. McCann, and X. M. Fang, “Strain Relaxation in PbSnSe and PbSe/PbSnSe Layers Grown by Liquid Phase Epitaxy on (100)-Oriented Silicon”, *Journal of Crystal Growth* **208**, 423 (2000).

41. D. W. McAlister, P. J. McCann, K. Namjou, H. Z. Wu and X. M. Fang, "Mid-IR Photoluminescence from IV-VI Layers Grown on Silicon", *Journal of Applied Physics* **89**, 3514 (2001).

42. H. Z. Wu, P. J. McCann, O. Alkhouli, X. M. Fang, D. McAlister, K. Namjou, N. Dai, S. J. Chung, and P. H. O. Rappl , "Molecular beam epitaxial growth of IV–VI multiple quantum well structures on Si(111) and BaF_2(111) and optical studies of epilayer heating", *Journal of Vacuum Science and Technology B* **19**, 1447 (2001).

43. H. Z. Wu, X. M. Fang, R. Salas, D. McAlister, and P. J. McCann, "Transfer of PbSe/PbEuSe Epilayers Grown by MBE on BaF_2-Coated Si(111)", *Thin Solid Films* **352**, 277 (1999).

44. P. J. McCann, "Apparatus for Fabricating Semiconductor Lasers", U.S. Patent Number 5,629,097, May 13, 1997.

45. U. P. Schiessl and J. Rohr, "60°C Lead Salt Laser Emission near 5 μm wavelength", *Infrared Physics and Technology* **40**, 325 (1999).

46. W. Heiss, T. Schwarzl, G. Springholz, K. Biermann, and K. Reimann, "Above-room-temperature mid-infrared lasing from vertical-cavity surface-emitting PbTe quantum-well lasers", *Appl. Phys. Lett.* **78**, 862 (2001).

47. F. Zhao, H. Wu, L. Jayasinghe, and Z. Shi, "Above-room-temperature optically pumped 4.12 μm midinfrared vertical-cavity surface-emitting lasers", *Applied Physics Letters* **80**, 1129 (2002).

48. J. Piprek and S. Nakamura, "Physics of high-power InGaN/GaN lasers", *IEE Proceedings of Optoelectronics* **149**, 145, (2002).

49. M. Beck, D. Hofstetter, T. Aellen, J. Faist, U. Oesterle, M. Ilegems, E. Gini, and H. Melchior, "Continuous-wave operation of a mid-infrared semiconductor laser at room-temperature", *Science* **295** , 301 (2002).

50. M. F. Khodr, P. J. McCann, and B. A. Mason, "Effects of Nonparabolicity on the Gain and Current Density in EuSe/$PbSe_{0.78}Te_{0.22}$/EuSe IV-VI Semiconductor Quantum Well Lasers", *IEEE Journal of Quantum Electronics* **32**, 236 (1996).

51. A. L. Yang, H. Z. Wu, Z. F. Li, Y. Chang, J. F. Li, P. J. McCann, X. M. Fang, "Raman Scattering Study of PbSe Grown on (111) BaF_2 Substrate", *Chinese Physics Letters* **17**, 606 (2000).

52. P. Kinslar, P. Harrison, and R. W. Kelsall, "Intersubband Terahertz Lasers using Four-Level Asymmetric Quantum Wells", *Journal of Applied Physics* **85**, 23 (1999).

53. S. Yuan, H. Krenn, G. Springholz, G. Bauer, and M. Kriechbaum, "Large Refractive Index Enhancement in PbTe/$Pb_{1-x}Eu_xTe$ Multiquantum Well Structures", *Applied Physics Letters* **62**, 885 (1993).

54. C. L. Felix, W. W. Bewley, I. Vurgaftman, J. R. Lindle, J. R. Meyer, H. Z. Wu, G. Xu, S. Khosravani, and Z. Shi, "Low-threshold optically pumped λ= 4.4 μm vertical-cavity surface-emitting laser with a PbSe quantum-well active region", *Applied Physics Letters* **78**, 3770 (2001).

55. P. H. O. Rappl and P. J. McCann, "Development of a Novel Epitaxial Layer Segmentation Method for Optoelectronic Device Fabrication", *IEEE Photonics Technology Letters* **15**, 374 (2003).

56. K. R. Lewelling and P. J. McCann, "Finite Element Modeling Predicts Possibility of Thermoelectrically-Cooled Lead-Salt Diode Lasers", *IEEE Photonics Technology Letters* **9**, 297 (1997).

57. D. W. McAlister, P. J. McCann, H. Z. Wu and X. M. Fang, "Fabrication of Thin Film Cleaved Cavities Using a Bonding and Cleaving Fixture", *IEEE Photonics Technology Letters* **12**, 22 (2000).

58. G. Kim, L. Shterengas, R. U. Martinelli, G. L. Belenky, "High-power room-temperature continuous wave operation of 2.7 and 2.8 μm In(Al)GaAsSb/GaSb diode lasers", *Applied Physics Letters* **83**, 1926 (2003).

59. J. F. Butler, A. R. Calawa, R. J. Phelan, T. C. Harman, A. J. Strauss, and R. H. Rediker, "PbTe Diode Laser", *Applied Physics Letters* **5**, 75 (1964).

60. T. M. Quist, R. H. Rediker, R. J. Keyes, W. E. Krag, B. Lax, A. L. MCWhorter, and H. J. Zeiger, "Semiconductor Maser of GaAs", *Applied Physics Letters* **1**, 91 (1962).

Mid-infrared Vertical Cavity Surface Emitting Lasers based on the Lead Salt Compounds

G. Springholz, T. Schwarzl and W. Heiss
Institute for Solid State and Semiconductor Physics
Johannes Kepler Universität, A-4040 Linz, Austria

1 Introduction

Mid-infrared lasers are of great importance for high-resolution gas spectroscopy due to the numerous absorption lines of almost all molecular gases in this spectral region. Thus, these lasers have a wide range of applications, including gas analysis, pollution monitoring, trace gas sensing, medical diagnostics, time resolved exhaust gas analysis, as well as space applications (see, *e.g.*, [1,2]). Lead salt or IV-VI semiconductor diode lasers [3-8] have long been used in these applications because they cover the whole 3 to 30 μm wavelength region [5-8] and feature the highest operation temperatures of conventional infrared band gap lasers. This is due to their favourable electronic band structure as well as the two orders of magnitude lower non-radiative Auger recombination rates as compared to those of III-V or II-VI narrow band gap semiconductors [9]. Based on advances in epitaxial growth techniques [10-12], lead salt diode lasers have reached operation temperatures up to 223 K in continuous-wave (CW) mode [13] and up to 60°C in pulsed mode [14] (see [7,10,15] for reviews). Although in recent years, III-V intersubband quantum cascade lasers [16] have been able to achieve even higher operation temperatures (see other book chapters or [17-19]), lead salt lasers still offer significant advantages due to their large inherent wavelength tunability and their narrow line widths [13,20]. Thus, lead salt lasers have been an excellent tool for gas sensing applications and will continue to be so in the future [1,7].

In this chapter, the realization of infrared vertical cavity surface emitting lasers (VCSELs) based on the lead salt compounds is described. In VCSELs [21-26], the laser resonator is created by epitaxial growth of Bragg mirrors and therefore, the laser light is coupled out perpendicular to the epilayer surface. VCSELs offer a variety of advantages such as a small beam divergence, single mode operation and the possibility of monolithic integration [22-26]. Furthermore, VCSELs have a great potential for reducing threshold currents and increasing the operation temperatures. Lead salt mid infrared VCSELs were first demonstrated by our group [27,28] in the year 2000 as well as by Shi, Bewley *et al.* [29-31] for wavelengths between 4.5 and 6 μm. This development was triggered by the realization of novel high-reflectivity epitaxial Bragg mirrors and ultra-high finesse microcavity structures for the infrared region [32-35]. Within a short time, the performance of optically pumped lead salt VCSELs was drastically improved [36], with pulsed laser emission now reaching up to 65°C [37,38]. In addition, we have achieved CW emission at very long wavelengths up to 8 μm [39-42], which represents the longest wavelengths of all VCSELs up to now. These lasers show ultra-narrow line widths of 12 μeV or 0.6 nm

[39,40], a large wavelength tuning range of 70 nm [42] and output powers of up to 5 mW in CW and 25 W in pulsed mode [40]. Apart from these achievements, we have also demonstrated the first mid-infrared quantum dot laser using a self-assembled PbSe quantum dot superlattice as active medium [43].

In this chapter, after a brief overview of the basic principles of VCSELs in Section 2, the design and properties of infrared Bragg interference mirrors will be described (Section 3). Representative results for quantum well as well as quantum dot VCSELs are presented in Sections 4 and 5, respectively, and a detailed comparison is given in Section 6. The results on long wavelength CW-VCSELs will be presented in Section 7, and the general aspects of the lasers are summarized in Section 8.

2 Vertical Surface Emitting Lasers

The first all-semiconductor vertical microcavity surface-emitting lasers were demonstrated in 1989 by Jewell *et al.* [21] using epitaxial Bragg mirrors. As shown schematically in Figure 1, these lasers have a completely different design as compared to edge emitters because the resonator axis is perpendicular to the growth plane of the layers. The Fabry-Perot cavity is formed by two high-reflectivity interference mirrors consisting of a stack of dielectric layers. Thus, the generated radiation is reflected and amplified many times across the active cavity region until it is coupled outward vertically through the top or bottom mirrors. Generally, the resonator length of VCSELs is comparable to the optical emission wavelength. This is because pumping can occur only over a limited thickness of the active region and because in this way single mode operation is easily obtained. On the other hand, to compensate the short length of the gain medium, the reflectivity of the VCSEL mirrors must be very high at about 99% as compared to typically only 30 % for edge emitters. Due to the rapid progress in research and development (see [24-26]), already in the mid-1990's VCSELs have become available on a commercial basis [44]. Up to now, however, most work has been focused on III-V VCSELs with emission in the visible to near infrared spectral region (400 nm [45,46] to 1.5 µm [47,48]). For longer wavelengths, until recently, little work has been done due to the tremendous difficulties in fabrication of high reflectivity infrared Bragg mirrors. The few exceptions are type II antimonide VCSELs [49] and CdHgTe VCSELs [50,51] emitting at 2.6 to 3 µm under optical pumping.

A particular feature of microcavity lasers is the large mode spacing of their resonator structure. For a dispersion free dielectric medium, the separation Δk_{mode} of these Fabry-Perot modes is constant in wave number or frequency space and is given by

$$\Delta k_{\text{mode}} = (2nL_{cav})^{-1} \quad (1)$$

where L_{cav} is the cavity length, which in a surface-emitting laser is the total layer thickness of the cavity region, and n is its refractive index. As the cavity length is

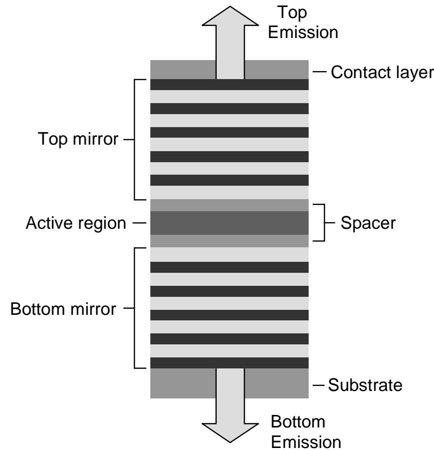


Figure 1. Schematic representation of a vertical cavity surface-emitting laser (VCSEL), consisting of two high-reflectivity Bragg mirrors of multiple λ/4 layer pairs with alternating high and low refractive index and an active region inserted between. The laser light (*arrows*) is emitted in the vertical direction through the top or bottom Bragg mirror.

comparable to the emission wavelength of the active medium, the laser emits on one of the first fundamental cavity modes k_m, which are given by

$$k_m = m(2nL_{cav})^{-1} \quad \text{or} \quad \lambda_m = (2nL_{cav})/m \tag{2}$$

where the integer number $m = 1, 2, 3 \ldots$ corresponds to the order of the mode. Thus, first order emission occurs for a microcavity laser with $L_{cav} = \lambda/2$, and for VCSELs usually the cavity length is typically in the ranges of $\lambda/2$ to 2λ.

An important parameter that characterizes the quality of the resonator is the finesse or quality factor F of the cavity given by

$$F = \pi \frac{\sqrt{R}}{1-R} = \frac{1}{m}\frac{k_m}{\delta k_m} \tag{3}$$

where R is the reflectivity of the mirrors at the resonance wavelength. The finesse describes how often the photons pass back and forth through the gain region before being emitted. Thus, the higher the finesse, the lower the threshold of the laser. As indicated by Equation (3), the finesse can be easily determined by measuring the width of the cavity resonance peaks δk_m in transmission or reflectivity spectra. Examples are shown in Section 4. Typical finesse values required for VCSELs are around 300, which corresponds to mirror reflectivities of about $R = 99\%$. Due to the large mode spacing of VCSELs, the gain spectrum of the active medium overlaps only with one resonator cavity mode. Thus, lasing occurs only on a single mode. Other advantages include simplified planar processing and monolithic integration in arrays and logic or control devices, and their symmetric circular output beams with very small beam divergence angles allow an easy coupling into optical fibers. Finally, very low threshold currents can also be obtained.

3 Lead Salt-based Bragg Interference Mirrors

High reflectivity epitaxial Bragg interference mirrors are crucial elements of vertical cavity surface emitting lasers. As already indicated in Figure 1, these mirrors consist of multilayer stacks of optically transparent dielectric layers with alternating low and high refractive indices and an optical thickness equal to one quarter of the target wavelength λ_T for which the reflectivity is maximized. The reflectivity of Bragg mirrors is purely caused by multiple-interference effects and can be tuned arbitrarily from 0 to 100% just by changing the layer sequence and/or number of layer pairs. In addition, Bragg mirrors exhibit very small intrinsic absorption unlike metallic mirrors, where the high reflectivity is due to extinction of light. Therefore, Bragg mirrors are particularly well suited for applications such as VCSELs where very high reflectivities without any losses are required.

For an ideal Bragg interference mirror consisting of N periods of dielectric $\lambda/4$ layer pairs of alternating low and high refractive indices n_1 and n_2, at the target wavelength λ_T all waves reflected at the interfaces are exactly in phase. Thus, the reflectivity is high and at this wavelength is given by

$$R(\lambda_T) = \left(\frac{1 - n_f / n_i \, (n_1/n_2)^{2N}}{1 + n_f / n_i \, (n_1/n_2)^{2N}} \right)^2 \tag{4}$$

where n_i and n_f are the refractive indices of the media in front and behind the mirror. Equation (4) shows that the reflectivity is essentially determined only by two parameters, namely, the total number N of $\lambda/4$ layer pairs and the refractive index ratio or *refractive index contrast* $\Delta n/\bar{n} = 2\,|n_2 - n_1|/|n_1 + n_2|$. Obviously, the condition that all reflected waves are in phase is fulfilled exactly only for the target wavelength. Therefore, with increasing deviation from this wavelength, the Bragg mirror becomes less ideal and the reflectivity decreases. If this deviation is only small, the mirror reflectivity remains high such that a high reflectivity spectral region called stop band is formed. The whole reflectivity spectrum of such a mirror can be calculated using the transfer matrix method [35,52]. The results show that the relative width of the stop band $\Delta k_{stop}/k_T$ depends only on the refractive index contrast according to [53]

$$\frac{\Delta k_{stop}}{k_T} = \frac{4}{\pi} \arcsin \left| \frac{n_1 - n_2}{n_1 + n_2} \right| \tag{5}$$

Thus, the stop band is wide only for Bragg mirrors with very high refractive index contrast and outside of this region the reflectivity rapidly falls off.

3.1 Mirror Materials

For high reflectivity Bragg mirrors, several requirements must be fulfilled: (i) the use of optically transparent dielectric layers with high purity and thus, negligible residual absorption, (ii) the achievement of smooth $\lambda/4$ layer interfaces to avoid diffuse light scattering, (iii) a precise control of the layer thicknesses to obtain the

desired target wavelength, and finally, (iv) the use of a sufficient number N of high refractive index contrast layers. For monolithic integration in a vertical laser device, the mirrors must be grown epitaxially onto the substrate and should be compatible with the growth of the active region. Thus, all layers should be reasonable lattice-matched to prevent defect formation. For laser applications, the mirror materials should exhibit a high thermal conductivity for efficient heat dissipation. For electrical pumping, the materials must also have a good electrical conductivity, *i.e.*, low resistance, whereas for optical pumping the mirrors must be optically transparent for the wavelength of the pump laser.

As is indicated by Equation (4), the refractive index contrast of the layers plays a crucial role for determining how many layers are required for a certain mirror reflection. This is illustrated in Figure 2(a), where the Bragg reflectivity at the center wavelength is plotted as a function of refractive index contrast for different numbers of $N = 3$, 5 10 and 20 of $\lambda/4$ pairs. Obviously, if the refractive index contrast is lower than 20%, a very large number of pairs of $N > 25$ are required to achieve a reflectivity of 99 %. As is shown in Figure 2(b), this applies for all well lattice-matched III-V or II-VI material combinations, where $\Delta n/\bar{n}$ is only about 12% for GaAs/AlAs and 14% for GaP/ AlP and GaSb/AlSb. For high reflectivity mid-infrared Bragg mirrors, this would result in total mirror thicknesses of more than 15 μm, *i.e.,* in VCSEL structures exceeding 30 μm in thickness. This obviously poses substantial problems for epitaxial growth and therefore, mirror materials with higher refractive index contrast are of crucial importance for infrared VCSEL fabrication.

As shown in Figure 2(b), for the lead salt compounds two types of dielectric materials are well suited as low refractive index layers. These are the wide band gap Eu-chalcogenide semiconductors EuSe and EuTe and the insulator BaF_2 with a band gap in the UV range. Due to their similar lattice structure and lattice parameters as compared to those of the lead salt compounds (see Table 1), these materials can be grown epitaxially with high quality on lead salt structures. Moreover, the refractive indices of the Eu-chalcogenides and BaF_2 of n ~ 2.5 and 1.5, respectively, are

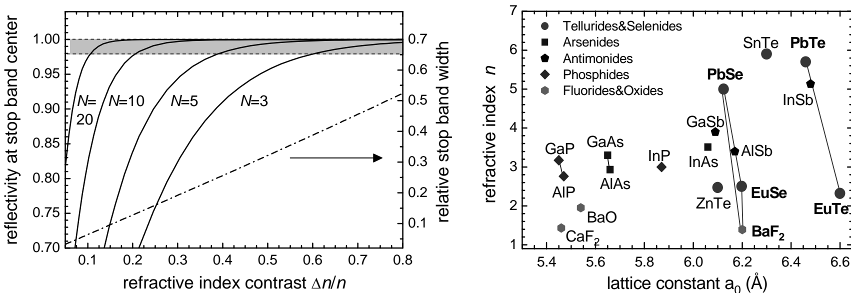


Figure 2. (a) Calculated stop band center reflectivity (*solid lines*) and relative stop band width $\Delta k_{stop}/k_T$ (*dash-dotted line*) of Bragg interference mirrors as a function of refractive index contrast $\Delta n/n$ for different number of $\lambda/4$ layer pairs of $N = 3, 5, 10, 20$ using Equation (1) and $n_1 = n_i = 2.5$, $n_f = 1$ (air) and n_2 varying between 2.5 and 6. (**b**) Refractive indices of various semiconductors and insulators (tellurides, selenides, arsenides, antimonides, phosphides, as well as fluorides and oxides) plotted *vs* lattice constant.

Table 1. Materials for infrared Bragg mirrors and their properties: Energy band gap E_g, refractive index n below E_g, lattice-constant a_0, thermal expansion coefficient β_{lin} [5,10,11].

Compound	$E_{g,300K}$ (eV)	$E_{g,77K}$ (eV)	a_0 (A)	$\Delta\beta_{lin}(10^{-6}K^{-1})$	n ($h\nu<E_g$)
PbTe	0.319	0.217	6.462	19.8	~5.6
PbSe	0.278	0.176	6.126	19.4	~4.8
PbS	0.410	0.307	5.936	20.3	~4.2
EuTe	2.2	2.2	6.598	13.6	2.3
EuSe	2.0	2.0	6.194	13.1	2.5
BaF_2	10.5	10.5	6.200	18.8	1.46

much lower than that of the lead salt compounds of 4.2 - 5.8, as shown in detail in Table 1. Thus, very high refractive index contrasts of up to 100% can be achieved. This drastically reduces the required number of $\lambda/4$ pairs to about $N = 3$. Therefore, these materials have formed the basis for VCSELs in the IV-VI material system.

For the proper design of Bragg mirrors, the spectral dependence of the refractive indices has to be taken into account. While for the wide band gap materials EuSe, EuTe and BaF_2, n is essentially constant in the mid-infrared region, the optical constants of the lead salt significantly vary in the vicinity of the fundamental energy gap E_g. In particular, the refractive index shows a pronounced cusp at the band gap and falls off to higher as well as lower energies (see, *e.g.,* [5,54-58]). As shown by H. Krenn *et al.* [58], this spectral dependence can be described by a complex dielectric function $\varepsilon(E)$ given by

$$\varepsilon(E) = \varepsilon_\infty - \frac{2B((E+i\Gamma_D)^2 - E_g^2)^{1/2}}{\pi(E+i\Gamma_D)} \cdot \operatorname{arctanh} \frac{(E+i\Gamma_D)}{((E+i\Gamma_D)^2 - E_g^2)^{1/2}} \tag{6}$$

where B characterizes the oscillator strength of the interband transition, Γ_D is a damping parameter to account for level broadening, and ε_∞ is the background dielectric constant. Since all of these parameters change as a function of temperature or layer composition they must be determined by transmission or reflection measurements of reference layers. As a general trend, n increases significantly from PbS to PbSe and PbTe [5,54,55] and decreases with increasing temperature [5,54-57]. For ternary lead salt alloys, the refractive index changes oppositely as compared to the energy band gap. Thus, n increases with increasing x_{Sn} in $Pb_{1-x}Sn_xTe$ or $Pb_{1-x}Sn_xSe$ [5], and decreases for $Pb_{1-x}Eu_xTe$ [56], $Pb_{1-x}Eu_xSe$ and $Pb_{1-x}Sr_xSe$ [59] with increasing x_{Eu} or x_{Sr}. This has to be taken into account in the mirror design.

3.2 Examples and Results

Highly efficient lead salt Bragg mirrors were first demonstrated by our group using the PbTe/EuTe material system [32-34], based on the well-controlled 2D heteroepitaxial growth obtained under optimized conditions [60,61]. Figure 3 shows the

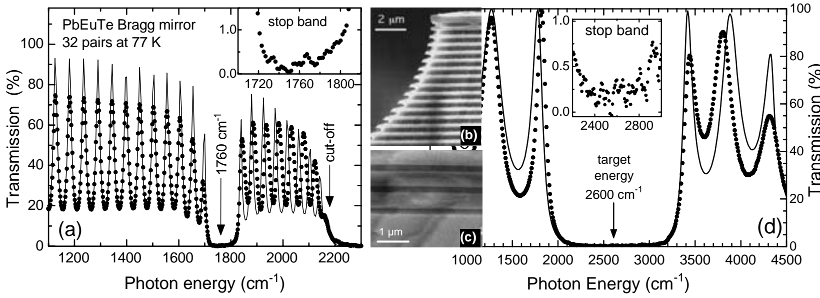


Figure 3. Transmission spectra and cross-sectional SEM images of two high-reflectivity IR Bragg mirrors formed by 32 pairs of 240 nm $Pb_{0.99}Eu_{0.01}Te$ and 259 nm $Pb_{0.94}Eu_{0.06}Te$ (*left hand side* (**a**) and (**b**)), and by three pairs of 205 nm $Pb_{0.93}Eu_{0.07}Te$ alternating with 458 nm EuTe λ/4 layers (*right hand side* (**c**) and (**d**)). The *inserts* show the stop bands on an enlarged scale. *Dots*: measured data, *solid lines*: calculated spectra using the matrix transfer method. See [32] and [33] for details.

transmission spectra for two such Bragg mirrors consisting of (a) nearly lattice-matched $Pb_{1-x}Eu_xTe$ layers with x_{Eu} alternating between 1 and 6%, and (b) $Pb_{0.93}Eu_{0.07}Te$ / EuTe layers. Clearly, in both cases a well-defined stop band with essentially zero transmission, *i.e.*, high reflectivity is formed at the target wavelengths of 5.68 and 3.85 µm, respectively. Outside of the stop band region, Fabry-Perot interference fringes appear due to multiple reflections over the Bragg mirror stack as a whole on the optically transparent BaF_2 substrates. As is evident from Figure 3 (d), for the PbEuTe/EuTe Bragg mirror, a high reflectivity ($R > 99$ %) and very wide stop band region is obtained already with three layer pairs [33], whereas in the other case as many as 32 layer pairs are required due to the much lower refractive index contrast [32]. The excellent quality of the PbEuTe/EuTe Bragg mirrors has been proven by Schwarzl *et al.* [33] by their use for realization of ultra-high finesse microcavity structures. For microcavities with five period bottom and top mirrors thus, exceedingly sharp cavity resonance were obtained with a line width as low as 0.63 cm^{-1} at k_T = $1877cm^{-1}$. This corresponds to a record cavity finesse of F_{eff} = 1700, which is more than a factor of two greater as compared to the best values reported for GaAs/AlAs mirrors [62]. Using Equation (3), this finesse value proves that mirror reflectivities exceeding 99.9 % can be obtained.

Other alternative material combinations, such as PbEuSe/EuSe [39-42], PbSrSe /BaF_2 [31,36,63,641] as well as PbEuSe/BaF_2 [65] have been also used. As summarized in detail in Table 2, these materials feature a similar high refractive index contrast and therefore, in all cases high reflectivity Bragg mirrors have been obtained. However, for the PbSrSe-based mirrors, heteroepitaxial growth seems to be less ideal, resulting in a higher interface roughness and thus lower reflectivity [36,63,64] as compared to the EuTe based mirrors. Apart from the lattice-mismatch between the layer materials, another important parameter is the high-energy cut-off wavelength where the intrinsic absorption of the lower band gap mirror material sets in (see Table 2). This is important because for optically pumped VCSELs, this cut-off determines which type of pump sources can be used for laser excitation.

Table 2. Material combinations and properties of high-reflectivity lead salt Bragg mirrors, including refractive index contrast $\Delta n/n$, lattice-mismatch $\Delta a/a$, reflectivity R_{max} at the stop band center wavelength λ_{cent} for N=3, relative stop band width $\Delta k_{stop}/k_{cent}$ and absorption cut-off wavelength $\lambda_{cut\text{-}off}$ of the lower energy band gap mirror material (room temperature data).

Layer combination	$\Delta n/n$ (%)	$\Delta a/a$ (%)	λ_{cent} (μm)	R_{max} for N=3(%)	$\Delta k/k$ (%)	$\lambda_{cut\text{-}off}$ (μm)	References
$Pb_{0.93}Eu_{0.07}Te/EuTe$	90	1.9	3.5-8	99.6	58	2.3	[33-35]
$Pb_{0.99}Eu_{0.01}Te$ / $Pb_{0.94}Eu_{0.06}Te$	6	0.2	5.4	99 (N=32)	4	3.6	[27,32]
$Pb_{0.94}Eu_{0.06}Se/EuSe$	55	1.1	6.5-8	98.5	42	2.7	[39,42]
$Pb_{0.99}Sr_{0.01}Se/BaF_2$	105	1.3	4.1-4.6	99	76	3.0	[31,36]
$Pb_{0.85}Sr_{0.15}Se/BaF_2$	87	1.3	4.5-4.7	98	54	1.5	[31,36]
$Pb_{0.65}Sr_{0.35}Se/BaF_2$	65	1.0	4.1	92	43	0.9	[63,64]

4 Lead Salt Vertical Cavity Surface Emitting Lasers

Lead salt VCSELs were simultaneously invented by our group [28,35] as well as by Shi and Bewley *et al.* [29-31]. The first VCSELs emitted at wavelengths between 4.5 and 6 μm, compared to the longest wavelength of 2.9 μm previously obtained for antimonide-based VCSELs [49]. Because for VCSELs no cleavage is required for laser fabrication, lead salt VCSELs can be grown on readily available substrates such as BaF_2 (111), whereas for conventional diode lasers high cost single crystalline lead salt substrates are needed. Many works have proven BaF_2 to be an excellent substrate material for lead salt heterostructures and multilayers [10,11,66-68]. It is not only readily available in high crystalline quality, but is also optically transparent and offers a five times larger thermal conductivity as compared to lead salt substrates. This allows a much better heat dissipation during laser operation. Also, BaF_2 has a much higher mechanical hardness, which facilitates the processing of the laser structures. As a consequence, up to now all lead salt VCSELs have been grown on BaF_2 substrates. On the other hand, BaF_2 exhibits (111) natural cleavage planes as opposed to the preferred (100) planes for the lead salt compounds. Thus, BaF_2 cannot be used for conventional edge emitters but this does not limit its used for vertical emitting lasers.

4.1 Design Issues and Resonator Structure

For the design of the VCSEL structures, several parameters must be optimized. First, the vertical resonator structure must have a high finesse in order to obtain low laser thresholds. This can be achieved using lead salt Bragg mirrors with high refractive index contrast (see Section 3). Second, the active region should have a high infrared luminescence efficiency. For this purpose, both lead salt quantum well structures or bulk-like binary epilayers can be employed because the luminescence efficiency has been found to be comparable in both cases [69]. For adjusting the

emission wavelengths, ternary alloys of the lead salts with Eu, Sr or Sn chalcogenide compounds can be used by which the energy band gap can be tuned over a wide range from below hundred to several hundred meV (see, *e.g.*, [4,5,10-12] for details). The third design issue is the matching of the cavity resonance modes to the emission of the active region. This is crucial because, in contrast to edge emitting lasers, the mode spacing of VCSELs is generally much wider than the width of the gain spectrum and because, evidently, lasing can occur only if the active region emits into a resonator mode. Finally, for the pumping of the VCSELs with ready to use low cost pump lasers, the pump light has to be efficiently coupled into the active region. In principle, this can be achieved by lateral patterning of the VCSELs and pumping over the side walls [27], but a preferable and much more efficient geometry is by pumping through the bottom or top Bragg mirror. This requires that the Bragg mirror to be transparent for the shorter pump laser wavelength outside of the stop band region, *i.e.*, the absorption edge of the mirror materials must be above the pump photon energy (*e.g.*, $E_g > 1.25$ eV for a 1-µm pump laser). As discussed already in Section 3, this narrows the material selection for the Bragg mirrors and reduces the refractive index contrasts that can be obtained.

Another essential design issue is the targeted operation temperature of the VCSELs because the optical constants as well as the energy band gaps and luminescence wavelengths of the lead salt compounds strongly vary with temperature. Figure 4, shows the comparison of the temperature dependence of the cavity mode positions and the PbTe band gap for the case of a high finesse microcavity structure formed by five period PbEuTe/EuTe Bragg mirrors and a 4λ thick PbTe/ PbEuTe MQW cavity region. As shown in Figure 4 (a), due to the high refractive index contrast of 80%, a very wide stop band is formed in the reflectivity spectrum with five very sharp cavity resonance peaks of m = sixth to tenth order and 29 meV mode spacing. Outside of the stop band, the usual Fabry-Perot interference fringes

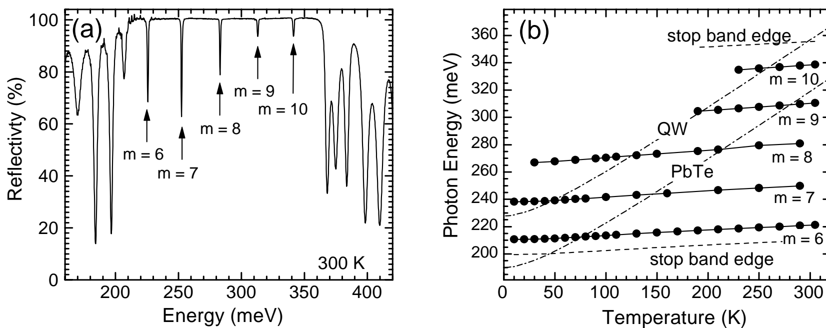


Figure 4. (**a**) Reflectivity at 300 K and (**b**) energetic positions of the cavity resonance peaks (*full dots*) as a function of temperatures of a 4λ microcavity structure consisting of two five period PbEuTe/EuTe Bragg mirrors and a cavity with four 20 nm PbTe/PbEuTe quantum wells. The *dash-dotted lines* in (**b**) show the temperature dependence of the absorption onset within the quantum wells as well as for pure PbTe and *dashed lines* the upper and lower edges of the mirror stop bands.

appear. The width of the resonance peaks is well below 1 meV, indicating a very high cavity finesse. In Figure 4(b), the position of these modes is plotted as a function of temperature. Clearly, all modes shift linearly to the red as the temperature decreases, with a total red-shift of about 12 meV when T decreases from 310 to 10 K. This is agrees well with the corresponding ~5% increase of the refractive index of the cavity material, giving an increase of the effective optical cavity length.

In comparison, for the same temperature range the PbTe energy band gap (dashed-dotted line in Figure 4(b)) decreases by as much as 134 meV from 325 meV at 310 K to 191 meV at 10K, and the same applies also to the interband transitions between the 20 nm PbTe quantum well energy levels (upper dashed-dotted line in Figure 4 (b)). Thus, the temperature coefficient of the interband optical transitions of $\alpha = \Delta E/\Delta T = -0.85$ meV/K is more than one order of magnitude larger than that of the resonator modes of only –0.05 meV/K. A similar large difference is also found for the other lead salt VCSEL material systems. The important consequence of this discrepancy is that the spontaneous emission from the active lead salt materials matches to a given cavity resonance only in a certain temperature range. For the m = eighth mode in Figure 4(b), *e.g.*, this is at around 100 K for the QW transitions, whereas for the m = seventh mode this matching range is shifted below 50 K. In general, the VCSELs will thus operate only in the temperature range where the cavity mode is just above the interband transition, and best laser operation is expected at the temperature when the resonator mode coincides exactly with the maximum of the gain spectrum of the active material. As shown below, this is indeed the behaviour observed for all lead salt microcavity lasers.

4.2 PbTe/EuTe Quantum Well VCSELs

As mentioned above, due to the well-controlled heteroepitaxial growth and the high refractive index contrast, the PbTe/EuTe material system is very well suited for lead salt VCSEL fabrication. In the following, representative results of these lasers are presented.

4.2.1 Structure and Optical Characterization

The structure of the VCSELs is formed by two EuTe/PbEuTe Bragg mirrors. Here we consider a sample (M1076) with a 2λ cavity region with four PbTe/PbEuTe quantum wells designed for low temperature operation at 4.8 μm wavelength [28]. Accordingly, the Bragg mirrors were optimized for a center stop band at 4.8 μm and consisted of three λ/4 pairs of 599 nm EuTe alternating with 272 nm PbEuTe, and the cavity length was set to 2170 nm. The 5 % Eu content in the ternary layers yields a refractive index contrast of 80 % to EuTe and therefore, a stop band reflectivity above 99.5%. At the same time, confinement barriers of 150 meV around the PbTe QWs are obtained. Figure 5 (b) shows a cross sectional SEM image of the complete VCSEL structure, in which the four 20nm PbTe quantum wells in the cavity region can be clearly distinguished. The high optical quality of the microcavity structure is demonstrated by the FTIR transmission spectrum depicted in Figure 5 (a).

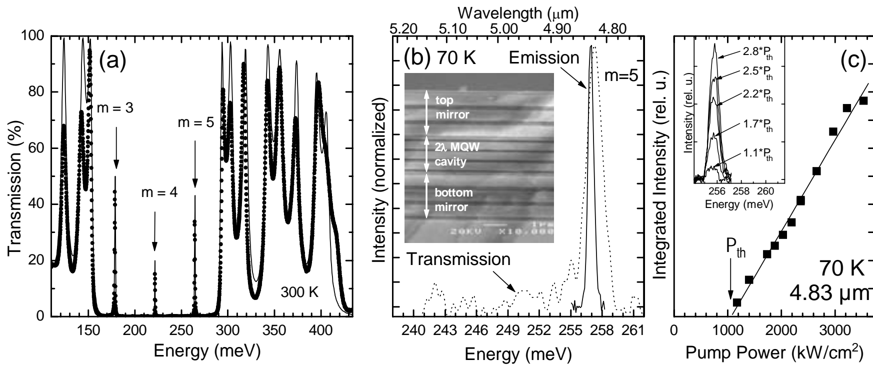


Figure 5. (**a**) Measured (*dots*) and calculated (*line*) transmission of the PbTe MQW VCSEL M1076 at 300K, showing the broad stop band region with three cavity resonance peaks of m = third-, fourth- and fifth-order. (**b**) Laser emission spectrum (*solid line*) at 4.83 μm and corresponding fifth cavity resonance peak (*dashed line*) at 70 K. Insert: VCSEL cross-sectional scanning electron micrograph showing the three-period EuTe/$Pb_{0.95}Eu_{0.05}Te$ Bragg mirrors and the four 20 nm PbTe quantum wells in the 2λ active region. (**c**) Integrated output intensity at 4.83 μm *vs* pump power, indicating an external laser threshold of 1050 kW/cm^2. *Insert*: Emission spectra at various pump powers. See [28] for details.

In the wide stop band region three sharp cavity resonance peaks at 179, 221, and 264 meV are observed, corresponding the m = third, fourth and fifth order modes. The FWHM of the peaks is only 0.3 - 0.5 meV, indicating a cavity finesse as high as 800 at 300 K.

4.2.2 Threshold and Laser Emission

Optical pumping of the VCSELs was performed at 1064 nm with a Nd:YVO_4 laser normal to the surface with a pulse length of 10 ns. The stimulated emission spectrum at 70 K is shown in Figure 5(b), together with the 70 K transmission spectrum (dashed line) in the same spectral region. Clearly, the laser emission at 257 meV (4.83 μm) coincides with the fifth-order cavity mode. The FWHM of the emission is only 0.4 meV, which is essentially limited by the resolution of the spectrometer setup. The narrow line width is not only much smaller than the photoluminescence of usual PbTe/$Pb_{1-x}Eu_xTe$ quantum wells (typically around 4 meV [68,69]), but also smaller than the width of the cavity mode (see Figure 5 (b)). This spectral narrowing is a clear indication for stimulated emission. The angular dependence of the emission shows a very rapid drop in intensity even for slight deviations of the emission angle off the surface normal direction. Such a narrow forward directed output is characteristic for a planar microcavity device [22].

The dependence of the integrated output intensity on the pump power density is shown in Figure 5(c). It exhibits the typical linear behaviour of optically pumped semiconductor lasers above a pump power threshold of P_{th} = 1050 kW/cm^2, with a narrowing of the line width with increasing pump power, as shown in the insert of Figure 5(c). Because the 1064 nm pump laser is strongly absorbed by the PbEuTe

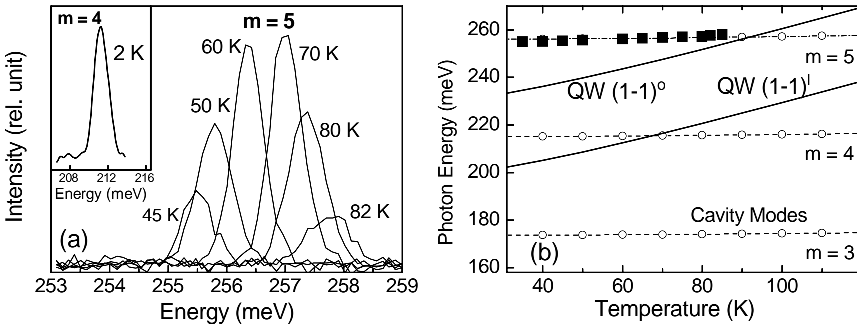


Figure 6. (**a**) PbTe QW VCSEL (M1706) emission spectra measured at different temperatures for a constant pump power of 2.8 P_{th}. (**b**) Measured cavity resonance positions (o) and calculated interband transition energies (*solid lines*) for the 20 nm PbTe quantum wells plotted *vs* temperature. The *full squares* indicate the peak positions of the VCSEL emission as shown in panel (**c**). See [28] for details.

layers in the top mirror (E_g=500 meV) with an estimated absorption coefficient of 5×10^4 cm^{-1}, less than 0.5 % of the incident light is actually coupled into the active cavity region. Thus, the *effective* threshold is only of the order of 5 kW/cm^2.

4.2.3 Temperature Dependence of Emission

Further evidence for microcavity lasing arises from temperature dependent emission measurements. Whereas at low temperatures around 4 K, emission is observed at the m = fourth cavity mode at λ = 5.87 μm (see insert of Figure 6(a)), at 35 K the emission switches to the m = fifth resonance at λ = 4.83 μm. As demonstrated in Figure 6(a), when the temperature is further increased, the output power strongly increases with little change in the emission wavelength. At 70 K, the emission intensity reaches its maximum but rapidly decreases for temperatures above 80 K and vanishes completely at 85 K. As discussed in Section 4.1, this can be readily explained by considering the temperature dependence of the intrinsic QW emission in relation to the energy position of the cavity modes as shown in Figure 6(b). Whereas the measured cavity mode positions (o) are almost constant with temperature, the interband transition energies of the 20 nm quantum wells (solid lines in Figure 6(b)) calculated using the envelope function method [57] rapidly increase as the temperature increases due to the increasing PbTe band gap.

At temperatures below 65 K, the m = fourth resonator mode overlaps with the 2D joint density of states of the ground state of the longitudinal energy valleys along the [111] growth direction. Therefore, low temperature emission is due to the $(1\text{-}1)^l$ QW transitions. As the temperature is raised, the quantum well transition energy of the oblique energy valleys $(1\text{-}1)^o$ approaches the m = fifth cavity mode and thus, the laser emission switches to the higher cavity mode. This is due to the fact that the joint density of states and thus, the gain of the three fold degenerate oblique valleys is about a factor of 10 larger than that of the longitudinal valley. From the calculations, the $(1\text{-}1)^o$ QW transition moves above the energy of the m = fifth cavity mode when the temperature reaches 85 K (see Figure 6(b)). This is exactly the temperature

where the laser action stops. In addition, the temperature where the maximum output intensity occurs corresponds to situation where the onset of the (1-1)^o transitions is just below the cavity mode. This behaviour clearly indicates that the maximum operation temperature of the VCSEL is not limited by intrinsic effects but rather by the characteristic temperature dependence of the optical transitions.

4.3 Room Temperature PbTe QW VCSELs

To obtain room temperature operation of PbTe VCSELs, the laser design has to be changed [35]. First, the central cavity mode has to be shifted to shorter wavelengths by decreasing the cavity length in order to match the 3 μm room temperature PbTe band gap emission. Second, to achieve a more efficient optical pumping, the top mirror has to be made transparent to the excitation wavelength. Thus, the Eu content in the PbEuTe layers of the top mirror has to be increased to about 20% to make the mirror transparent for the 2 μm pump laser. In addition, the number of PbTe quantum wells was increased to achieve a higher laser gain. Figure 7 (a) shows the cross sectional SEM image of such a modified VCSEL, consisting of the same bottom mirror as above, but a shorter cavity length and nine 20 nm PbTe double quantum wells at the antinodes of the electric field distribution in the cavity region. To compensate the reduction in refractive index contrast, four instead of three λ/4 layer pairs were used in the top Bragg mirror to retain a reflectivity above 99%. For the following laser samples M1210 and M1208, the central cavity resonances were λ = 3.1 and 3.6 μm, respectively.

Figure 7 (b) displays the room temperature reflectivity spectrum of the VCSEL M1210, showing the wide stop band region from 300 to 494 meV with three micro-cavity resonance peaks of *m*= third-, fourth- and fifth-order. The strong and narrow *m*=third resonance at 320 meV exhibits a finesse of 150, whereas the higher *m*=fourth and fifth modes are weaker and broader due to the stronger damping by

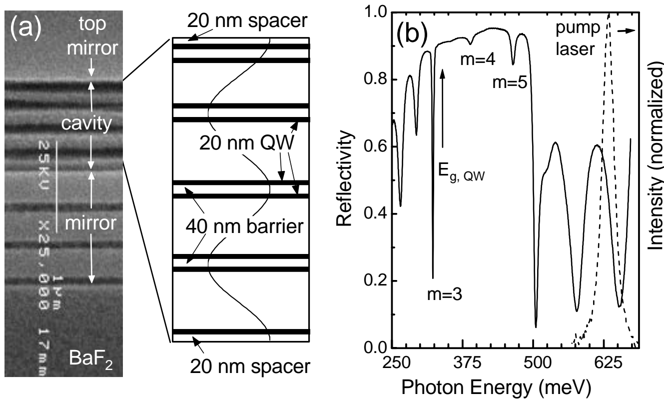


Figure 7. (**a**) Cross sectional SEM of VCSEL #M1210 together with a sketch of the cavity region. The chemical contrast between distinct layers is visualized by selective etching. The *dotted line* represents the standing electromagnetic wave in the cavity. (**b**) Reflectivity spectrum compared with that of the pump pulses. The arrow indicates the absorption edge of the 20 nm wide PbTe quantum wells. See [35,37] for details,

the absorption in the PbTe QWs starting at about 322 meV at room temperature. The damping is strongest at the *m*=fourth cavity mode at 386 meV, exhibiting a line width of 8.6 meV.

For laser excitation, we used a 1.97 μm 100 fs laser generated in an optical parametric amplifier, pumped by a Ti:sapphire laser [70]. The spectrum of the pump laser pulse directed at an angle of 15° onto the sample is shown as dashed line in Figure 7(b). The pump peak is centered at about 628 meV photon energy, which is significantly below the 788 meV absorption edge of the top mirror. Thus, optical pumping creates electron-hole-pairs only within the active region. At the 628 meV pump laser photon energy, the top mirror has a reflectivity of about 50%. The second VCSEL M1208 shows a similar reflectivity spectrum, except that the Bragg mirror stop band as well as the cavity resonances are shifted to lower energies such that the central mode of is tuned to 354 meV.

Room temperature emission spectra of VCSEL M1210 for various pump energy densities are shown in Figure 8(a). For an excitation energy of 0.8 mJ/cm^2, the emission spectrum shows a Lorentzian shaped peak centered at 400 meV (λ=3.1 μm) with a width of 20 meV. Increasing the excitation density to 1 mJ/cm^2 results in a considerable narrowing of the emission spectrum with a drastic rise of the luminescence intensity. Both effects indicate the onset of stimulated emission. For excitation powers higher than 1 mJ/cm^2 the line width becomes larger again and the integrated emission intensity increases linearly with increasing pump power. This is shown in Figure 8(b) for both VCSELs at T = 18 and 50°C. Thus, a laser threshold of 0.83 mJ/cm^2 is obtained for sample M1210 at room temperature.

In contrast to the 77 K QW VCSEL described in Section 4.1, the emission spectra of the RT VCSELs are about 12 eV blue shifted with respect to the central cavity resonance in the reflectivity measurements (see Figure 8(a)). Furthermore, the width of the emission is considerably larger than that of the cavity resonance peak. Both effects are caused by the excitation with high intensity fs pulses. Taking into account that the carrier lifetime in PbTe is larger than the excitation pulse duration [9], the carrier concentration in the microcavity layer can be estimated to exceed 1x10^{19} cm^{-3} immediately after the excitation pulse and subsequently decreases due to photon emission and nonradiative recombination. This dynamic band filling modulates the effective band gap and concomitantly the refractive index, which in term modulates the cavity resonance position. This obviously leads to a broadening of the time-averaged emission spectra from the sample. The observed 12 meV blue shift of the emission spectra corresponds to a reduction of the refractive index of 3%, a value that has been observed in PbTe epilayers already under ps excitation with excited carrier concentrations of around 10^{17} cm^{-2} [71].

As shown in Figure 8(b), at room temperature VCSEL M1208 exhibits superior lasing characteristics as compared to M1210, with a lower laser threshold of only 0.32 mJ/m^2 and a much steeper slope of the linearly rising output intensity with increasing pump power. In spite of the fact that for pump powers exceeding 2 mJ/cm^2 the output intensity starts to level off due to saturation effects (see inset in Figure 8(b)), M1208 yields a more than three times larger output intensity than

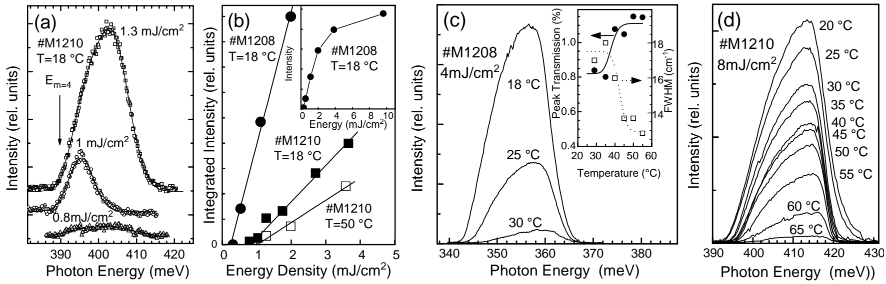


Figure 8. (**a**) Emission spectra of VCSEL M1210 at 18°C for various pump pulse energy densities, below and above the laser threshold. The arrow shows the resonance energy of the m=4 cavity mode. (**b**) Integrated output intensity *vs* pump energy density for VCSELs M1210 and M1208. The *inset* shows the output intensity of M1208 to higher powers. The emission spectra of the VCSELs at higher temperatures are shown in (**c**) for sample M1208 and in (**d**) for M1210 up to 65 °C. The *inset* in (c) shows the transmission (*circles*) and the FWHM (*squares*) of the m=4 mode of M1208, determined from transmission measurements. See [35,37] for details,

M1210, even up to the highest pump pulse energy of 19 mJ/cm^2. This superior laser performance is a result of the lower density of states and the higher population density after carrier thermalization at the energy of the microcavity resonance, which is just about 12 meV above the band edge of the PbTe quantum wells.

The dependence of the emission spectra of both samples as a function of operation temperature is shown in Figure 8(c). For VCSEL M1208 and an excitation power of 4 mJ/cm^2 (Figure 8(c)), the stimulated emission at λ = 3.5 µm or 354 meV strongly decreases when the sample temperature is raised to 30°C, and it completely quenches at 35°C. This is not due to an increase of nonradiative recombination but rather to the 9 meV blue shift of the band edge of the PbTe quantum wells caused by the 20°C temperature increase. Thus, the QW gain spectrum shifts out of the cavity resonance mode. This is confirmed by the temperature dependent CW FTIR experiments shown in the insert of Figure 8(c), where a step-like increase of the cavity resonance transmission as well as a decrease in the line width is observed when the temperature rises above 35°C. This indicates an abrupt decrease of the absorption coefficient due to the shift of the QW transitions above the cavity mode.

For sample M1210, this does not limit the maximum operation temperature because for this sample the central cavity mode is 50 meV higher than for M1208 and thus the QW transitions are below the cavity mode even up to T = 150°C. The emission spectra of this VCSEL at 8 mJ/cm^{-2} pump power are shown in Figure 8(d). With increasing temperature the laser output intensity at first decreases only slowly. Above 55° C the intensity decreases more rapidly and completely quenches at 70° C. As demonstrated by Figure 8(b), up to 50° C the laser threshold increases only slightly to 1 mJ/cm^{-2}and the absorption coefficient at the cavity resonance does not show any significant change from transmission measurements. Therefore, the observed quenching of the laser emission at 70°C seems to be caused by nonradative processes like Auger recombination and multiphonon emission.

4.4 PbSe/PbSrSe VCSELs

Apart from the PbTe/EuTe VCSELs, vertical cavity lasers have been also produced based on the PbSe-PbSrSe-BaF_2 material system (Z. Shi *et al.* [29-31,36]). In this case, epitaxial Bragg mirrors are fabricated using $Pb_{1-x}Sr_xSe$ and BaF_2 as quarter wavelength layers with x_{Sr} ranging from 3 to 35% with refractive indices of $n = 4$ - 4.7 for PbSrSe and 1.46 for BaF_2. This yields an index contrast of up to 100%, but the MBE growth of these materials seems to be less well controlled as compared to the PbTe-EuTe system. As a result, the reported Bragg mirror reflectivities in the 92 to 99% range [36,63,64] and microcavities with finesse values of ~70 to 200 [31,36,63] are somewhat lower as compared to those of PbEuTe/EuTe structures.

The PbSe VCSELs usually consist of three period $Pb_{1-x}Sr_xSe/BaF_2$ Bragg mirrors, where $x_{Sr} = 3$ % for the bottom mirror and 15 or 35 % for the top mirror in order to make them transparent for 1- 2 μm pump lasers. As active region, either bulk-like PbSe or 20 to 30 nm PbSe/$Pb_{0.97}Sr_{0.03}Se$ multi-quantum wells were used with λ or λ/2 cavity lengths. The target wavelength of the lasers was between 4.1 to 4.6 μm to match the spontaneous PbSe emission in the range of 200 to 300 K. Figure 9 shows representative results for these VCSEL consisting of a 1.03 μm PbSe active region obtained by optical pumping with a 100 ns 2.098 μm Ho:YAG laser. Plotting in Figure 9(a) the output *vs* pump power at temperatures between 199 and 256 K, a clear threshold behaviour is found with a linearly increasing output power above thresholds of 100-200 kW/cm^2. The threshold values are plotted in Figure 9(b) as a function of operation temperature, showing a minimum threshold of 69 kW/cm^2 at T = 240 K. This is obviously the temperature, where an optimum overlap of the PbSe gain spectrum with the cavity resonance is obtained.

The highest operation temperature for this VCSEL was 290 K beyond which the PbSe band gap energy shifts above the cavity mode. Improved results were reported for PbSe quantum well VCSELs, for which lower pump thresholds of 10.5 kW/cm^2 were obtained [36] as well as higher operation temperatures of 310K [63].

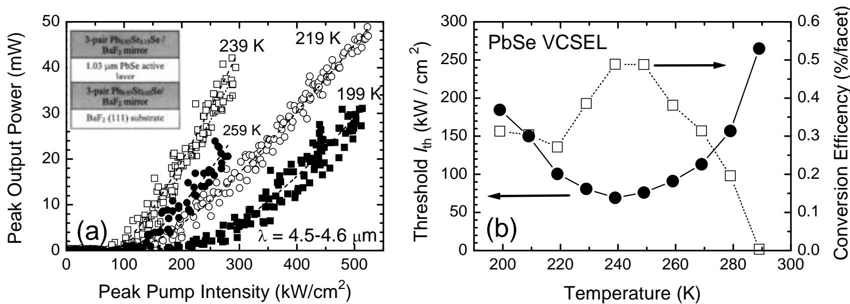


Figure 9. (**a**) Peak output power *vs* pump intensity for a PbSe VCSEL emitting at 4.5–4.6 μm at four different temperatures. The VCSEL structure shown in the *inset* consists of two three period PbSrSe/BaF_2 Bragg mirrors and a λ PbSe as active cavity medium. (**b**) Temperature dependence of the threshold pump power I_{th} and the conversion efficiency, indicating a lowest threshold of 69 kW/cm^2 at T = 240K (adopted from Z. Shi *et al.* [31]).

The maximum peak pulsed output powers were up to 200 mW, with conversion efficiencies of about 0.5% (see Figure 9(b)). Also, pumping using a 1.064 μm pulsed Nd:YAG laser could be used when x_{Sr} in the PbSrSe top mirror layers was increased to 35% [63]. The wavelength tuning rate of the lasers was around –1.4 nm/K [31,36,63], giving a total tuning range of up to 100 nm in one case [31]. Although compared to the PbTe-EuTe based lasers, the PbSe VCSELs show similar thresholds and operation temperatures, the emission line widths of 21 nm [31,36] to 46 nm [63] have been significantly larger and even more, no line width narrowing at increasing pump powers was reported. In addition, the angular emission behaviour exhibited a one order of magnitude broader beam divergence of 24 – 31° [36] as compared to the sharp 1° emission cones [38,42] seen for the PbTe VCSELs as expected for high finesse microcavity structures. This implies a lower quality of the PbSrSe/BaF_2 Bragg mirrors, presumably caused by a substantially larger interface roughness [63,64].

5 Self-assembled Infrared Quantum Dot Lasers

Due to their highly peaked zero-dimensional density of states, semiconductor quantum dots offer great advantages for opto-electronic devices [72]. Thus, quantum dot lasers have been predicted to yield not only strongly increased material gain and differential gain [73], but also lower threshold currents, higher modulation band widths and better temperature stability [74] as compared to 2D quantum well or bulk-like lasers. The experimental realization of the first quantum dot lasers [75] became possible through the development of self-assembled quantum dots [76] based on the Stranski-Krastanow growth mode that is typical for highly lattice-mismatched heteroepitaxial systems [77,78]. As demonstrated by recent work, in this way, self-assembled InAs quantum dot lasers with ultra-low threshold current densities were achieved [79] and a wider spectral range can be covered with a given material system [80]. Up to recently, almost all work on quantum dot lasers has been focused exclusively on III-V semiconductor quantum dots emitting in the visible or near infrared spectral region [81]. To apply this concept for mid-infrared lasers, quantum dots of narrow gap semiconductor are required. We have therefore, pursued this approach by growing self-assembled PbSe quantum dots [43,82] grown on PbTe or PbEuTe (111) using molecular beam epitaxy.

5.1 Self-assembled PbSe Quantum Dots

Strained-layer heteroepitaxy has been the method of choice for direct fabrication of self-assembled quantum dots because due to the layer/substrate mismatch, three-dimensional (3D) nanoislands are spontaneously formed on the surface of a thin 2D wetting layer once a certain critical layer thickness is exceeded [77,78]. In the initial stages of growth, these self-assembled dots are coherent, *i.e.*, defect free and thus exhibit excellent optical and electronic properties as required for optoelectronic device applications [76]. The driving force for 3D island formation is the lattice

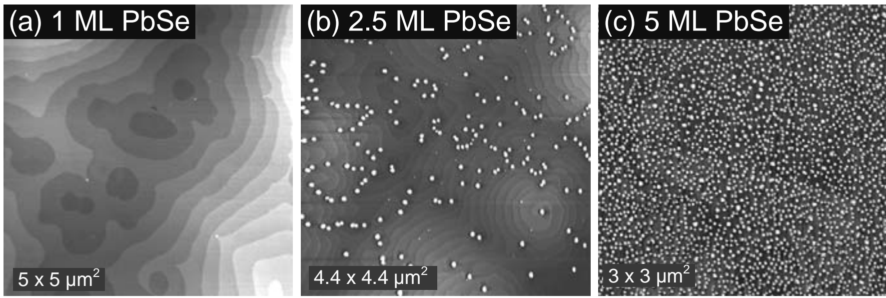


Figure 10. Self-assembled PbSe quantum dot formed on PbTe (111) illustrated by atomic force microscopy images of PbSe layers with layer thickness of 1, 2.5 and 5 monolayers for (**a**) to (**c**), respectively. Because for layer thicknesses below 2 ML only a 2D wetting layer is formed only single monolayer surface steps and terraces are observed in (a). See [82] for details.

mismatch strain between the layer and substrate material, which can be effectively relaxed by the lateral elastic expansion or compression of the islands in the directions of the free side faces, thus leading to a significant lowering of the total free energy of the system [77].

Concerning the lead salt compounds, we have developed the growth of self-assembled PbSe dots on PbTe (111) with a 5.4% lattice-mismatch [82-85]. As shown in Figure 10, the same characteristic 2D/3D Stranski-Krastanow growth mode transition occurs as observed, *e.g.* for InAs on GaAs dots [76]. This is illustrated by the atomic force microscopy (AFM) images of PbSe layers with d increasing from 1 to 5 monolayers (ML) depicted in Figure 10. Up to a layer thickness of 1.5 ML, PbSe growth clearly proceeds in a 2D manner [82] and thus, the epilayer surface exhibits only the usual monolayer surface step structure (Figure 10(a)). At 1.5 ML thickness, however, 3D nanoislands suddenly start to nucleate (Figure 10(b)), and by 5 ML PbSe thickness (Figure 10(c)), the dot density saturates and further deposition only leads to a growth in dot size [82]. The size of the dots is typically around 100 Å in height and 230 Å in width, and all dots exhibit a well-defined facetted island shape that is defined by the formation of {100} side facets. Thus, three sided pyramids with an aspect ratio of $1/\sqrt{6}$ are formed [82,83]. This is illustrated in Figure 11(d). The statistical analysis shows that the dot size distribution is quite narrow, with a variance as low as ±12 % [82]. The actual density and size of the dots can be tuned effectively by the changing the deposited PbSe layer thickness [82] or by changing the growth temperature [85]. By the latter, the dot density can be varied from 2000 down to 100 dots per μm^2 and the average dot height from 40 to 180 Å when the substrate temperature is changed from 280 to 420°C, respectively [85]. Further tuning of density and size can be achieved by post-growth annealing [84].

A particular feature of self-assembled PbSe dots is their efficient lateral ordering in PbSe/PbEuTe superlattices [86-89]. Figure 11(a)-(c) shows a sequence of AFM images of the PbSe surface dots for superlattices with increasing number of superlattice periods N, where the thickness of the PbEuTe spacer thickness was kept constant at 480Å. Whereas for the first dot layer (Figure 11(a)), the PbSe islands are

distributed randomly on the surface, with increasing number of SL periods a rapidly progressing ordering of the dots occurs. After 10 periods (Figure 11(b)) the dots are preferentially aligned in single and double rows along the $<\bar{1}10>$ directions and small patches of hexagonally ordered regions are formed. With further increasing period number larger and larger ordered regions are formed such that for $N \geq 30$ nearly perfect 2D hexagonal dot arrays are formed (Figure 11(c)). The ordering process is also clearly reflected by the FFT power spectra of the AFM images shown in the inset of Figure 11(a)-(c). Whereas for the single dot layer exhibits a broad and diffuse ring in the FFT spectrum, for $N = 10$, six side maxima appear, which rapidly sharpen with increasing period number such that many higher order satellite peaks appear (Figure 11(c)). This indicates the formation of well-ordered 2D hexagonal dot lattices. Correlated with this lateral ordering, also the dot size distribution narrows from ±14 % for the initial single dot layer to ±10% for superlattices with 100 periods [87].

A particularly interesting feature is the special interlayer stacking of the dots in the SL structures. As proven by X-ray diffraction [86] as well as transmission electron microscopy (see Figure 12(c)) [88], the PbSe dots are anticorrelated along

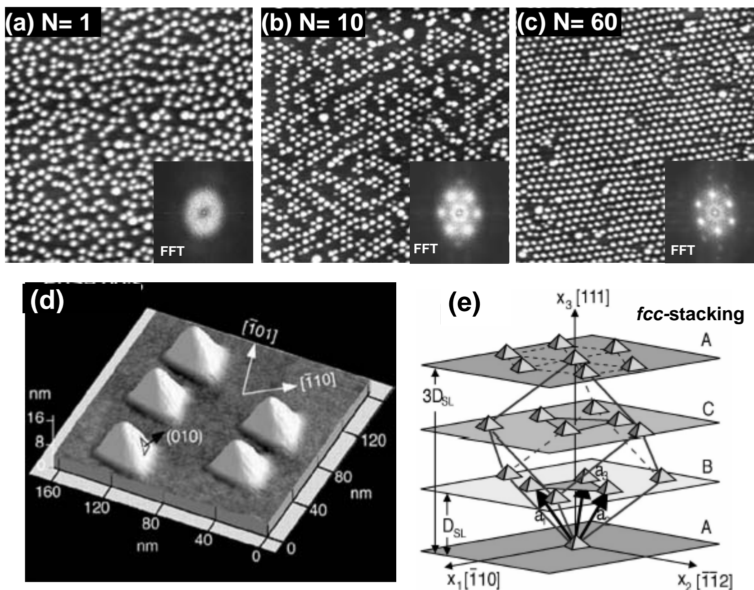


Figure 11. Lateral ordering and interlayer stacking in self-assembled PbSe/PbEuTe quantum dot superlattices: (**a**)-(**c**): AFM images (1.5 × 1.5 μm^2) of superlattices with $N = 1$, 10 and 60 SL periods. Each superlattice period consists of 5 ML PbSe alternating with 480 Å PbEuTe. The efficient lateral ordering is revealed by the FFT power spectrum of the AFM images shown as inserts. (**d**) 3D AFM surface image of PbSe dots showing their well defined pyramidal shape with {100} side facets. (**e**) Schematic illustration of the *fcc*-like *ABCABC*... interlayer dot stacking formed in the superlattices due to the high elastic anisotropy of the lead salt compounds. See [86,87] for details.

the growth direction forming an *fcc*-like *ABCABC…* stacking sequence [86] as shown schematically in Figure 11(e). The efficient lateral ordering as well as the *fcc* dot stacking can been explained on the basis of the interactions of the dots via their elastic strain fields [86,90]. These are strongly modified as compared to those in III-V dot multilayers due to the very high elastic anisotropy of the lead salt compounds [90]. Further work has shown that the lateral PbSe dot spacings can be effectively controlled by the spacer layer thickness such that self-organized 3D quantum dot crystals with tunable lattice constant are obtained [86,88]. Also, for small spacer layer thicknesses a transition from the *fcc*-stacking to a vertical dot alignment was found [88], which could be explained on the basis of strain field interactions as well [88]. As shown by various studies, these structures also show a rather good in-plane hexagonal ordering [89], which is however rather of short-range ordering type [91]. More details can also be found in [86-91].

5.2 Quantum Dot VCSELs: Growth and Characterization

For fabrication of mid-infrared quantum dot VCSELs, we have embedded PbSe/ PbEuTe quantum dot superlattices into the active region of vertical cavity structures. Due to the strong increase of the PbEuTe band gap with Eu content ($\Delta E_g / \Delta x_{Eu}$ = 4.48 eV at 4 K [92]), a quantum confinement in the PbSe dots can be achieved already for Eu concentrations of a few percent. Here, we have chosen x_{Eu} = 5%, which yields E_g= 420 meV for PbEuTe as compared to 145 meV for PbSe at 4 K in the unstrained case. As shown in Figure 12, the VCSEL structure consists of two high reflectivity PbEuTe/EuTe Bragg mirrors and a cavity designed for a central mode at 290 meV (λ = 4.27 μm). This corresponds to the low temperature onset of quantum dot absorption. The bottom and top Bragg mirrors consisted of three $\lambda/4$ layer pairs of 513 nm EuTe alternating with 253 nm $Pb_{0.94}Eu_{0.06}Te$.

For the active region, first a 3 μm thick $Pb_{0.95}Eu_{0.05}Te$ buffer layer was deposited, followed by a 140 period PbSe quantum superlattice of 5 ML PbSe dots alternating with 480 Å $Pb_{0.95}Eu_{0.05}Te$ spacer layers. This yields dots with an average height and width of 120 and 300 Å, respectively, and a dot density of about $5x10^{10}$ cm^{-2} in each layer. Due to the large number of superlattice periods, a large total cavity length of 10.3 μm is obtained. This yields a large number of cavity modes that may overlap with the inhomogeneously broadened quantum dot gain spectrum, which facilitates laser emission. In addition, this also allows to obtained spectroscopic information on the electronic transitions of the dot ensembles. Figure 12(b) shows a cross sectional scanning electron micrograph of the whole VCSEL structure, and the well ordered PbSe dot arrangement and *fcc*-stacking in the superlattice is illustrated in Figure 12(c) by the cross-sectional and plan-view TEM image of a superlattice reference sample.

The FTIR transmission spectrum of the cavity structure at 300 K is shown in Figure 13(a), illustrating the high optical quality of the cavity structure. Within the wide stop band region, a large number of sharp cavity resonance peaks are observed with a mode spacing of 11 meV. The central m = 28th- order cavity mode is at 288 meV (4.3 μm) with a FWHM of 0.2 meV, indicating a high cavity finesse. At 300 K, the absorption edge of the PbSe quantum dots is at about 390 meV (arrow in Figure

13(a)). This is well above the cavity stop band region and causes a damping of the Fabry Perot interference fringes above the stop band. As a consequence, the whole structure represents an unfilled optical cavity. This is also evident from the symmetric distribution of the resonance peaks in the stop band region. As the temperature is lowered, the dot absorption shifts into the stop band due to the decreasing PbSe band gap. As a result, the higher cavity modes become damped by the dot

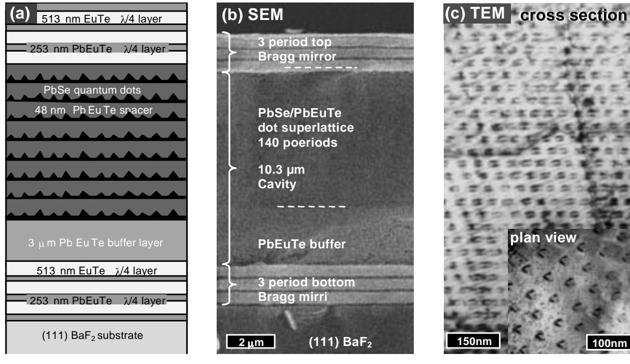


Figure 12. (**a**) Schematic representation and (**b**) cross sectional scanning electron micrograph of the PbSe/$Pb_{1-x}Eu_xTe$ quantum dot VCSEL structure M1220. (**c**) Cross sectional and plan-view transmission electron micrograph of a PbSe dot superlattice reference sample with 5 ML PbSe alternating with 480 Å $Pb_{1-x}Eu_xTe$. See also [43].

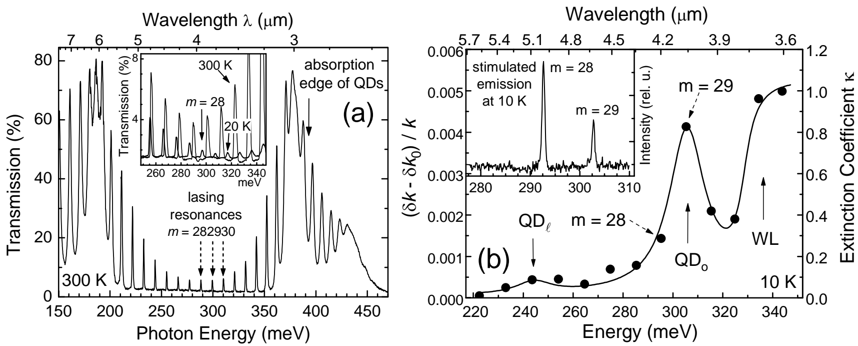


Figure 13. (**a**) FTIR transmission spectrum of the PbSe dot VCSEL M1220 at 300 K, showing the broad stop band region with the large number of cavity resonance peaks. The central cavity mode is of m = 28th- order and the onset of PbSe dot absorption at about 390 meV is indicated by an *arrow*. *Inset*: Cavity spectrum on an enlarged scale for T = 20 and 300 K (*thick and thin solid lines*, respectively), showing the strong damping of the $m \geq 28$ cavity modes at 20 K due to the shift of the quantum dot absorption into the stop band region. (**b**) Dispersion of the extinction coefficient κ (*right hand scale*) of the quantum dot superlattice at 10 K determined from the normalized width $(\delta k - \delta k_0)/k$ of the cavity resonances as shown on the *left hand scale*. The *full line* shows the fitted calculation of the dispersion of the extinction coefficient. *Inset*: 10 K microcavity emission spectrum with the two laser lines corresponding to the m=28th- and 29th- cavity modes. See [93] for details.

absorption. This is demonstrated by the inset of Figure 13(a), which shows the transmission spectra at 20 K and 300 K around the stop band center on an enlarged scale. As indicated by the arrow, at 20 K all cavity resonance peaks with $m \geq 28$ are strongly damped. This indicates an onset of the PbSe dot absorption at about 290 meV, which is blue-shifted by about 145 meV with respect to the bulk PbSe energy band gap due to the quantum confinement.

The microcavity structure can be also used as a sensitive detector for determining the dispersion of the dot absorption. Because the width of the cavity resonance peaks is linearly related to the cavity absorption [93], from the normalized difference of the damped resonances widths at low temperature δk and those of the undamped resonances at room temperature δk_0, the spectral dependence of the PbSe dot absorption is obtained. As shown in Figure 13(b), it exhibits an absorption onset at about 240 meV due to the electronic transitions of the longitudinal valleys. At higher energies, the spectrum shows a peak at 305 meV and a step-like increase at 328 meV. The peak is attributed to the dot absorption caused by the three-fold degenerate oblique valleys, whereas the step is attributed to the absorption of the 2D wetting layers. The peak at 305 meV has a full width at half-maximum of about 20 meV, which corresponds to the inhomogeneous broadening of the electronic transitions due to the dot size variations. For more details, see [93].

5.3 Quantum Dot Laser Emission

The PbSe dot VCSEL was pumped by a 1064 nm pulsed Nd:YAG laser. The stimulated emission spectra measured at 1.5 K are shown in Figure 14(a) for various pump powers. At this temperature, simultaneous emission at the m = 28th- and 29th-order cavity modes at λ = 4.24 and 4.09 μm occurs, with a line width of only 700 μeV. This two-mode laser operation is a result of the inhomogeneous broadening of the quantum dot gain spectrum due to dot size fluctuations. Figure 14 (b) shows the integrated output intensity of the two lines as a function of pump power. Above an external threshold of P_{th} = 510 kW/cm^2, the output intensity increases linearly with a transition to an even stronger nonlinear increase at a power of about 900 kW/cm^2. Below threshold no signal was found due to insufficient detector sensitivity. Because the 1064 nm pump laser is strongly absorbed by the $Pb_{1-x}Eu_xTe$ layers in the top Bragg mirror, less than 5×10^{-3} of the incident light is actually coupled into the cavity region. Therefore, the effective threshold power density is estimated to be only 2.6 kW/cm^2 for this structure.

The measured temperature dependence of the laser emission is shown in Figure 14(c). At very low temperatures, laser emission occurs predominantly at the m = 28th central cavity mode with a weaker emission also at the 29th mode. As the temperature is increased, the intensity of the 29th mode increases whereas that of the 28th mode decreases and eventually disappears at 40 K where the 29th mode shows a maximum of emission. As the temperature further increases, the 29th emission in turn decreases and at 60 K the next higher laser mode turns on. At 70 K, the 29th mode completely disappears, whereas the 30th mode emission persists up to 90 K.

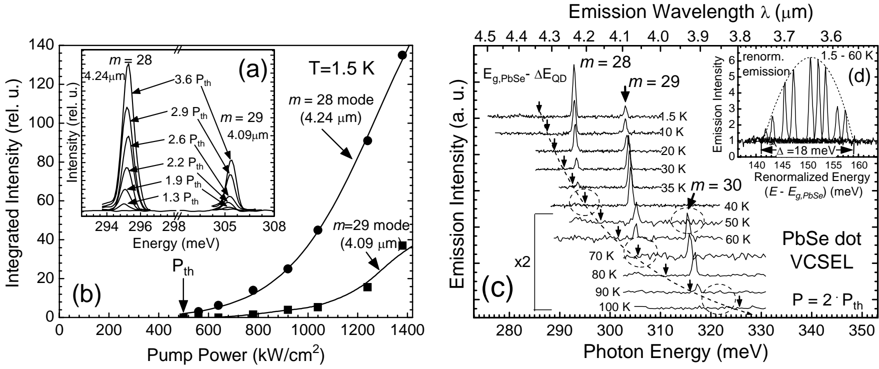


Figure 14. (**a**) PbSe dot VCSEL emission spectra at various pump powers. (**b**) Integrated output intensities at the m = 28th mode at 4.24 µm (filled circles) and the 29th mode at 4.09 µm (filled squares) *vs* pump power at 1.5 K, indicating an external threshold of 510 kW/cm^2. (**c**) Emission spectra at T = 1.5 to 100 K for a pump power of 2 × P_{th}, showing the switching of the laser emission to higher cavity modes as the temperature increases. The *arrows* and *dashed line* indicate the low energy edge of the quantum dot gain spectrum given by $E_{g,PbSe}(T) + \Delta E_{QD}$, where ΔE_{QD} = 141 meV. The *insert* shows the emission spectra on a renormalized energy scale, where for each spectrum the band gap energy of bulk PbSe was subtracted. The envelope of these spectra (*dashed line*) corresponds to the inhomogeneously broadened quantum dot gain spectrum with a width Δ of about 18 meV. See [43] for details.

The successive switching of the laser emission to higher modes is explained by the increase of the PbSe band gap with temperature T, given by [5]

$$E_{g,PbSe}(\text{meV}) = 125 + \sqrt{(400 + 0.256 \cdot T^2)} \qquad (7)$$

where T is given in Kelvin. As a result, the interband transition energies within the quantum dots are increasingly blue shifted with rising temperature, with a temperature dependence equal to that given by Equation 7 because for T < 100 K the effective electron and hole masses are almost constant. On the other hand, for the laser lines the wavelengths are fixed by the optical Eigenmodes of the cavity structure and are thus almost temperature independent. This means that as the temperature increases, the quantum dot gain spectrum is tuned through the cavity mode spectrum with a quenching of the emission each time when the low energy edge of the gain spectrum moves out of resonance, and with a turn on of a new mode when the upper edge of the gain spectrum moves into a higher cavity mode.

Using the m = 28th mode at 294 meV, which just disappears at 35 K and subtracting the corresponding 35 K PbSe band gap of 153 meV, a quantization energy ΔE_{QD} of 141 meV is obtained for the lowest energy dot Eigenstate. Adding $E_{g,PbSe}(T)$ from Equation (7), the temperature dependence of this low energy onset of the gain spectrum is obtained as represented by the dashed line and downward arrows in Figure 14(c). Obviously, these energies agree well with the turn-off of the laser modes at the different temperatures. Similarly, using the turn-on of the m = 30 mode at about 55 K (see Figure 14), we get an upper energy edge of the gain spectrum for a total dot quantization energy of 159 meV. Thus, a total width of the quantum dot gain spectrum of about 18 meV is obtained. The same result is obtained when we

plot the emission spectra on a reduced energy scale E- $E_{g,PbSe}(T)$. As shown in Figure 14(d), the dashed envelope of these spectra directly corresponds to the quantum dot gain spectrum for laser emission. Obviously, we get the same width of 18 meV from this plot, which agrees well with the width of the dot absorption spectrum of Figure 13(b). From these arguments, it becomes evident that the quenching of the dot laser emission at temperatures higher than 100 K is due to the shift of the dot gain spectrum to the edge of the cavity stop band region where the optical cavity finesse is significantly reduced.

6 Lead Salt VCSELs with Different Active Regions

As mentioned above, for optoelectronic devices the dimensionality of the active region strongly influences the performance due to the resulting modifications of the electronic density of states. For III-V lasers, reducing the active material from bulk, to 2D quantum wells and 0D quantum dots has resulted in drastically reduced threshold currents and better temperature stability [72-74,94]. With respect to lead salt VCSELs, lasers with active regions consisting of bulk-like layers, quantum wells as well as quantum dots have been realized as well (see sections above), but a comparison of the performance of these lasers has been hindered by the fact that different layer materials, different cavity structures as well as different pump lasers were used in these works. To allow a comparison of the dimensionality effects on the properties of lead salt VCSELs, we have therefore fabricated a series of lasers with nearly identical optical cavity structure but with different types of active material [38]. In addition, the same optical setup and pump source was used for laser excitation and characterization in these investigations.

6.1 VCSEL Structures and Optical Properties

The laser samples were grown by molecular beam epitaxy onto BaF_2 (111) substrates under comparable conditions. The vertical cavity structure is formed by two high-reflectivity EuTe/$Pb_{1-x}Eu_xTe$ Bragg mirrors. For the three-period bottom mirror, x_{Eu} = 5 % was used to obtain a refractive index contrast of 80 % and a mirror reflectivity above 99 %. As listed in Table 3, different active regions were grown on top, namely, (i) a bulk-like 1115 nm PbTe-layer (M1504), (ii) a nine period PbTe/PbEuTe multi quantum well (M1208) and (iii) a 236 period PbSe quantum dot superlattice (M1235). This yields different respective cavity lengths of 2.1, 1.4 and 14 μm. The optical cavities were designed such that for each VCSEL the central cavity mode matches the room temperature spontaneous emission of the active material. Thus, the bulk-like VCSEL was optimized for a target wavelength of 3.7 μm, the QW VCSEL for 3.5 μm and the QD VCSEL for 3.3 μm. To make the top mirror transparent for the 1.907 μm pump wavelength, the Eu-content in the top PbEuTe layers was increased to 30 % for M1504, and to 20 % for M1208 and M1235. Thus, four layer pairs were used to compensate for the reduced refractive index contrast and retain a reflectivity above 99 %.

Table 3. Structure of three different lead salt VCSELs with different active regions of bulk-like PbTe (VCSEL M1504), PbTe multi quantum wells (VCSEL M1208) and self-assembled PbSe quantum dots (VCSEL M1235). The same type of bottom and top Bragg mirrors were used in all cases, consisting of three EuTe/$Pb_{1-x}Eu_xTe$ layer pairs with x_{Eu} = 0.05 for the bottom mirror and four layer pairs with x_{Eu} = 0.2 to 0.3 for the top mirror.

Sample	Active material	Barriers in active region	Central mode cavity length	Buffer layer in cavity	λ/4 Bragg mirror layers
Bulk-VCSEL M1504	PbTe 1115 nm	None	λ =3.82 μm L =2.1 μm	$Pb_{0.94}Eu_{0.06}Te$ 960 nm	EuTe: 400 nm PbEuTe: 183 nm
QW-VCSEL M1208	Nine 20 nm PbTe QWs	$Pb_{0.94}Eu_{0.06}Te$	λ =3.54 μm L =1.4 μm	-	EuTe: 385 nm PbEuTe: 178 nm
QD-VCSEL M1235	5 ML PbSe dots (236 periods)	$Pb_{0.94}Eu_{0.06}Te$ 48 nm	λ =3.82 μm L =14 μm	$Pb_{0.94}Eu_{0.06}Te$ 1800 nm	EuTe: 370 nm PbEuTe: 160 nm

All three VCSELs were characterized by FTIR transmission measurements and the resulting 300K transmission spectra are shown in Figure 15(a)-(c). Owing to the strong absorption of the active medium above the stop band center, the higher energy parts of the spectra are shown on an enlarged scale. The spectrum of the bulk-like VCSEL M1504 (Figure 15(a)) exhibits three cavity resonance peaks of m= fourth-, fifth- and sixth-order at 244 meV, 285 meV and 325 meV, with line widths of 2.9, 1.0 and 2.7 meV, respectively. The same is observed for the QW-VCSEL M1208 shown in Figure 15(b) with resonance peaks at 291, 350 and at 421 meV, but due to the shorter cavity length, the mode spacing of 65 meV is somewhat larger than that of the bulk-like VCSEL. For the QD VCSEL M1235 (Figure 15(c)), a large number of cavity resonance peaks around the central 325 meV mode are observed in the stop band with line widths between 0.5 and 1.0 meV. The significantly narrower 8 meV mode spacing is due to the much larger cavity thickness of 14 μm.

6.2 Laser Emission

For laser pumping and spectroscopic emission measurements, the same optical setup was used. The first Stokes shifted line of a Q-switched Nd:YAG laser, produced in a hydrogen Raman cell [95] was used as pump source, resulting in a 1.907 μm pump wavelength and 10 ns pulses. The pump beam was focused on the sample to a 320 μm spot size and the emission was recorded with an InSb detector through a grating spectrometer. As shown in Figure 15(d)-(f), all three VCSELs show laser emission at photon energies around 325 meV or 3.8 μm wavelength.

6.2.1 Bulk-like PbTe VCSEL

For the bulk-like VCSEL M1504, the emission spectra of Figure 15(d) recorded at constant pump power of 96 kW/cm^2 show two lines corresponding to the m = fifth and sixth cavity mode of Figure 15(a). Owing to the increase of the PbTe band gap from 217 meV at 77 K to 320 meV at 300 K, the fifth cavity mode at 282 meV (4.4

μm) is excited only at temperatures below 207 K, where the PbTe band gap energy is below the cavity mode. The m = fifth emission reaches a maximum intensity at about 195 K where the PbTe band gap coincides exactly with the cavity mode. At higher temperatures, the PbTe band gap shifts to higher energies and therefore, this mode is quenched and the m = sixth laser emission at 330 meV or 3.75 μm appears. The two modes are never observed simultaneously because the 40 meV mode spacing is much bigger than the typical 10 meV width of the PbTe spontaneous photoluminescence. At 290 K, the sixth order laser mode reaches its maximum intensity and its 1.7 meV width is a factor of 1.8 smaller than that of the resonance peak in the transmission spectrum, clearly indicating a narrowing due to stimulated emission.

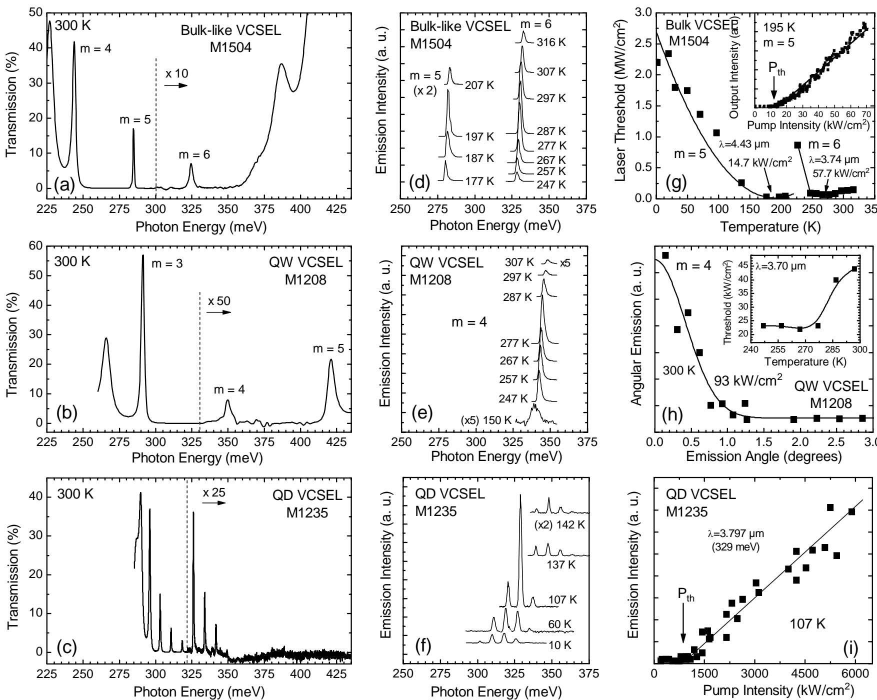


Figure 15. Optical properties and laser characteristics of three lead salt VCSELs with similar cavity structure but different active regions (see Table 3): *Top row*: VCSEL M1504with bulk-like PbTe as active region. Center row: QW VCSEL M1208 with nine 20 nm PbTe quantum wells. *Bottom row*: QD-VCSEL M1235 with a self-assembled PbSe quantum dot superlattice. *Left hand column* (**a-c**): Room temperature transmission of the VCSELs showing the stop band region with the cavity resonance modes. *Center column* (**d-f**): Laser emission spectra at different operation temperatures. *Right hand column*: (**g**) Threshold pump power as a function of temperature for the bulk-like VCSEL. The *inset* shows the output intensity at T= 195K *vs* pump power. (**h**) Angular emission behaviour of the QW-VCSEL: emitted intensity as a function of emission angle at T=300 K and 93 kW/cm^2 pump power together with the Gaussian line fit (*solid line*). The *insert* shows the laser threshold as a function of temperature. (**i**) Output intensity of the QD-VCSEL *vs* pump power at T = 107 K. See [38] for details.

At the different operation temperatures, the laser threshold was obtained by measuring the output intensity as a function of the pump power. In all cases, an essentially linear increase of the output power was observed above the threshold as is shown in the insert of Figure 15(g) for $T = 195$ K. The temperature dependence of the laser threshold is plotted in Figure 15(g). At low temperatures, the threshold intensity is very high because the cavity resonance is much above the PbTe band gap. However, when the optimum temperature for excitation of the m = fifth mode is approached, the threshold decreases and reaches a minimum of 14.7 kW cm^{-2} at 195 K (insert of Figure 15(g)). The m = sixth mode, on the other hand, has a minimum threshold of 57.7 kW/cm^2 at a temperature of 290 K, above which the threshold gradually increases due to non-radiative recombination processes. Because the top mirror of the VCSEL reflects about 73 % of the pump laser, the internal thresholds are even smaller at 4 kW/cm^2 and 15.6 kW/cm^2 at 195 and 290 K, respectively. These values are much smaller than those reported for III-V mid infrared VCSELs of 235 kWcm^{-2} at 260 K [49] as well as those reported for a similar bulk-like PbSe/PbSrSe VCSELs of about 70 kW/cm^2 at $T = 239$ K [31]. The highest laser operation temperature for the VCSEL M1504 was 317 K, above which the PbTe band gap shifts above the m = sixth resonance energy of 330 meV.

6.2.2 PbTe/PbEuTe Quantum Well VCSEL

The emission spectra observed for the multi quantum well VCSEL M1208 are shown in Figure 15(e) for a constant pump power of 93 kW/cm^2. In this case, only one laser emission mode at m = fourth is observed because the m = third and fifth modes are too close to the edge of the Bragg mirror stop band where the cavity finesse is too small to sustain stimulated emission. As shown by Figure 15(e), laser emission for this sample is observed over a temperature range from 247 to 307 K. The emission seen at 150 K is very weak and broad and therefore is attributed to spontaneous QW emission. At a temperature of 247 K, the laser turns on at a photon energy of 345 meV or 3.6 μm wavelength. As shown in Figure 15(e), it reaches a maximum intensity at 277 K and persists up to 307 K. At this temperature, the lower edge of the QW gain spectrum shifts out of the cavity resonance energy thus quenching the laser operation.

From the evaluation of the emission spectra, again the FWHM of the laser line is found to be significantly smaller than the resonance width in the FTIR transmission spectrum. The maximum emission intensity of the QW VCSEL is of the same order of magnitude as that of bulk-like VCSEL M1504. As shown in the insert of Figure 15(h), the laser threshold of the QW VCSEL is nearly constant at a value of 22.9 kW/cm^2 in the temperature range from 247 to about 280 K, and then increases to 44 kW cm^{-2} at 297 K. Taking again into account the reflectivity of the top mirror for the pump laser of 58 %, the internal threshold is 9.6 kW/cm^2 at 267 K and 18.5 kW/cm^2 at 297 K. Figure 15(h) also shows the angular emission behaviour of this VCSEL by plotting the relative output intensity as a function of emission angle. Obviously, the measured laser intensity essentially drops to zero for emission angles larger than 1°, demonstrating an extremely narrow beam divergence of less than 1°. The same characteristics were also observed for the other lasers of this sample series.

6.2.3 Lasing Properties of the Quantum Dot VCSEL

The emission spectra of the QD VCSEL M1235 at different operation temperatures are shown in Figure 15(f). Due to the 14 μm cavity length, the mode spacing of 8meV is so small that simultaneous lasing occurs on two to three modes at each temperature. Therefore, the width of the underlying inhomogeneously broadened spontaneous quantum dot emission can be estimated to be roughly 20 meV, in agreement with the results of Section 5.3. As shown in Figure 15(f), the second lowest laser line at 319 meV or $\lambda = 3.92$ μm reaches its maximum intensity at 60 K and the next higher mode at 329 meV ($\lambda = 3.8$ μm) at 107 K with an external threshold of 800 kW/cm^2 (see Figure 15(i)). This is significantly higher than the thresholds of the other two VCSELs. An explanation for this observation could be a possible type II band alignment between the PbSe dots and the PbEuTe barrier layers. This would strongly decrease the optical matrix elements due to the reduced overlap of the electron and hole wave functions. Up to now, an experimental determination of the band alignment in this strained heterosystem is still lacking, but in this case higher barriers with higher Eu-content are required for improving the laser performance. Another possibility to be considered is the reduction of the PL emission efficiency due to intermixing of the PbSe dots with the surrounding matrix material. Thus, the highest temperature for laser emission was only 146 K for this QD VCSEL.

6.2.4 Tuning Properties

As shown in Figure 15, the emission lines of all three lasers show a notable blue shift with increasing temperature. This is due to the decrease of the refractive index of the cavity material with increasing temperature, which results in an increase in the optical cavity length. For the bulk-like VCSEL, the m = fifth mode changes by about 2 meV and the m = sixth mode by about 4 meV in the temperature range shown in Figure 15 (d). This yields as tuning range of +0.08 and +0.059 meV/K, respectively. Thus, the laser wavelength changes from 4.431 to 4.393 μm or –1.26 nm/K for the fifth order mode and from 3.779 to 3.733 μm or –0.67 nm/K for the sixth order laser mode. For the QW VCSEL, the laser line changes by 6 meV from $T = 247$ to 307 K. This tuning rate of +0.091 meV/K corresponds to a wavelength change from 3.623 to 3.565 μm and a tuning rate of –0.967 nm/K. For the laser modes of the QD VCSEL, similar values for the tuning rate of around +0.05 meV/K are obtained. thus these lasers are indeed suited for spectroscopic applications.

The bulk-like and the QW PbTe VCSELs show very similar thresholds as well as above room temperature operation, whereas lasing from self-assembled PbSe quantum dots was restricted to below 146 K. Thus, up to now, the reduction of the dimensionality has not yet resulted in a comparable performance improvement as witnessed for III-V lasers. Since bulk-like and QW lead salt VCSELs show a comparable performance, it seems that the nonradiative recombination due to defects is still the major limiting factor for the laser process at higher operation temperatures.

7 CW-VCSELs Emitting at 6-8 µm

For spectroscopic applications, CW operation of infrared lasers with narrow emission line widths is preferred. CW operation generally requires lasers with low thresholds to prevent overheating effects. In addition, for optically pumped lasers suitable high power CW pump sources are required. For demonstration of principle, we have employed a 5.28 µm CO laser for this purpose, pumping specially designed VCSELs with very low thresholds of 25 W/cm^2. Because in our previous photoluminescence studies [69] PbSe has shown superior luminescence efficiency as compared to PbTe, Se-based compounds were used for all layers of these VCSELs, in contrast to the PbTe based lasers described above. The VCSELs thus, consist of PbSe active cavity regions sandwiched in between two high-reflectivity EuSe/ $Pb_{0.94}Eu_{0.06}Se$ Bragg mirrors. Due to the refractive index of EuSe of 2.5 in the 5 – 10 µm wavelength range compared to ~ 4.3 for $Pb_{0.94}Eu_{0.06}Se$, a refractive index contrast of 50 % is obtained. This yields a reflectivity above 99 % with five Bragg mirror layer pairs.

7.1 Structure and Optical Properties

Two different CW-VCSEL structures were designed and fabricated for different operation temperatures and emission wavelengths. The first one (M1753) was optimized for operation at liquid He temperature. Thus, the cavity resonance wavelength was set to 7.9 µm to match the PbSe band gap emission at 4 K. The second VCSEL M1781 was designed for operation at 85 K with a corresponding PbSe emission at about 6.7 µm. Each VCSEL consisted of five period top and bottom EuSe /$Pb_{0.94}Eu_{0.06}Se$ Bragg mirrors and a two-wavelength cavity of a $Pb_{0.94}Eu_{0.06}Se$ buffer layer followed by PbSe as active material. The cross sectional electron microscopy image of the 6.7 µm VCSEL is shown in the insert of Figure 16(b) and the structural parameters of the two lasers are summarized in Table 4. The cavity design ensures that the 5.28 µm pump laser is not absorbed by the top mirror and does not coincide with the mirror stop band.

Table 4. Parameters of 6 to 8 µm CW-PbSe VCSELs with 2λ cavity of a $Pb_{0.94}Eu_{0.06}Se$ buffer and a PbSe active layer and two five period EuSe/ $Pb_{0.94}Eu_{0.06}Se$ Bragg mirrors

CW-VCSEL	Target wavelength	Active region of PbSe	Cavity buffer $Pb_{0.94}Eu_{0.06}Se$	EuSe λ/4 layer	$Pb_{0.94}Eu_{0.06}Se$ λ/4 layer
M1753	7.9 µm	1434 nm	2395 nm	814 nm	487 nm
M1781	6.7 µm	1121 nm	2161 nm	735 nm	443 nm

Figure 16(a) shows the reflectivity spectrum of the 7.9 µm VCSEL measured at 300 K. At this temperature, the spectrum exhibits three narrow cavity resonances of m = third-, fourth-, and fifth-order at 133, 163, and 190 meV (9.33, 7.61, and 6.53 µm), respectively. The measured line width of the third resonance mode is 0.25 meV, whereas that of the central fourth order mode is only 0.10 meV, demonstrating a high effective cavity finesse of about 400, taking into account the order of the

mode. The calculated reflectivity spectrum, shown by the solid line in Figure 16(a), is in very good agreement with the experimental data. Owing to the decrease of the energy band gap of PbSe from 279 meV at 300 K to 146 meV at 2 K, at low temperatures the higher energy cavity resonances are damped by the interband PbSe absorption. Thus, at 60 K (see insert of Figure 16(a)), the fourth-order cavity resonance is broadened from 0.10 meV at room temperature to 0.13 meV. At even lower temperatures, the fourth-order resonance completely disappears because the PbSe active region in the cavity absorbs strongly.

Comparing the 300 K and 60 K spectra one can see that the spectral position of the resonances shifts to lower energies at a rate of 22 μeV/K due to the increasing refractive index of the cavity. Thus, at 60 K the cavity resonance is at 158 meV or 7.8 μm. For the 6.7 μm VCSEL M1781, the room temperature transmission spectrum is shown in Figure 16 (b). Again, three cavity resonance peaks are observed, but now at energies of 155, 190 and 223 meV or $\lambda = 8.0$, 6.53, and 5.23 μm, respectively. The line width of the central fourth-order cavity mode is only 0.17 meV or 7 nm, indicating a cavity finesse of about 300. As the temperature is lowered, again the resonance positions are red-shifted due to the increasing refractive index and below 100 K the central cavity mode becomes damped due to the PbSe absorption.

7.2 Laser Emission

Both VCSELs were optically pumped with a CW-CO laser emitting at a 5.28 μm wavelength (235 meV). The pump beam was focused onto the sample surface to a spot size of about 200 μm and the emitted light was detected with an HgCdTe detector through a grating spectrometer using the lock-in technique. The total emitted power was measured with a calibrated detector using an InSb long pass filter. Figure 16(c) shows the emitted laser line for CW as well as pulsed excitation with 400 mW pump power in both cases. At 2 K, the stimulated emission of the fourth-order resonance of the VCSEL M1753 is found at 156.51 meV *i.e.,* $\lambda = 7.92$ μm. This represents the longest emission wavelength of all VCSELs reported to date. For CW operation, an extremely narrow line width of only 12 μeV (0.6 nm) is observed, as found from the Lorentzian line fits shown as solid lines in Figure 16(c). This line width is 10 times smaller than the 130 μeV width found for the unpumped cavity resonance in the 60 K transmission measurements of Figure 16(a); and it is also much narrower than that observed for pulsed mode emission that is shown by the solid triangles in Figure 16(c). At 400 mW peak power excitation, the FWHM of the pulsed laser emission is 45 μeV (2.3 nm), but increases up to 100 μeV when the pump power is increased. This is still below the width of the unpumped cavity and is attributed to dynamical broadening effects [37]. For CW-excitation, the measured line width is exactly the resolution of the spectrometer setup. Thus, the true CW line width must be actually much smaller than 0.6 nm.

The emission spectra of the 6.7 μm VCSEL M1781 at 85 K are shown in Figure 16(d). At a low pump power of 43 mW, a weak emission at 6.685 μm with an 11 nm FWHM is found (open squares in Figure 16(d) on a strongly enlarged intensity scale). This corresponds to the expected wavelength and line width of the fourth-order emission filtered out by the cavity mode.

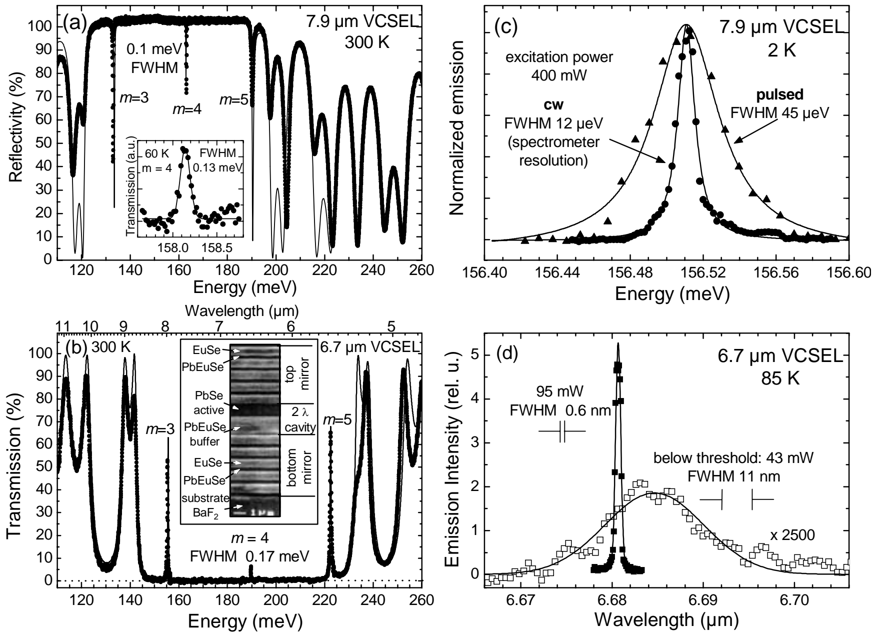


Figure 16. (**a**) Reflectivity spectrum of the CW-VCSEL M1753 designed for 7.9 µm wavelength, and (**b**) transmission spectrum of the CW-VCSEL M1781 designed for 6.7 µm, both measured at 300 K. The *solid lines* correspond to the theoretical spectra calculated using the transfer matrix method. The *insert* in (a) shows the transmission around the central cavity mode at 60 K on an enlarged scale, and the *insert* in (b) the cross-sectional scanning electron microscopy image of the VCSEL structure. *Right hand side*: (**c**) Emission spectra of M1753 at 2 K in CW (*dots*) and pulsed mode (*triangles*), both at 400 mW pump power. (**d**) Emission of VCSEL M1781 at 85 K at a pump power of 95 mW (*filled squares*) above threshold and 43 mW (*open squares*) below threshold. See [39,40,42] for details.

At a doubled excitation power of 95 mW (filled squares in Figure 16(d)), the emitted intensity increases drastically by a factor of 5000 and it is a factor of 20 narrower. The line width as determined by a Gaussian line fit (solid line) is only 0.6 nm or 16 µeV, again limited by the spectrometer resolution (denoted by –||– in the figure). Thus, above threshold we observe a very strong spectral narrowing not only with respect to the width of the passive cavity mode, but also compared to the spontaneous emission signal.

Figure 17(a) shows the CW power emitted by the 7.9 µm VCSEL as a function of excitation power. For weak excitation, only spontaneous emission (open circles) is observed. Above threshold, the emitted power rapidly increases (full circles). As can be seen by the inflection point of the S-shaped curve of the double logarithmic plot in Figure 17(a), the external laser threshold is 25 mW. Since the pump laser was focused to a diameter of 200 µm, this corresponds to an internal threshold of only 25 W/cm^2, where the 63 % reflectivity of the laser structure at the pump energy is taken into account. At an excitation power of 1.2 W, the CW power emitted through one

Bragg mirror is 4.8 mW. This is the highest observed from any CW VCSEL in the mid-infrared.

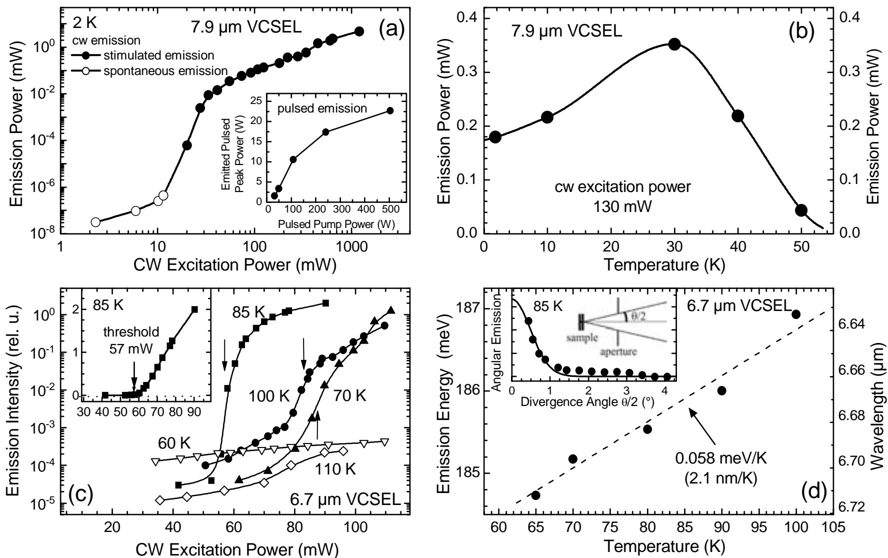


Figure 17. (**a**) CW output power of the 7.9 μm VCSEL M1753 plotted as a function of pump power at 2 K on a double logarithmic scale. *Open circles*: spontaneous emission, *full circles*: CW laser emission. The laser threshold is determined as 25 mW from the inflection point. *Inset*: pulsed laser output as a function of excitation power. (**b**) CW output power of the same VCSEL at a constant 130 mW pump power plotted *vs* operation temperature. (**c**) CW output intensity of the 6.7 μm VCSEL M1781 plotted as a function of pump power at temperatures between 60 and 110 K on a semi logarithmic scale: *full symbols*: stimulated emission, *open symbols*: spontaneous emission. The *insert* shows the output intensity *vs* pump power at 85 K on a linear scale. The laser threshold at each temperature is indicated by *arrows*. (**d**) Tuning characteristics of the 6.7 μm VCSEL: emission energy (*left scale*) and emission wavelength (*right scale*) plotted *vs* operation temperature. This yields a linear tuning coefficient of +0.058 meV/K or –2.1 nm/K (*dashed line*). The *insert* shows the measured angular emission characteristics of the laser (*dots*). The *solid line* represents the Gaussian line fit to the data points. See [39,42] for details.

For excitation powers higher than 1.2 W, pulsed mode excitation was used in order to prevent thermal damage of the sample. The inset of Figure 17(a) shows the emitted pulse peak power as a function of pump power. For 500 W pump pulses, a peak power of 23 W was emitted from one side of the VCSEL. To deduce the conversion efficiencies, the total emission of the structure from both sides of the VCSEL was taken into account. Due to the symmetric cavity design, the total emitted power is twice the power measured from one side. Therefore, the conversion efficiencies for CW and pulsed operation at maximum output powers are 0.8 % and 9 %, respectively. The maximum efficiency in pulsed mode is 20 %, at 100 W pump power. Below 100 W, the maximum slope efficiency determined by linear fits is 24 % whereas for CW operation a value of 1% was found.

The temperature dependence of the 7.9 μm VCSEL emission is shown in Figure 17(b) for a constant 130 mW excitation power. Up to 30 K, the spontaneous emission of the PbSe active material shifts closer to the fourth-order cavity mode. Thus, the emitted intensity increases. At 30 K, the maximum of the spontaneous emission coincides with the cavity resonance, manifested in best laser performance. Above 30 K, the emitted intensity is reduced because the maximum of the spontaneous emission shifts above the cavity mode. At 57 K, the laser emission is quenched due to complete detuning of resonance and spontaneous emission.

The threshold behaviour of the 6.7 μm VCSEL sample is shown in Figure 17(c) for different temperatures. Again, the data is plotted on a logarithmic scale to reveal also the sub-threshold signals. For clarity, the measurement for 85 K is also plotted in the inset on a linear scale. The threshold pump power was evaluated for operation temperatures of 70 (triangles), 85 (squares) and 100 K (circles) to be 87, 57 and 83 mW, respectively, as indicated by the arrows at the inflection points of the logarithmic plots. Thus, the minimum threshold is at the design temperature, where the maximum of the spontaneous emission and the cavity mode position coincide. The minimum pump threshold of 57 mW at 85 K corresponds to an internal threshold power density of 67 W/cm^2. In the logarithmic plots one can not only see the sub-threshold signals, but also the emission at 60 and 110 K (open symbols in Figure 17(c)), at which only weak and much broader spontaneous emission is observed. Thus, this VCSEL operates only at temperatures between 65 and 100 K because at higher as well as lower temperatures, the PbSe emission is above or below the cavity mode. The maximum observed total CW output power at 85 K is 1.2 mW at a pump power of 230 mW, which corresponds to a conversion efficiency of 0.5 %. We have also fabricated a third VCSEL with a design wavelength of 6.5 μm. This laser shows essentially the same performance as the 6.7 μm laser, but due to the shorter wavelength it lases up to a temperature of 120 K.

Recently, another group [96] has also reported CW emission from a PbSe/PbSrSe microcavity structure. Although a nonlinear increase of emission with increasing pump power was observed, the reported results do not clearly prove the stimulated nature of emission because the emission was rather broad (44 nm FWHM) and no line width narrowing was observed. Furthermore, apart from the emission from the active PbSe material, simultaneous emission from the PbSrSe buffer layer was also observed, with an intensity that was even higher than that attributed to laser emission. Therefore, in this case most probably only a spontaneous emission that is filtered by the microcavity resonance was actually observed

For applications of the lasers for molecular spectroscopy, the tuning ability and beam divergence are of crucial importance. These are characterized in detail in Figure 17(d) for the 6.7 μm VCSEL. The total tuning range by temperature variation was found to be 70 nm, with a linear tuning coefficient of +0.058 meV/K or –2.1 nm/K. This large tuning range is particularly advantageous for spectroscopy and it is substantially larger as compared to that of III-V quantum cascade lasers operating in the same wavelength range with a typical tuning coefficient of 0.01 meV/K [97]. In addition, we have recently shown that the emission wavelength can also be tuned quite elegantly by applying external magnetic fields [41]. In this case, we have found that the laser emission splits into two circularly polarized lines with opposite

polarization. Therefore, the lasing action involves only spin polarized electronic states, *i.e.*, these VCSELs can be viewed as a novel kind of spin laser. We have also determined the angular emission characteristic of the CW VCSELs. As demonstrated in the inset of Figure 17(d) for an emission angle of only 1° off the surface normal (see sketch in the inset), the emission intensity (filled squares) essentially drops to zero, and the Gaussian line fit to the emission profile (solid line) yields a half width at half maximum of only 0.5°. This narrow forward directed emission is a clear advantage of the VCSELs as compared to edge emitters. Up to now, the CW-operation temperatures were limited by the pump laser wavelength as well as the temperature dependence of the PbSe photoluminescence emission. It is therefore anticipated that already minor changes in the cavity structure and active region design as well as the use of other pump sources will soon lead to higher CW-laser operation temperatures.

8 Conclusions

In conclusion, lead salt vertical-cavity surface-emitting lasers offer attractive properties as coherent infrared laser sources. In particular, they feature single mode operation, emit circularly shaped parallel beams with extremely small beam divergence, and exhibit very sharp emission lines widths below 12 μeV. Based on the use of high finesse infrared microcavity structures, pulsed mode operation has been achieved well above room temperature as well as CW-operation up to 120 K, which is expected to be significantly increased in the near future. In comparison with quantum cascade lasers, the lead salt VCSELs show a substantially larger wavelength tunability, which is of crucial importance for spectroscopy applications. In addition, lead salt VCSELs can be grown on readily available substrate materials. This not only drastically reduces costs and facilitates the laser fabrication, but also offers improved heat dissipation due to the higher substrate thermal conductivity. Up to now, lead salt VCSELs have operated only under optical excitation but with some technological efforts electrically pumped lasers should also become feasible. Alternatively, one can envision hybrid structures in which low cost and readily available NIR pump lasers are integrated in one package with the lead salt VCSELs to obtain easy to use and cost-efficient mid-infrared laser sources. This would certainly open many promising applications in a variety of different fields.

Acknowledgements

We thank J. Fürst, M. Aigle, H. Pascher, G. Bauer, M. Böberl, E. Kaufmannn, J. Roither, H. Krenn, K. Biermann and K. Reimann for their collaboration, as well as the Fonds zur Förderung der Wissenschaftlichen Forschung and its START program and the Gesellschaft für Mikroelektronik, Vienna, for financial support.

References

1. Grisar R, Böttner H, Tacke M, Restelli G (eds). Monitoring of Gaseous Pollutants by Tuneable Diode Lasers, Kluwer Academic Publishers, Dordrecht, 1992.
2. Tittel F K, Richter D and Fried A. In: Sorokina I T, Vodopyanov K L (eds) Solid State Mid-Infrared Lasers Sources. Springer-Verlag, Berlin, 2003, pp 445-510.
3. Preier H. Appl Phys 1979; 20:189-206.
4. Partin D L. IEEE J Quantum Electron 1988; 24:1716-1726.
5. Katzir A, Rosman R, Shani Y, *et al.* In: Cheo P K (ed) Handbook of Solid State Lasers, Marcel Dekker, New York, 1989, pp 227-347.
6. Preier H. Semicond Sci Technol 1990; 5:S12-S20.
7. Tacke M. Infrared Phys Technol 1995; 36:447-463; and Tacke M. In: Helm M (ed) Long Wavelength Infrared Emitters based on Quantum Wells and Superlattices, Gordon and Breach Science Publishers, Amsterdam, 2000, pp 347-396.
8. Ishida A and Fujiasyu H. In: Khokhlov D, Lead Chalcogenides: Physics and Applications Taylor and Francis, New York, 2003, pp 533-554.
9. Findlay P C, Pidgon C R, Kotitschke R, *et al.* Phys Rev B 1998; 58:12908-12915.
10. Springholz G, Shi Z, and Zogg H. In: Liu W K and Santos M B (eds) Thin Films: Heteroepitaxial Systems,World Scientific, Singapore, 1999, pp 621-688.
11. Springholz G, In: Khokhlov D, Lead chalcogenides: Physics and applications, Taylor and Francis, New York, 2003, pp. 123-207.
12. Partin D L and Heremans J. In Mahajan (ed) Handbook on Semiconductors Vol. 3a, North Holland, Amsterdam, 1994, pp. 369- 450.
13. Feit Z, Kostyk D, Woods R J, and Mak P. Appl Phys Lett 1991; 58:343-345.
14. Schießl U P and Rohr J. Infrared Phys Technol 1999; 40:325-328.
15. Bauer G, Kriechbaum M, Shi Z, *et al.* J Nonlinear Opt Phys Mat 1995; 4:283-312.
16. Faist J, Capasso F, Sivco D L, *et al.* Science 1994; 264:553-556.
17. Helm M (ed). Long Wavelength Infrared Emitters based on Quantum Wells and Superlattices, Gordon and Breach Science Publishers, Amsterdam, 2000, pp 1-12.
18. Gmachl C, Capasso F, Sivco D L, and Cho A Y. Rep Prog Phys 2001; 64:1533-1601.
19. Hofstetter D and Faist J. In: Sorokina I T and Vodopyanov K L (eds) Solid State Mid-Infrared Lasers Sources, Springer-Verlag, Berlin, 2003, pp 61-96.
20. Shani A, Katzir A, Bachem K H, *et al.* Appl Phys Lett 1986; 48:1178-1180.
21. Jewell J L, Huang K F, Tai K, *et al.* Appl Phys Lett 1989; 55:424-426.
22. Jewell J L, Harbison J P, Scherer A, *et al.* IEEE J Quantum Electron 1991; 27:1332-1346.
23. Iga K, Koyama F, and Kinoshita S. IEEE J Quantum Electron 1988; 24:1845-1855.
24. Koch S W, Jahnke F, and Chow W W. Semicond Sci Technol 1995; 10:739-751.
25. Blum O. In: Miller A, Ebrahimzadeh M, Finlayson D M (eds), Semiconductor Quantum Optoelectronics, Institute of Physics Publishing, London, 1999, pp 265-294.
26. Li H and Iga K, Vertical-Cavity Surface-Emitting Laser Devices, Springer-Verlag, Berlin, 2003.
27. Schwarzl T, Heiss W, Springholz G, *et al.* Electron Lett 2000; 36:322-324.
28. Springholz G, Schwarzl T, Aigle M, *et al.* Appl Phys Lett 2000; 76:1807-1809.
29. Bewley W W, Felix C L, Vurgaftman I, *et al.* Electron Lett 2000; 36:539-540.
30. Shi Z, Xu G, McCann P J, *et al.* Mater Res Soc Symp Proc 2000; 607:181-185.
31. Shi Z, Xu G, McCann P J, *et al.* Appl Phys Lett 2000; 76:3688-3690.
32. Springholz G, Schwarzl T, Heiss W, *et al.* J Crystal Growth 1999; 201/202:999-1004.

33. Schwarzl T, Heiss W, and Springholz G. Appl Phys Lett 1999; 75:1246-1248.
34. Schwarzl T, Springholz G, Seyringer H, *et al.* IEEE J Quantum Electron 1999; 35:1753-1758.
35. Heiss W, Schwarzl T, Roither J, *et al.* Prog Quantum Electron 2001; 25:193-228.
36. Felix C L, Bewley W W, Vurgaftman I, *et al.* Appl Phys Lett 2001; 73:770-3772.
37. Heiss W, Schwarzl T, Springholz G, *et al.* Appl Phys Lett 2001; 78:862-864.
38. Fürst J, Pascher H, Schwarzl T, *et al.* Appl Phys Lett 2002; 81:208-210.
39. Fürst J, Schwarzl T, Böberl M, *et al.* Appl Phys Lett 2004; 84: 3268-3270.
40. Fürst J, Schwarzl T, Böberl M, *et al.* Electron Lett 2004, 40:966-968.
41. Fürst J, Pascher H, Schwarzl T, *et al.* Appl Phys Lett 2005; 86:021100-2.
42. Schwarzl T, Springholz G, Böberl M *et al.* Appl Phys Lett 2005; 86:031102-4.
43. Springholz G, Schwarzl T, Heiss W, *et al.* Appl Phys Lett 2001; 79:1225-1227.
44. Whitaker T. Compound Semiconductors, 1998; Winter II:18-29.
45. Redwing J M, Loeber D A S, Anderson N G, *et al.* Appl Phys Lett 1996; 69:1-3.
46. Krestnikov I L, Lundin W V, Sakharov A V, *et al.* Appl Phys Lett 1999; 75:1192-1194.
47. Koeth J, Dietrich R, and Forchel A. Appl Phys Lett 1998; 72:1638-1640.
48. Baba T, Yogo Y, Suzuki K, Koyama F, and Iga K. Electron Lett 1993; 29:913-914.
49. Felix C L, Bewley W W, Vurgaftman I, *et al.* Appl Phys Lett 1997; 71:3483-3485.
50. Hadji E, Bleuse J, Magnea N, Pautrat J L. Appl Phys Lett 1996; 68:2480-2482.
51. Roux C, Hadji E, and Pautrat J L. Appl Phys Lett 1999; 75:3763-3765.
52. Gerard A, Burch J M. Introduction to Matrix Method in Optics, John Wiley and Sons, 1975; or Le Roy-Brehonnet F and Le Jeune B. Prog Quant Electron 1997; 21:109-151.
53. Rakic A D, Majewski M L. In: Li H and Iga K, Vertical-Cavity Surface-Emitting Laser Devices, Springer-Verlag, Berlin, 2003, pp 259-301.
54. Zemel J N, Jensen J D, and Schoolar R B. Phys Rev 1965; 140:A330-A342.
55. Bauer G and Krenn H. In: Palik E D (ed) Handbook of Optical Constants of Solids, Academic Press, New York, 1985, pp 517-523, 535-546.
56. Yuan S, Krenn H, Springholz G, Bauer G. Phys Rev B 1993; 47:7213-7226.
57. Yuan S, Springholz G, Bauer G, Kriechbaum M. Phys Rev B 1994; 49:5476- 5489.
58. Krenn H, Yuan S, Frank N, Bauer G, Phys Rev B 1998; 57:2393-2401.
59. Majumdar A, Xu H Z, Zhao F. J Appl Phys 2004; 95:939-943.
60. Springholz G, Bauer G. Appl Phys Lett 1993; 62:2399-2401.
61. Frank N, Springholz G, Bauer G. Phys Rev Lett. 1994; 73:2236-2239.
62. Zhao F, Wu H, Jayasinghe L, Shi Z. Appl Phys Lett 2002; 80:1129-1131.
63. Wu H, Zhao F, Jayasinghe L, Shi Z. J Vac Sci Technol B 2002; 20:1356-1360.
64. Fang X M, Wu H Z, Shi Z, *et al.* J Vac Sci Technol B 1999; 17:1297-1300.
65. Stanley R P, Houdre R, Oesterle U, *et al.* Appl Phys Lett 1994; 65:1883-1885.
66. Holloway H and Walpole J N. Prog. Crystal Growth Charact 1979; 2:49-94.
67. Springholz G, Bauer G and Ihninger G. J Cryst Growth 1993; 127:302-307; and Ueta A Y, Springholz G and Bauer G. J Cryst Growth 1997; 175/176:1022-1027.
68. Springholz G, Kriechbaum M, Hofmann W, *et al.* Superlatt Microstruc 1993; 13:25-28.
69. Böberl M, Heiss W, Schwarzl T, *et al.* Appl Phys Lett 2003; 82:4065-4067.
70. Kaindl R A, Wurm M, Reimann K, *et al.* J Opt Soc America B 2000; 17:2086-2092.
71. Klann R, Hofer T, Buhleier R, *et al.* J Appl Phys 1995; 77:277-283.
72. Bimberg D, Grundmann M and Ledentsov N N, Quantum Dot Heterostructures, Wiley, Chichester, 1998.
73. Asada M, Miyamoto Y and Seimitsu Y. IEEE J Quantum Electron 1986; 22:1915-1919.

74. Arakawa Y and Sakaki H. Appl Phys Lett 1982; 40:939-941.
75. Kirstaedter N, Ledentsov N N, Grundmann M, *et al.* Electron Lett 1994; 30:1416-1418.
76. Leonard D, Krishnamurty M, Reaves C M. Appl Phys Lett 1993; 63:3203-3205.
77. Ratsch C and Zangwill A. Surf Sci 1993; 293:123-126. Srolowitz D J. Acta Metall 1989; 37:621-625.
78. see, e.g., Shchukin V A, Ledentsov N N, and Bimberg D, Epitaxy of Nanostructures, Springer Verlag, Berlin, 2003, and references therein.
79. Huang X, Stintz A, Hains C P, *et al.* Electron Lett 2000; 36:41-43.
80. see, e.g., Ledentsov N N. In: Grundmann M (ed) Nano-Optoelectronics, Springer Verlag, Berlin, 2002, pp 317-337.
81. Grundmann M. Physica E 2000; 5:167-181, and references therein.
82. Pinczolits M, Springholz G, Bauer G. Appl Phys Lett 1998; 73:250-252.
83. Pinczolits M, Springholz G, Bauer G. J Crystal Growth 1999; 201/202:1126-1130.
84. Raab A, Springholz G. Appl Phys Lett 2000; 77:2991-2993.
85. Raab A, Springholz G. Appl Phys Lett 2002; 81:2457-2459.
86. Springholz G, Holy V, Pinczolits M, Bauer G. Science 1998; 282:734-737.
87. Pinczolits M, Springholz G, Bauer G. Phys Rev B 1999; 60:11524-11529.
88. Springholz G, Pinczolits M, Mayer P, *et al.* Phys Rev Lett 2000; 84:4669-4672.
89. Raab A, Lechner R T and Springholz G. Appl Phys Lett 2002; 80:1273-1275.
90. Holy V, Springholz G, Pinczolits M and Bauer G. Phys Rev Lett 1999; 83:356-359.
91. Lechner R T, Schülli T, Holy V, *et al.* Appl Phys Lett 2005; 84:885-888.
92. Krenn H, Herbst W, Pascher H, *et al.* Phys Rev B 1999; 60:8117-8122.
93. Schwarzl T, Heiss W, Springholz G, *et al.* Phys Rev B 2002; 65:245321-9.
94. Alferov Z. In: Mirua N and Ando T (eds), Proc. 25th Int Conf Phys Semicond, Osaka 2000, pp 14-21 .
95. Heuer W and Zacharias H. IEEE J Quant Electron 1988; QE-24:2087-2100.
96. Zhao F, Wu H, Majumdar A and Shi Z. Appl Phys Lett 2003; 83:5133-5135.
97. Allen T, Blaser S, Beck M, *et al.* Appl Phys Lett 2003; 83:1929-1931.

Optically Pumped MIR Lasers

R. Kaspi, G.C. Dente and A.P. Ongstad
Air Force Research Laboratory, Directed Energy Directorate
Kirtland AFB, Albuquerque, NM USA

1 Introduction

Optical pumping, rather than electrical injection, can be employed to establish population inversion in the active region of semiconductor lasers. Such an optically pumped laser is an optical converter, converting incoherent pump light into a highly coherent output beam.

While far less common in practice relative to electrical injection, optical pumping offers several distinct advantages that can be exploited to produce lasers with high efficiency and good beam quality. For example, the photo-generation of carriers in a semiconductor laser heterostructure is preferred in cases where factors within the heterostructure, such as band-alignment or conductivity, are not favorable to electrical injection. In addition, the geometry of optical excitation can provide flexibility in controlling the lateral mode. The availability of high power linear diode arrays at various wavelengths makes the optical pumping scheme even more attractive. Bandgap engineering techniques are then employed to design the semiconductor heterostructure such that wavelength conversion occurs while achieving efficient photon-to-photon conversion rates.

A number of studies have demonstrated optically pumped lasers that emit in the mid-infrared wavelength region [1-16]. The 3-5μm range has proven to be particularly difficult to produce high power emission using electrical injection, and therefore has benefited from optically pumped laser designs.

Goodhue *et al.* discuss a generic design for a diode-laser pumped semiconductor laser concept referred to as SHOC (semiconductor heterostructure optical conversion) [17,18]. In this chapter, we describe in detail the embodiment of such a laser where an array of diode lasers emitting in the near-IR pump a semiconductor laser whose emission is in the mid-infrared.

The laser heterostructure is based on the GaSb family of III-V alloys and optical gain is generated by Type-II electronic transitions in the active region. We show that very high photon-to-photon conversion efficiencies and excellent beam-quality can be achieved when the laser is properly designed. Moreover, the optical pump as well as the absorbing medium can remain unchanged even when the active region is redesigned to emit at any wavelength in the ~2 to ~10 micron range.

2 Optically Pumped Laser Design

2.1 Guidelines for Optical Pumping and the MIR-OPSL

Consider an optically pumped semiconductor laser in which pump absorption occurs in layers that are integrated into the heterostructure. It is possible that these layers

also generate gain, such as in an optically pumped double-heterostructure laser [19,20]. Alternatively, the absorbing layers are placed to make photo-generated carriers drift into adjacent layers where gain is generated. In this section, we give details of a mid-IR laser heterostructure following the latter scheme, where mid-IR quantum wells are inserted into the absorbing medium and together they form the core of the waveguide. We refer to this device as a “mid-infrared optically pumped semiconductor laser” (MIR-OPSL), and note that it embodies many of the SHOC principles. A device of this Type was first demonstrated by researchers at the Lincoln Laboratory [16].

A generic design for an optically pumped device is shown in Figure 1. The design principles that must be followed when optimizing a MIR-OPSL laser are described below, and these are accompanied with a description of the specific choices we have made for our GaSb-based devices.

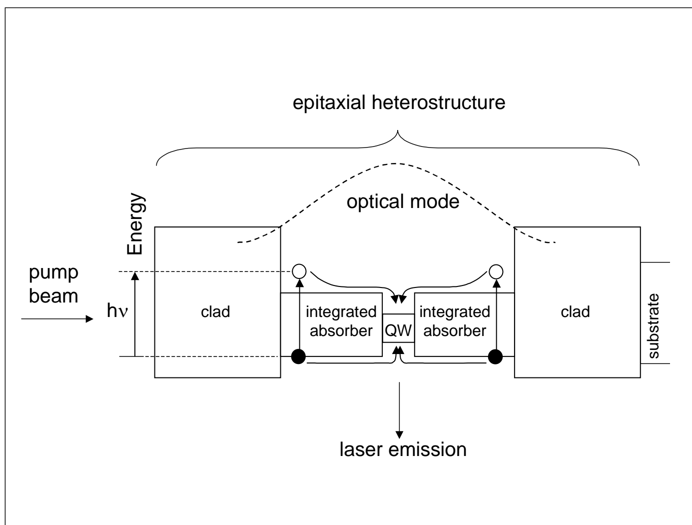


Figure 1. Generic design for optically pumped semiconductor laser

The first design principle is the proper identification of the pump wavelength in relation to the absorbing layers in the heterostructure. Naturally, the pump photon energies must be higher in energy than the absorbing layer bandgap, *i.e.* $h\nu - E_g$(absorber)>0. In designing the semiconductor heterostructure, care must be taken such that the pump beam primarily illuminates the intended absorbing region without being attenuated elsewhere in the heterostructure. This dictates that the typically thick clad layer must be transparent to the pump wavelength, *i.e.*, $h\nu - E_g$ (clad)<0. For substrate side pumping of the heterostructure, $h\nu - E_g$(substrate)<0 must also hold true.

In our case, we select a pump emission wavelength of $1.8<\lambda<2.0$ μm ($620<h\nu<690$ meV) due to the availability of semiconductor diode laser bars that exhibit high efficiency at room temperature. Consequently, both the GaSb substrate (E_g ~ 710 meV at RT) as well as any lattice-matched $Al_xGa_{1-x}As_ySb_{1-y}$ alloy candidate for the clad layers transmits the pump emission with little attenuation. On the other hand, we use the lattice-matched quaternary $(GaSb)_x(InAs_{0.89}Sb_{0.11})_{1-x}$ alloy as the integrated absorber layer. For x=0.8 *i.e.* $In_{0.2}Ga_{0.8}As_{0.18}Sb_{0.82}$, E_g(absorber)~590 meV, resulting in high absorption.

A second design principle is the proper choice of overall thickness of the absorbing medium in relation to the pump wavelength, such that a reasonably high density and a reasonably good distribution of carriers into the active region can be achieved. The absorption coefficient, α of a typical direct bandgap III-V semiconductor alloy is approximately given by $\alpha(h\nu)\sim 2\times10^5(m_c/m)^{1/2}(h\nu-E_g)^{1/2}$ cm^{-1}; for $In_{0.2}Ga_{0.8}As_{0.18}Sb_{0.8}$ we find $(m_c/m)\sim0.034$ and α(646meV)~8730 cm^{-1}. An exponential drop in absorption with layer thickness dictates that, for a given absorption coefficient, the absorbing volume must be optimized. Alternatively, a resonant cavity may be built to effectively capture the pump emission [21,22].

For the current embodiment of the MIR-OPSL, we use $In_{0.2}Ga_{0.8}As_{0.18}Sb_{0.82}$ as the absorbing medium, and from FTIR absorbance measurements find that α(646meV)~9200 cm^{-1}; this compares well with the calculated value. Since a highly reflective metallic layer typically lies beneath the laser, we assume that the pump beam effectively makes a double-pass through the heterostructure. This suggests that in order to absorb more than 95% of the pump energy, the absorption layers would need to have a total thickness of approximately 1.5 μm.

A third design concern, in the case of lasers with quantum wells imbedded in the absorbing media, is the relative band alignment between these such that the transfer of photo-generated carriers into the wells is favored.

In the current MIR-OPSL design, we have opted to use Type-II quantum wells for their flexibility in designing for a wide range of emission wavelengths. A simple Type-II well that exhibits good electron/hole wave-function overlap consists of a "W" arrangement in which a hole well, typically $In_xGa_{1-x}Sb$, is surrounded by two coupled InAs electron wells. Given the choice of $In_{0.2}Ga_{0.8}As_{0.18}Sb_{0.82}$ as the absorbing medium, the band alignment is indeed favorable for the transfer of electrons into the two coupled InAs wells. Hole transfer into the hole well is somewhat more problematic, as it will strongly depend on the alloy composition as well as the thickness of the $In_xGa_{1-x}Sb$ layer. The interplay of the details of the quantum well in relation to the surrounding absorbing medium is critical to the MIR-OPSL design, and will be discussed in more detail in the next section.

A fourth design principle is to minimize, as much as possible, the quantum defect between the pump photon energy and the laser photon energy. This difference in energy inevitably translates to heat and degrades the laser. Our choice of pumping at $1.8<\lambda<2.0$ μm is the longest wavelength readily available in a high power stacked array format that also reasonably satisfies all other design criteria.

A final but critical design point must be considered when the absorbing medium is a substantial portion of the waveguide, as is the case with the current MIR-OPSL. The effective refractive index in the waveguide must allow for sufficient optical confinement for lasing. It has been shown that any lattice-matched $Al_xGa_{1-x}As_ySb_{1-y}$

layer used as a clad is able to provide optical confinement in a waveguide that is primarily $In_{0.2}Ga_{0.8}As_{0.18}Sb_{0.82}$ (n~3.85) [23]. Therefore there is a choice in designing an MIR-OPSL with either a high mode confinement factor using Al-rich clad layers (n~3.20 for $Al_{.9}Ga_{.1}As_{.07}Sb_{0.83}$), or a very low mode confinement factor using a binary GaSb (n~3.82) clad layers. While there are advantages and disadvantages to both, we opt for the low confinement factor in the current MIR-OPSL design [24,25]. One advantage to low confinement factor design is the improvement in lateral beam quality, and we discuss this in more detail below.

A schematic drawing with approximate band alignments of the particular MIR-OPSL that we describe in this chapter is shown in Figure 2. This design complies with all the criteria described above, and can be fine-tuned to emit within a wide range of wavelengths by varying the InAs layer thickness.

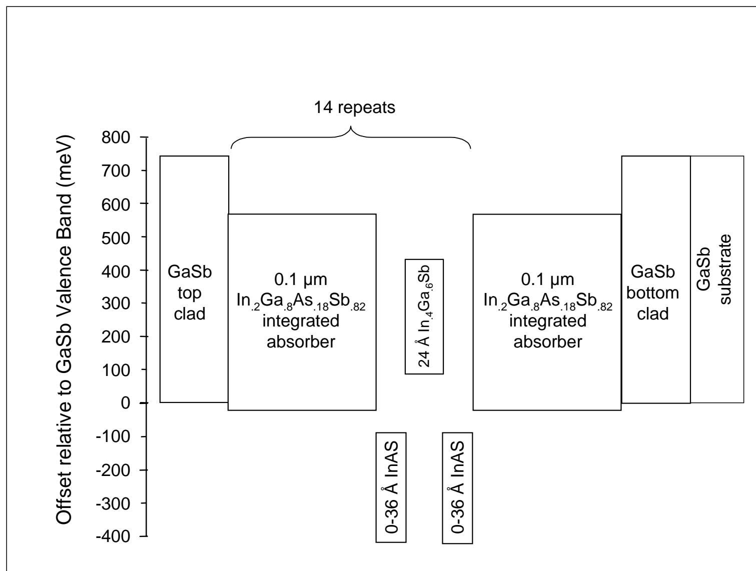


Figure 2. Schematic of MIR-OPSL described in this chapter

2.2 Type-II Quantum Wells

When two adjacent layers such as InAs and GaSb that exhibit a Type-II band alignment are formed, then an inter-band transition can occur where the emitted photon will have an energy that is smaller than either of the band-gaps associated with these layers. With the addition of quantum confinement as well as compositional alloying, a remarkable flexibility in the design of mid-infrared emission is achieved. The "W" arrangement of quantum wells, first suggested by Meyer *et al.* [26] is one in which holes in a compressively strained InGaSb well are confined by InAs barriers that in turn form a coupled quantum well for electrons. Surrounding layers may be used to enhance the overlap between the electron and the hole wave-functions. The ability of modern epitaxial technologies, such as molecular

beam epitaxy to form extremely thin layers with good precision is key to realizing such Type-II quantum well designs with adequate reproducibility.

In the case of MIR-OPSL lasers described here, we opt to use a 24 Å thick $In_{0.4}Ga_{0.6}Sb$ hole wells surrounded by InAs in order to achieve a large wave-function overlap. As shown in Figure 2, the InAs layers are in turn surrounded by the $In_{0.2}Ga_{0.8}As_{0.18}Sb_{0.82}$ absorbing layers. In order to predict the more salient features of this quantum well, we invoke the pseudopotential technique as a more accurate alternative to the effective mass approach.

Although the most widely used theoretical tools for calculating the electronic and optical properties of these materials have been based on **k.p** perturbation theories in combination with envelope function approximations, [27] their accuracy for Type-I and Type-II short period superlattices, remains in question [28,29]. Our superlattice empirical pseudopotential model (SEPM), as fully described in [28,29], makes the critical assumption that the heterointerface charges redistribute, forming charge and dipole sheets near the interface, so that the layer pseudopotentials remain as bulk-like as possible. This assumption allows us to form the SEPM potential of the superlattice using only the empirical pseudopotential (EPM) form factors for the constituent materials, as well as a band offset for each constituent pair; these form factors are determined by fitting the band diagram of each material in the superlattice.

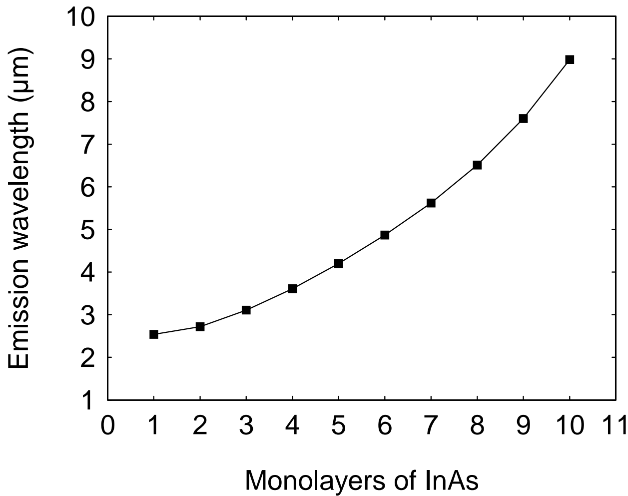


Figure 3. SEPM prediction of emission wavelength as a function of InAs layer thickness when $In_{0.2}Ga_{0.8}As_{0.18}Sb_{0.82}$ / InAs / 8ML $In_{0.4}Ga_{0.6}Sb$ / InAs / $In_{0.2}Ga_{0.8}As_{0.18}Sb_{0.82}$ is used as the mid-IR quantum well

The SEPM modeling is used to predict the emission wavelength for this quantum well structure as a function of InAs layer thickness. As shown in Figure 3, the wavelength is expected to vary from ~2.5 to ~9µm when the InAs electron well layer thickness is increased from 2 to 10 monolayers.

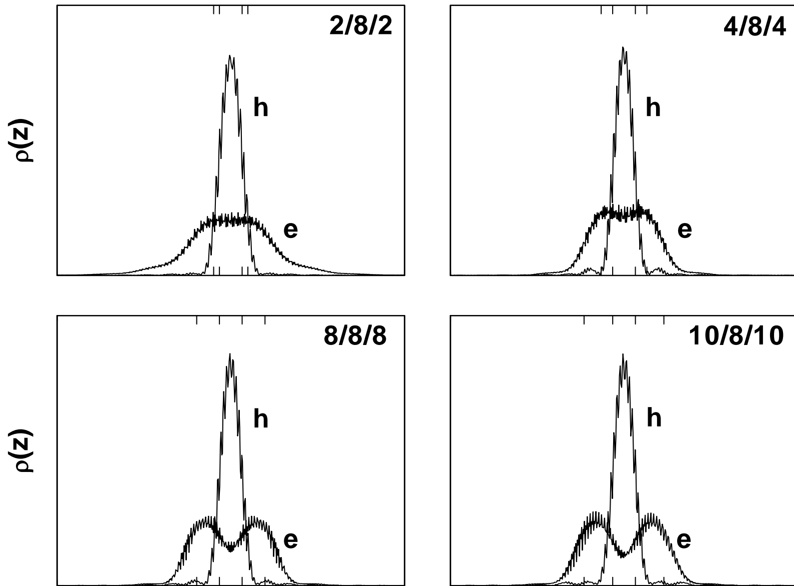


Figure 4. Wave-function planar density as a function of distance for the Type-II well when the InAs layer thickness is 2, 4, 8, and 10 monolayers. The position of the InAs layers is indicated by the *tick marks.*

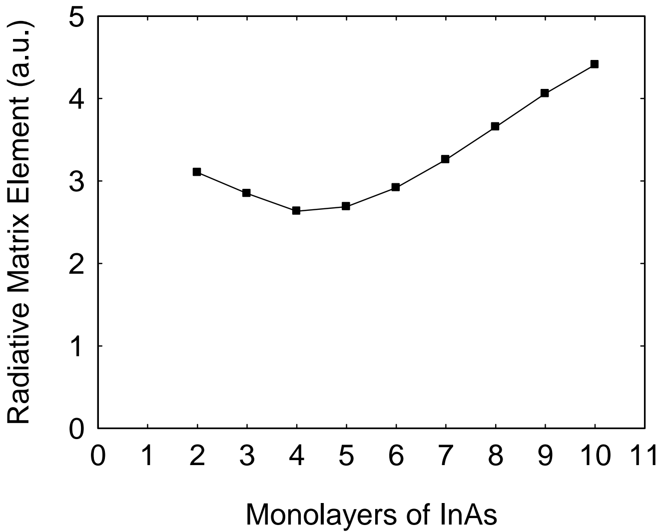


Figure 5. Calculated radiative matrix element for Type-II well as a function of InAs layer thickness

In addition, the SEPM calculations suggest that the electron and hole wave-function overlap is only moderately reduced as the InAs layer thickness is increased from 2 to 10 monolayers, spanning the entire range of interest. This is primarily due to the close proximity of the coupled electron wells that allow a consistently large portion of the electron wave-function to reside in the InGaSb hole well. In Figure 4, the planar density of both holes and electrons as a function of distance are shown for our different cases; the InAs layer thickness is 2,4,8, and 10 monolayers while the InGaSb hole well thickness is held fixed at 8 monolayers. As a result of the strong wave-function overlap, the radiative matrix element of this particular Type-II well is calculated to remain large even when it is redesigned to emit photons at longer wavelengths. This is shown in Figure 5.

In summary, the Type-II well presented here exhibits the flexibility to engineer the wavelength over a large range without an accompanying loss of gain.

2.3 Dielectric Waveguide Design for the MIR-OPSL

For many types of semiconductor lasers, low-threshold pumping is a primary design consideration. This one factor has motivated most dielectric waveguide designs in the past, leading to epitaxial designs with a significant index of refraction increase in the active region layers. This then gives a high confinement factor for the optical waveguide mode, which in turn leads to the typical large transverse angular divergence for the radiation propagating from the facet of the lasers. For longer wavelength lasers in the mid-IR region, this fast-axis divergence can present considerable optical engineering problems as the device is incorporated into a system. Furthermore, high confinement will generate large radiation intensity at the laser facet, contributing to concentrated facet heating and accelerated device degradation. Beyond these obvious transverse effects created by high confinement, it can also exacerbate lateral filament formation, leading to lateral beam quality degradations as well as further lifetime and facet degradations; the more subtle connection between confinement factor and filamentation is derived in the next section.

These effects, taken together, lead us to consider low-confinement factor MIR-OPSL semiconductor lasers, which offer manageable fast-axis divergence, improved lifetimes, reduced filamentation, and greatly improved lateral beam quality. Although we pay a small price in increased threshold, we gain considerably by incorporating low-confinement dielectric waveguides into our optically pumped mid-IR lasers. The following sections will detail both the benefits of low-confinement devices, along with several problematical side effects.

2.3.1 Connection Between Confinement Factor and Beam Quality

Although most semiconductor laser applications require only a few milliwatts to a few tens of milliwatts of power, there are applications, such as IRCM, that require much higher powers. The simplest method of obtaining higher output powers requires increasing the lateral width of the pumped region. However, filament formation brought on by self-focusing in the active region of broad-area devices

tends to produce spatially incoherent light at the laser output facet. This results in lateral far-field divergences that are many times diffraction-limited. Various methods have been invoked for controlling these filament-formation mechanisms and consequently promoting single lateral mode behavior. Index- and gain-guided arrays offered limited improvements [30]. Laterally tapered devices yielded improved beam quality, since the natural lateral divergence of the tapered devices tends to counteract the self-focusing filamentation process [31]. Other approaches included lateral index guides and lateral unstable resonators [32]. Although beam quality improvements were achieved with many of the lateral mode-control mechanisms, ultimately, at some modest level above threshold, lateral mode break-up through filamentation continued to degrade lateral beam quality.

Here, we analyze the filament formation process near steady-state in a broad-area semiconductor laser. First, we derive an expression for the filament gain at a given steady-state mode gain. In this expression, we show that while the mode gain decreases with the confinement factor in a linear manner, the gain for filaments decreases far more rapidly. This suggests that epitaxial designs promoting good beam quality should deliberately minimize the overlap of the active layer with the transverse waveguide mode.

We proceed by calculating the growth rate of small sinusoidal perturbations placed on the carrier distribution and the radiation field. The derivation closely follows [33]. Since most semiconductor laser active regions use quantum wells, it is convenient to quantify carriers as an areal density in the plane of the epitaxy. The steady-state carrier equation for the optically pumped laser is then given as

$$\frac{\eta_i P}{h\nu_P} - \Re(N) - \frac{G(N)\Gamma}{(h\nu)} U(x,z)U^*(x,z) + D\frac{\partial^2 N}{\partial x^2} = 0, \tag{1}$$

in which $N(x,z)$ = areal density (carriers/cm^2), η_i = injection efficiency, P = absorbed pump intensity, $h\nu_P$ = energy per pump photon, $h\nu$ = energy per laser photon, $\Re$ = recombination rate as a function of N, G = gain as a function of N, Γ = confinement factor and D = diffusion constant. Also, we neglect diffusion in the longitudinal z-direction since the significant spatial variations are only in the lateral x-direction. The local lateral intensity is given in units of watts per lateral cm as UU^*. These intensity units are particularly convenient for calculating broad-area semiconductor laser properties. For example, a 100-µm wide device emitting 1 W outcouples a lateral intensity of 100 W per lateral cm. The lateral field satisfies a paraxial wave equation given by

$$\left[\frac{1}{2ik}\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial z}\right]U = \frac{\Gamma}{2}[G(N)] + i\frac{2\pi}{\lambda}\Gamma n_c(N)\cdot U - \frac{a}{2}U, \tag{2}$$

in which $k = 2\pi n / \lambda$, with n the effective index of the mode, n_c = the carrier-induced index of refraction, and $a = a_{wg} + 1/L \ln(1/R)$ (waveguide loss and distributed mirror losses for equal facet reflectivities, R).

We place sinusoidal perturbations on the carriers and field as

$$N = \overline{N} + \varepsilon(z)\sin(k_F x)$$
$$U = U_0 + f(z)\sin(k_F x), \tag{3}$$

in which $\overline{N}$ = the steady-state areal density and U_0 = the steady-state complex field, while ε = the carrier perturbation and f = the field perturbation at spatial frequency k_F. We anticipate that these perturbations can vary as a function of propagation distance, z. We substitute these expressions into the carrier equation and the paraxial wave equation. Isolating terms that are linear in the perturbations yields

$$-\varepsilon\left\{\frac{\partial\Re}{\partial N} + \frac{\Gamma}{(h\nu)}\frac{\partial G}{\partial N} + Dk_F^2\right\} - \frac{\Gamma}{(h\nu)}G(\overline{N})\cdot U_0\left(f + f^*\right) = 0, \tag{4}$$

and

$$\left\{\frac{\partial f}{\partial z} - \frac{k_F^2}{2ik}f - \frac{\Gamma}{2}\frac{\partial G}{\partial N}U_0\varepsilon - i\frac{2\pi}{\lambda}\Gamma\frac{\partial n_c}{\partial N}U_0\cdot\varepsilon\right\} = 0\,. \tag{5}$$

We arrive at these equations by assuming that $\overline{N}$ and U_0 satisfy the steady-state equations with the gain equal to the local loss condition,

$$\Gamma G(\overline{N}) = a\,. \tag{6}$$

To further simplify Equation (5), we introduce the antiguiding parameter as

$$-\alpha \equiv \frac{2\pi}{\lambda}\frac{\partial n_c}{\partial N}\Bigg/\left(\frac{1}{2}\frac{\partial G}{\partial N}\right) \tag{7}$$

We first solve for the carrier fluctuation, ε, in terms of the field fluctuations, f, and substitute into the equation for the field. This gives an equation for the complex field fluctuations as

$$\frac{\partial f}{\partial z} = \frac{k_F^2}{2ik}f - \frac{B}{2}(1 - i\alpha)\left(f + f^*\right) \tag{8}$$

in which $B \equiv \dfrac{\Gamma\, I_0/I_s\, G(\overline{N})}{\left[1 + I_0/I_s + s^2k_F^2\right]}$,

I_0 = (local lateral intensity) = $U_0U_0^*$, s^2 = (squared diffusion length) = $D\Big/\dfrac{\partial\Re}{\partial N}$

and I_s = (saturation intensity) = $\dfrac{(h\nu)\dfrac{\partial \Re}{\partial N}}{\Gamma \dfrac{\partial G}{\partial N}}$. We solve this equation by writing the complex field fluctuation as

$$f(z) = (g + il)\exp(\gamma z), \tag{9}$$

in which g = Re($f(z = 0)$), l = Im($f(z = 0)$) define the complex filament amplitude at z = 0 and γ is the exponential growth rate, or filament gain. We then substitute into the equation for the field fluctuations. This gives two coupled equations to determine the quantities g, l, γ :

$$\begin{aligned} \gamma g &= -Bg + \frac{k_F^2}{2k} l \\ \gamma l &= \left(-\frac{k_F^2}{2k} + \alpha B \right) g \end{aligned} \tag{10}$$

In order to avoid the trivial solution, $g=l=0$, the filament gain must satisfy the characteristic equation as

$$\gamma = \frac{-B + \sqrt{B^2 + 4\alpha B\left(k_F^2 / 2k\right) - 4\left(k_F^2 / 2k\right)}}{2} \tag{11}$$

In the limit of small carrier diffusion, we find that the maximum filament gain occurs at a spatial frequency given as

$$\left(\frac{k_F^2}{2k} \right) = \frac{\alpha B}{2} \tag{12}$$

At this spatial frequency, the gain and period of the filaments are given by

$$\gamma_{\max} = \frac{\left(\sqrt{\alpha^2 + 1} - 1\right)}{2} \; \frac{\left(I_0 / I_s \right)}{1 + \left(I_0 / I_s \right)} \; a$$

(13)

$$P = 2\pi \Big/ \sqrt{\frac{\alpha\left(I_0/I_s\right) a\,k}{1+\left(I_0/I_s\right)}}$$

in which the saturation intensity was defined above. Equations (13) quantify the connection between confinement factor Γ, and lateral beam quality. Should filaments develop, a lateral beam quality at best focus will be approximately given by the ratio of lateral pump width to filament period, *W/P*.

As the confinement factor is reduced, the necessary threshold gain is increased while the differential gain is also reduced; both of these effects lead to a significant saturation intensity increase. If the design allows for large saturation intensity so that the ratio of circulating intensity to saturation, I_0/I_s, is small, then Equations (13) predict a decreased filament gain while the period or size of any filaments that do form is increased. Improvements in lateral spatial coherence and device beam quality should follow. Results for low-confinement factor mid-IR semiconductor lasers with reduced fast-axis divergence and reduced filamentation will be shown in Section 3.

2.3.2 Ghost Modes in the MIR-OPSL

Mid-IR semiconductor lasers operate at wavelengths for which the laser substrate can be essentially transparent to the lasing wavelength. In these cases, with the bandgap of the active region less than the bandgap of the substrate, it is possible for modes in the substrate to compete with the dominant mode that is guided by the epitaxial waveguide layers. The possibility of competition becomes even more pronounced in lasers with low confinement dielectric waveguides. Although these substrate modes generally have poor overlap with the active layer and only modest gains relative to the dominant mode, they can also have small facet losses due to their high angle-of-incidence on the facets. Recent data on mid-infrared optically pumped semiconductor lasers grown on GaSb substrates have indicated competition between the dominant laser mode and the low-loss modes of the substrate, which are nearly totally internally-reflected. We refer to these low outcoupling substrate modes as *ghost modes.* Figure 6 gives calculated transverse intensity profiles for the two mode types. The arrows indicate that the dominant mode essentially propagates from facet to facet, while the ghost mode circulates throughout the structure, propagating towards the facets at a high angle of incidence.

The probability of ghost mode competition becomes more pronounced in optically pumped lasers incorporating a low-loss undoped substrate. In these devices, the totally internally reflected (TIR) ghost modes suffer essentially no facet loss, and very little free-carrier absorption in the substrate. Therefore, even with very little overlap with the gain, the possibility of reaching a threshold and competing with the main mode is present. If competition exists between the main out-coupled mode and the ghost mode, the effect is most evident in plots of differential quantum efficiency, DQE *vs* device length. Figure 7 shows such a plot for a mid-IR low confinement laser structure grown on a transparent GaSb:Te substrate. The DQE is observed to oscillate as a function of device length. This can be compared to a laser operating

with a single transverse mode for which the standard DQE *vs* length dependence is shown by the dashed line in Figure 7.

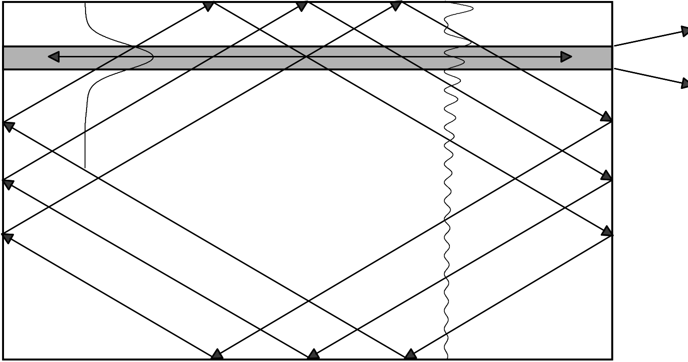


Figure **6.** Dielectric waveguide structure with a main mode and a ghost mode extending into the substrate

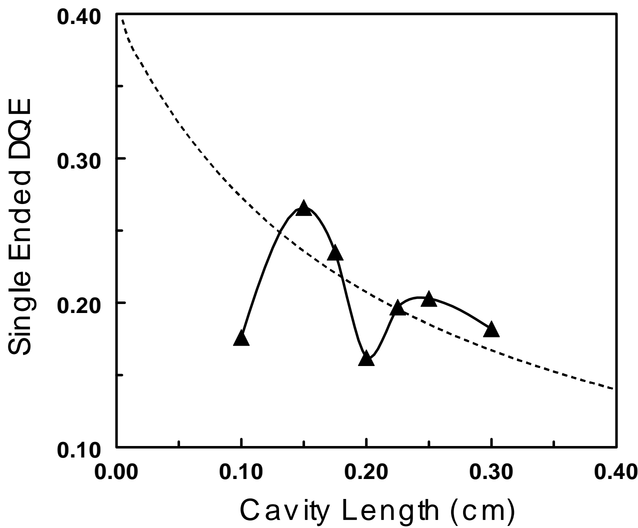


Figure 7. Differential quantum efficiency *vs* cavity length for a low-Γ laser grown on a transperant substrate. The *dashed line* shows the expected dependence for a single mode laser.

The complex DQE for a device corrupted by ghost modes can be explained in terms of power exchange between the competing modes. Whenever a ghost mode reaches threshold and competes with the dominant out-coupled mode, the modes are able to share power while saturating the gain to the point where gain equals loss. Since only the power in the out-coupled mode contributes to DQE, while the ghost power is dissipated in or scattered out of the laser, one could expect complications in DQE measurements. The quantitative details of this power-sharing depend very sensitively on the thresholds for the competing modes. For the device shown in Figure 7, the active region incorporated ten W-IA regions with a total active region thickness of 1 μm. For this design, the ghost mode and dominant mode were extremely competitive. We find that by increasing to fourteen W-IA regions with a total thickness of 1.4 μm, the overlap of the dominant mode is increased, and ghost modes are suppressed.

3 Laser Characteristics

3.1 Spectral, Power and Loss Measurements

MIR-OPSL devices require a minimum amount of post-epitaxial processing. We typically thin the substrate side and mount a rectangular device that is cleaved at the front and back facets onto a copper heatsink that is placed in a LN2 dewar. The Fabry-Perot cavity length of the device is typically between 0.5 and 4 mm.

A pump stripe of approximately ~250 μm in width is generated using an aspheric lens in front of a multi-bar stacked array emitting at a wavelength near 1.93 μm. Depending on the mounting, the gain-guided device can either be pumped from the epi-side or the substrate side.

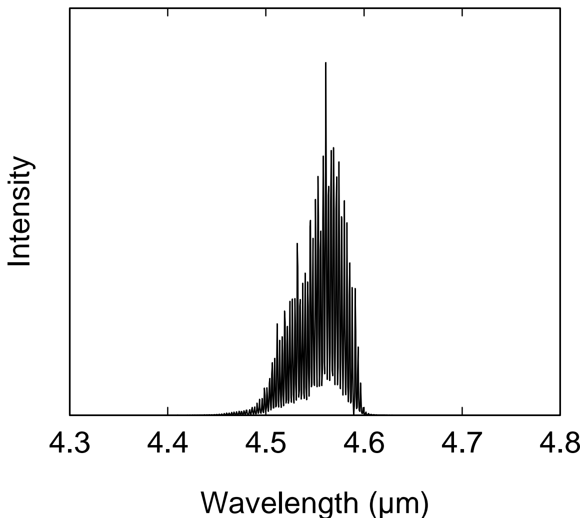


Figure 8. Emission spectrum from MIR-OPSL laser with six monolayer thick InAs electron wells

Here, we present characteristics of lasers that feature a very low confinement waveguide design, in which the clad layers are GaSb. Keeping everything else the same, and varying only the InAs layer thickness, lasers with emission wavelength spanning the ~2.4 to ~9.3 µm range are produced. Broad area lasers typically exhibit a multi-longitudinal mode spectral width of ~5meV. The emission spectrum collected at ~80K from a laser emitting at ~ 4.55 µm is shown in Figure 8, and is typical of all other devices. With nominally six monolayer thick InAs layers, this device, as well as all others, are very consistent with predictions made by SEPM.

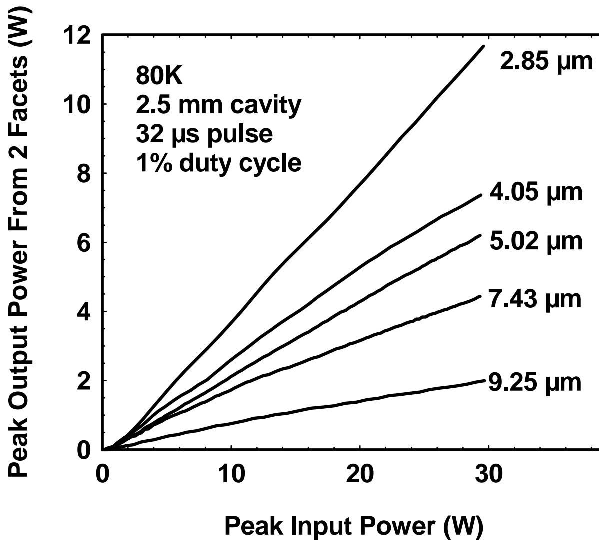


Figure 9. Power-power curves collected from a variety of MIR-OPSLs at ~80K

Power-power curves collected at ~80 K for a set of uncoated 2.5 mm long cavity devices that were epi-up mounted are shown in Figure 9. To avoid heating, 32µs long pulses at 1% duty cycle are used. Note that, in each case, the maximum peak power measured from both facets is limited only by the availability of pump power. The photon-to-photon conversion efficiency is quite high near 60% in each case, with the exception of ~40% for the longest wavelength laser. The fact that this conversion efficiency remains essentially unchanged within the large wavelength range is remarkable, and suggests that wavelength dependent loss mechanisms are not dominant in this type of active region. The threshold power densities are also similar, and range between 40 and 100 W/cm^2. In all cases, the photon-to-photon conversion efficiency drops as the operating temperature is increased. This drop in efficiency is also observed to be somewhat independent of emission wavelength as shown in Figure 10. Two lasers, at 3.6 µm and 5.1µm, are typical and show a very similar drop in the efficiency as temperature is increased.

For optimized single ended emission, both high-reflection and anti-reflection coatings are applied on the facets, and an anti-reflection coating is also applied on the top surface of the device through which it is pumped. The latter allows more

pump energy to be coupled into the device. Power output at ~80K from two such lasers are shown in Figure11, demonstrating ~11 W at ~3.6 μm and ~7 W at ~4.55 at the highest pump power available.

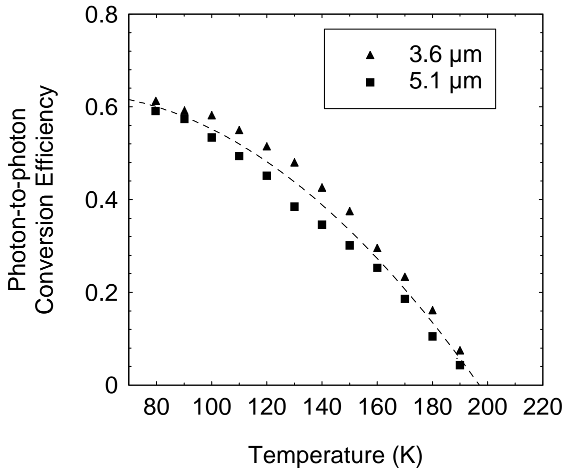


Figure 10. Conversion efficiency *vs* temperature in different MIR-OPSLs

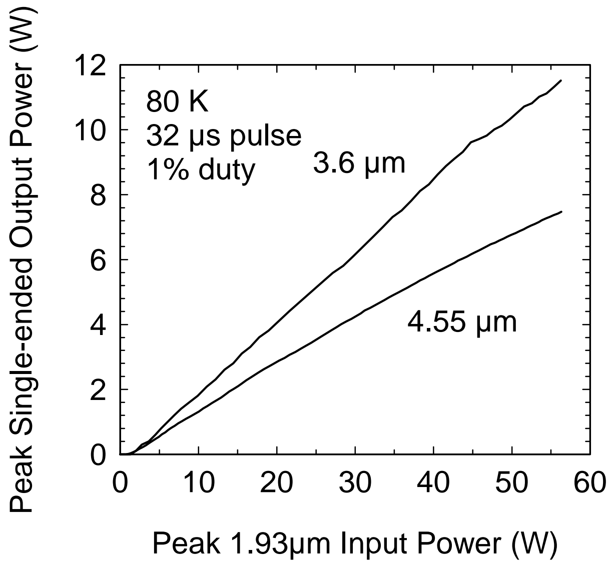


Figure **11.** Single-ended output power from coated MIR-OPSL devices

Sub-threshold Hakki-Paoli gain measurements, [25,34] as well as standard DQE *vs* cavity length measurements are used to determine waveguide loss in MIR-OPSL devices, and this is typically very modest; less than 5 cm^{-1}. In Figure 12, the temperature dependency of the waveguide loss coefficient is shown for the 3.6 μm laser. Interestingly, the waveguide loss is not observed to vary with temperature. The decrease in photon-to-photon conversion rate as a function of temperature is therefore primarily due to decreasing internal efficiency, as also shown in Figure 12.

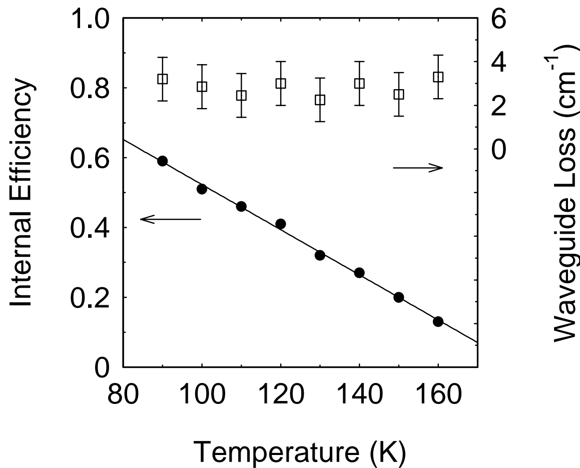


Figure **12.** Internal efficiency and waveguide loss for the 3.6 μm laser

A decrease in internal efficiency with temperature may indicate one of several loss mechanisms. It is possible, for instance, that the carrier lifetime in the absorbing regions is reduced with increasing temperature. Alternatively, holes may become less likely to be trapped in the Type-II wells at elevated temperatures due to inadequate confinement.

We believe, however, that the disparate effective masses of electrons and holes in the Type-II well alone lead to a significant reduction in gain as a function of temperature. Due to the large effective mass difference, the momentum spread of carriers at elevated temperatures results in a much-reduced rate of direct recombination. Evidence for this is given by the relatively good agreement between experimental measurements and SEPM calculations of gain as a function of temperature.

The gain at the peak wavelength, measured at a fixed input power of 865 mW using the Hakki-Paoli technique, is shown as a function of temperature in Figure 13. The gain for this nominally 3.6μm laser is measured to rapidly decrease with temperature. In comparison, SEPM calculations of gain assuming a fixed inversion (6.08×10^{12} cm^{-2}) and ignoring all other losses are also shown for the same temperature range in Figure 13. While a discrepancy between experiment and theory

exists, it can be completely removed if the assumption of a fixed inversion is relaxed to allow a 20% carrier reduction during the temperature change from 100 to 130 K. An increase in the carrier recombination rate suggested by this observation is likely due to an increase of the rates of non-radiative loss mechanisms, such as Auger and Shockley-Read recombination.

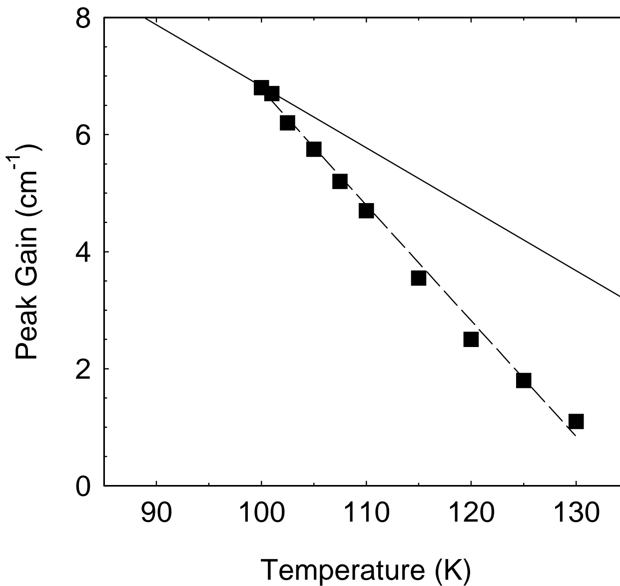


Figure **13.** Peak gain *vs* temperature for a fixed pumping power of 865 mW. *Filled squares* are experimental data; *solid line* is SEPM theory with fixed inversion; *dashed line* is SEPM theory with linearly decreasing inversion *vs* T.

3.2 Beam Quality

A distinct advantage of optical pumping is the ability to incorporate a waveguide design in which the transverse optical mode is weakly confined. This allows the transverse mode to extend far beyond the active region without paying the penalty of increased losses due to the overlap with conductive clad layers. In the case of the devices described here, the transverse mode extends well into the GaSb substrate, reaching approximately 10 μm below the surface of the epilayer. Consequently, the transverse divergence in these devices is reduced when compared to similar MIR-OPSLs designed with high optical confinement containing $Al_xGa_{1-x}As_ySb_{1-y}$ clad layers with x~0.9. This comparison is shown in Figure 14a, where the device with low optical confinement has approximately one third the divergence as the high confinement device in the transverse direction.

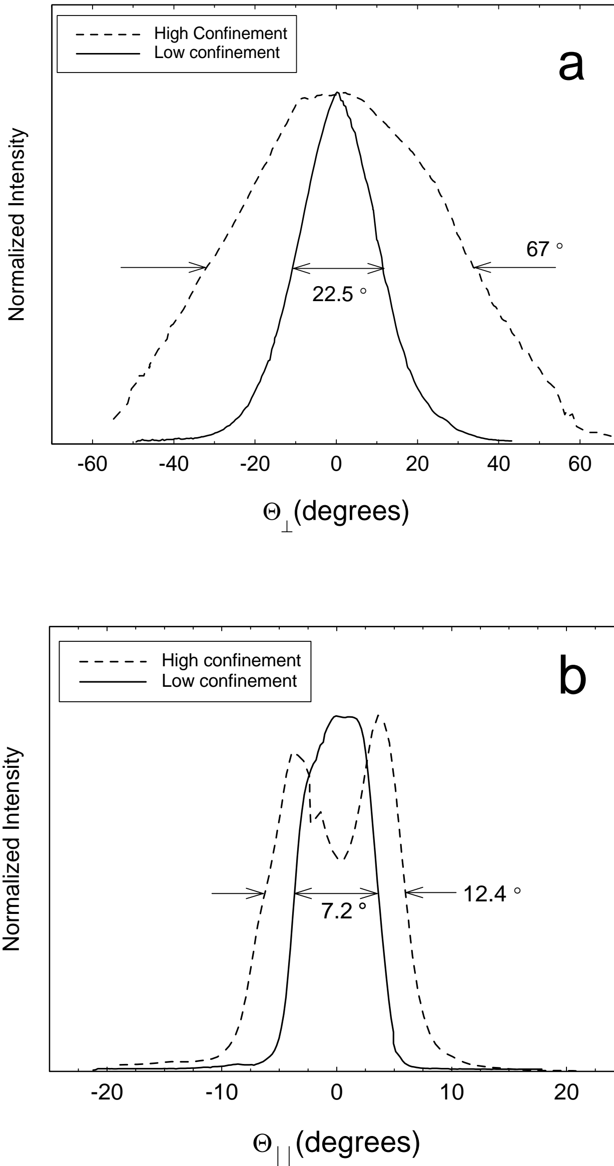


Figure **14.** (**a**) Comparison of MIR-OPSL transverse (fast-axis) divergence in devices with high and low optical confinement factor; (**b**) Comparison of lateral (slow-axis) divergence in the same devices.

For reasons discussed in Section 2, the reduced transverse optical confinement factor is also expected to reduce the tendency for lateral filament formation in broad area devices. A consequence of this is a reduced lateral divergence from these devices. In Figure 14b, the lateral divergence from low- and high-confinement devices of the

same ~250μm width is compared. A drastically reduced divergence from the low confinement device indicates that the average filament size becomes much larger, relative to the case with high confinement. Our best estimate for the filament period is in the range of 60-80 μm. As a result, the lateral beam quality is greatly improved. We estimate a three to four times diffraction limited beam in the lateral direction in the MIR-OPSL. Moreover, an improved saturation of the active region when fewer filaments are present helps to extract the high power levels observed.

4 Conclusion

In summary, the MIR-OPSL devices presented here provide a remarkable flexibility in designing for any emission wavelength in the 2.5-9.5μm ranges. A very efficient photon-to-photon conversion is maintained across this wavelength range as long as the devices are operated below ~150K. The fact that many of the device characteristics such as the threshold power and internal efficiency appear to be wavelength independent is probably due to the fact that only the InAs layer thickness is varied by a few monolayers to traverse this large wavelength range. The much-reduced transverse divergence in low confinement devices allows for a beam that is captured with much less effort in a system application. The multi-Watt power levels that are produced are accompanied by a much-improved lateral beam quality, and consequently result in very high brightness levels.

References

1. Le HQ, Turner GW, Ochoa JR, IEEE Photonics Technol. Lett. 1998; 10: 663-665.
2. Biefeld RM, Phillips JD, Kurtz SR, J. Crystal Growth 2000; 211: 400-404.
3. Kellermann K, Zimin D, Alchalabi K, Gasser P, Pikhtin NA, Zogg H, J. Applied Physics 2003; 94: 7053-7058.
4. Springholz G, Schwarzl T, Heiss W, Bauer G, Aigle M, Pascher H, Vavra I, Applied Physics Letters 2001; 79: 1225-1227.
5. Felix CL, Bewley WW, Vurgaftman I, *et al.* Appplied Optics 2001; 40: 806-811.
6. Malin JI, Felix CL, Meyer JR, *et al.* Electronics Letters 1996; 32: 1593-1595.
7. Kaspi R, Ongstad A, Moeller C, Dente GC, Chavez J, Tilton ML, Gianardi D, Applied Physics Letters 2001; 79: 302-304.
8. Felix CL, Bewley WW, Olafsen LJ, IEEE Photonics Technol. Lett. 1999; 11:964-966.
9. Le HQ, Turner GW, Ochoa JR, Electronics Letters 1996; 32: 2359-2360.
10. Bewley WW, Felix CL, Aifer EH, *et al.* Applied Physics Letters 1998; 73: 3383 -3835.
11. Bewley WW, Felix CL, Vurgaftman I, et al, Applied Physics Letters 1999; 74: 1075-1077.
12. Vurgaftman I, Bewley WW, Canedy CL, *et al.* IEE Proceedings Optoelectronics 2003;150:322-326.
13. Kurtz SR, Allerman AA, Biefeld RM, Applied Physics Letters 1997; 70: 3188 3190.
14. Le HQ, Lin CH, Pei SS, Applied Physics Letters 1998; 72: 3434-3436.
15. Bewley WW, Kim CS, Kim M, *et al.* Applied Physics Letters 2003; 83: 5383-5385.

16. Goyal AK, Turner, GW, Choi HK, Foti PJ, Manfra MJ, Fan TY, Sanchez A, LEOS 2000 IEEE Annual Meeting Conference Proceedings, 2000; 1: 249-250.
17. Goodhue WD, Le HQ, DiCecca S, J. Vacuum Science and Technol. B 1993; 11: 948-951.
18. Le HQ, Goodhue WD, DiCecca S, Applied Physics Letters 1992; 60: 1280-1282.
19. Le HQ, Turner GW, Eglash SJ, Choi HK, Coppeta DA, Applied Physics Letters 1994;64:152-154.
20. Turner GW, Choi HK, Le HQ, J. Vacuum Science and Technol B 1995; 13: 699-701.
21. Bewley WW, Felix CL, Vurgaftman I, Stokes DW, Meyer JR, Lee H, Martinelli RU, IEEE Photonics Technol. Lett. 2000; 12:477-479.
22. Felix CL, Bewley WW, Vurgaftman I, Olafsen LJ, Stokes DW, Meyer JR, Yang MJ, Applied Physics Letters 1999; 75: 2876-2878.
23. Munoz M, Wei K, Pollak FH, Freeouf JL, Wang CA, Charache GW, Journal of Applied Physics 2000; 87: 1780-1787.
24. Kaspi R, Ongstad A, Dente GC, Chavez J, Tilton ML, Gianardi D, Applied Physics Letters 2002; 81: 406-408.
25. Ongstad AP, Kaspi R, Moeller CE, Tilton ML, Chavez JR, Dente GC,Journal of Applied Physics 2004; 95: 1619-1624.
26. Meyer JR, Hofman CA, Bartoli FJ, Ram-Mohan LR, Applied Physics Letters 1995; 67: 757-759.
27. Smith DL, Mailhiot C, Reviews of Modern Physics 1990; 62: 173-234.
28. Dente GC, Tilton ML, Journal of Applied Physics 1999; 86: 1420-1429.
29. Dente GC, Tilton ML, Physical Review B 2002; 66: 165307
30. Ackley DE, IEEE Journal of Quantum Electronics 1982; 11: 1910-1917.
31. Choi HK, Walpole JN, Turner GW, Connors MK, Missaggia LJ, Manfra MJ, **:** IEEE Photonics Technol. Lett. 1998; 10: 938-940.
32. Tilton ML, Dente GC, Paxton AH, *et al.* IEEE Journal of Quantum Electronics 1991; 27: 2098-2108.
33. Dente GC, IEEE Journal of Quantum Electronics 2001; 37: 1650-1653.
34. Suchalkin S, Westerfeld D, Donetski D, *et al.* Applied Physics Letters 2002; 80: 2833-2835.

Mid-infrared Quantum Cascade Lasers

J. Cockburn
Department of Physics and Astronomy, University of Sheffield, Sheffield S3 7RH, United Kingdom

1 Introduction

It is now more than ten years since Faist, Capasso, and co-workers [1] reported the invention of the quantum cascade laser (QCL), a truly revolutionary mid-IR source exploiting electronic wavefunction engineering at an unprecedented level of sophistication. During this time, QCL device physics has developed into a highly active and flourishing field, and significant progress has been made towards realising the technological potential of these devices. The aim of this chapter is to present an overview of the basic concepts of mid-IR QCL design[1], and to provide a summary of recent developments. In particular, improvements in active region design which have provided the basis for current state of the art device performance will be reviewed, as will recent progress towards high performance short wavelength QCLs emitting in the highly technologically significant 3–5μm region. In addition, the recent emergence of metal organic vapour phase epitaxy (MOVPE) as a viable route towards relatively high volume, low cost production of QCLs will be discussed.

1.1 General QCL Concepts and Current Status

In contrast to conventional interband semiconductor lasers, the radiative transitions in a QCL occur not between electron and hole states, but between electron subbands confined within conduction band quantum wells alone (*e.g.,* E3 and E2 in Figure 1). Thus, free electrons injected into the structure are not lost in the radiative transition process, and may be recycled many times by subsequent injection into "downstream" active regions, forming the electron cascade (through typically 25-40 active periods) that gives the device its name. This scheme of operation provides significantly enhanced optical gain. The active regions (Figure 1) must, of course, be suitably designed to produce an intersubband population inversion, as discussed in Section 2. In addition, the injector region must be doped to an appropriate level to prevent the formation of high electric field domains under operating conditions, to ensure equal potential drop across each successive active period.

The intersubband nature of the lasing transition provides QCLs with a number of key advantages compared with conventional mid-IR diode lasers. First, the dominant non-radiative intersubband process is electron-longitudinal optic (LO) phonon scattering. This is relatively insensitive to temperature compared with the non-radiative Auger recombination processes which dominate in mid-IR interband lasers.

---

[1] For a more detailed treatment of the physics of QCL operation see Faist *et al.* [2] or Gmachl *et al.* [3]

Thus, the high temperature performance of QCLs is significantly superior at long wavelengths, with high power, room temperature continuous wave (CW) operation[2] being possible in the 5-10μm range with optimised device design and fabrication [5, 6, 7]. The best of these devices have threshold current densities (J_{th}) of less than $2kAcm^{-2}$ at 300K, and produce average optical powers of several hundred milliwatts at room temperature.

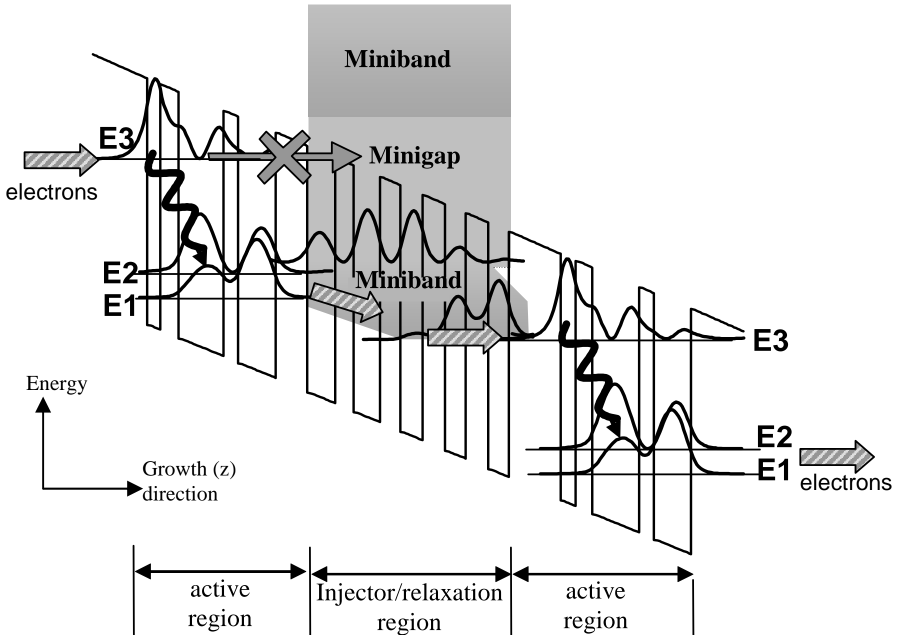


Figure 1. Schematic illustration of two active periods of a typical QCL, linked by a superlattice injector/relaxation region. Electrons are injected into the upper laser level (E3), undergo the laser transition indicated by the *wavy arrow*, and follow the path indicated by the striped arrows to the next active period. The *crossed arrow* represents the inhibition of electron tunnelling out of E3 by the minigap of states in the injector/relaxation region [4].

A further benefit of the QCL concept is that the emission wavelength can be varied over a very wide range within a single materials system, by making appropriate changes to the widths of the quantum well and barrier layers in the active regions. Thus, wavelengths ranging from ~3.5μm [8] to 24μm [9] have been achieved for QCLs based on InGaAs/AlInAs/InP, while an even wider range of 7.3μm [10] to beyond 100μm [11] has been demonstrated for GaAs/AlGaAs QCLs. Such wavelength versatility is impossible with conventional diode lasers, where the emission energy is intimately linked to the bandgap of the material. It should be

[2] Unless explicitly stated otherwise, all descriptions of laser performance in this chapter should be taken to refer to pulsed, rather than CW, operation.

noted, however, that QCL performance in the $\lambda \approx 3$ to 5μm region is generally inferior to that displayed by QCLs operating at longer mid-IR wavelengths. The quest for higher performance at short wavelengths is a major area of activity in current QCL research, and is discussed in detail in Section 3.

QCLs also have the unique feature that the joint density of states for the lasing transition is quasi-atomic. Thus, in contrast to a conventional interband semiconductor laser, where absorption of above-bandgap radiation occurs into a continuum of states, absorption in a QCL active region is only possible for a small range of photon energies close to the intersubband separation (assuming negligible subband non-parabolicity). This contrasting behaviour is illustrated in Figure 2. This has the important consequence that lasing can be arranged to occur at two or more widely spaced wavelengths within a single QCL cavity. This is not possible in a conventional semiconductor laser, where the shorter wavelength radiation is always absorbed into the above bandgap continuum. Such a mode of operation has potential technological applications in differential absorption techniques for trace gas sensing, *etc.*

The first demonstration of dual wavelength QCL operation was reported by Faist *et al.* [12]. In this work, a QCL with a strongly diagonal laser transition was used so that the transition energy displayed a significant Stark shift under bias. By employing a split contact, different current densities were injected into two sections of the laser ridge of unequal length. Since the short section required higher current density to reach the lasing threshold, the transitions in each section experienced different Stark shifts, enabling laser action to be observed at two separate wavelengths (λ=6.2 and 7μm). Further work showed that multi-wavelength

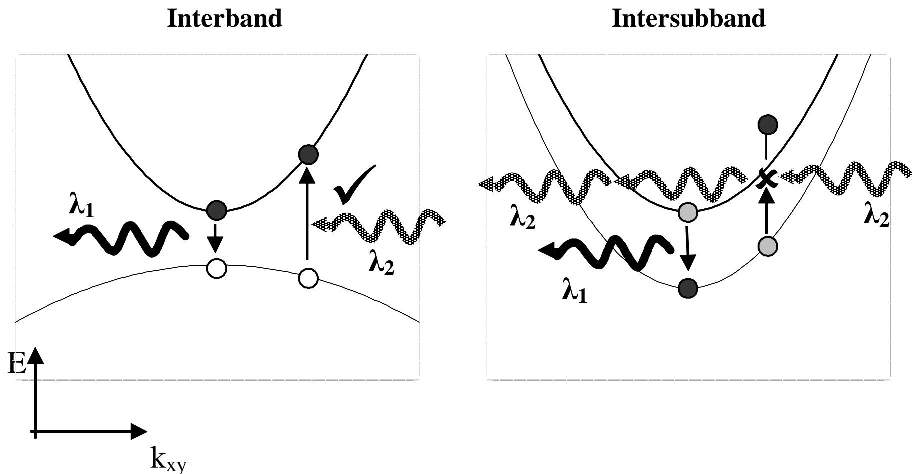


Figure 2. Schematic electron energy *vs* in-plane wave vector diagrams illustrating the possibility of multi-wavelength operation in an intersubband device. In contrast to the conventional interband case, the intersubband gain region emitting at wavelength λ_1 is transparent to radiation at the shorter wavelength λ_2, due to the parallel nature of the subbands in k-space.

operation could also be achieved by transitions between discrete levels within quasi-minibands of states [13], and by employing active regions with separate substacks designed to emit at two different wavelengths [14]. The latter concept was further developed when Gmachl *et al.* [15] demonstrated QCLs in which the active quantum well widths were varied throughout the structure to give an almost continuous range of transition energies and a broad gain spectrum. The result was an ultra-broadband intersubband laser, emitting simultaneously over a range of wavelengths from 6 to 8μm.

An interesting corollary of multi-wavelength operation is the potential for exploiting the very large non-linear susceptibilities of intersubband transitions [16] in order to produce intracavity non-linear optical effects such as second harmonic and sum/difference frequency generation. Such effects have been demonstrated in a dual wavelength QCL [17] consisting of λ=7.1μm and λ=9.5μm active regions separated by a "mixing" superlattice region. In addition to emission at the fundamental lasing wavelengths, this structure displayed efficient generation of sum frequency ($\lambda \approx$4.1μm) and second harmonic radiation arising from both fundamental emission wavelengths. In this case, the optical non-linearities were provided by resonant intersubband transitions within the laser active regions themselves (from the upper laser level to higher-lying states) as well as within the mixing superlattice. Second harmonic generation has also been observed in GaAs-based QCLs grown on ‹111› oriented substrates, exploiting the bulk optical non-linearity of the substrate material [18]. In general, this is an area of significant potential, not only for novel technological applications but also for fundamental physical investigations.

2 QCL Active Region Design

Over the past few years significant advances have been made in both the performance and functionality of quantum cascade lasers, made possible to a large extent by continual improvements to active region design, combined with development of devices in new materials systems (Section 3). In the present section, the fundamental design principles of mid–IR QCL gain regions will be discussed together with a brief review of the most successful examples produced to date.

The material gain coefficient for the active core of a QCL designed for laser emission between an upper subband i and a lower subband f is proportional to both the square of the intersubband dipole matrix element $\langle\psi_f|z|\psi_i\rangle$ and the intersubband population inversion (n_i-n_f). Here ψ_i and ψ_f are the z-direction envelope functions and n_i and n_f the electronic sheet densities for the upper and lower subbands respectively. As QCL design has evolved, the dipole matrix element has been maximised by moving towards active regions with more spatially direct (vertical) intersubband transitions, and by the utilisation of materials systems with low electron effective mass in the quantum wells (see Section 2.5). In this section we will concentrate on aspects of active region design that control the intersubband population inversion.

2.1 Intersubband Population Inversion

In QCLs, the radiative efficiency for spontaneous intersubband emission is very low, and the sub-threshold carrier dynamics are dominated by non-radiative intersubband scattering. For mid–IR wavelengths, the intersubband separation is well above the longitudinal optic (LO) phonon energy in the III-V materials used ($\hbar\omega_{LO}$ ~30meV), and the lifetime of an electron in a particular subband is therefore dominated by LO phonon scattering. This typically occurs on a picosecond timescale, at least 3-4 orders of magnitude faster than the rate for spontaneous radiative transitions. The LO-phonon mediated intersubband scattering rate may be calculated using Fermi's Golden Rule:

$$\frac{1}{\tau_{if}} = \frac{2\pi}{\hbar}\left|\langle\psi_i|V(r)|\psi_f\rangle\right|^2 \delta(\varepsilon_f + \hbar\omega_{LO} - \varepsilon_i) \quad (1)$$

For details of such calculations, the reader is referred to Ferreira and Bastard [19] for example, but for present purposes it is useful to consider qualitatively the form of the perturbing potential V(r) applicable to the electron-LO phonon interaction. In III-V polar crystals a strong interaction arises due to the polarisation field produced by the relative displacement *r* of the group III cations and group V anions moving in antiphase in the LO phonon mode. Since this field *E* is proportional to $e^{i(q.r-\omega t)}$,

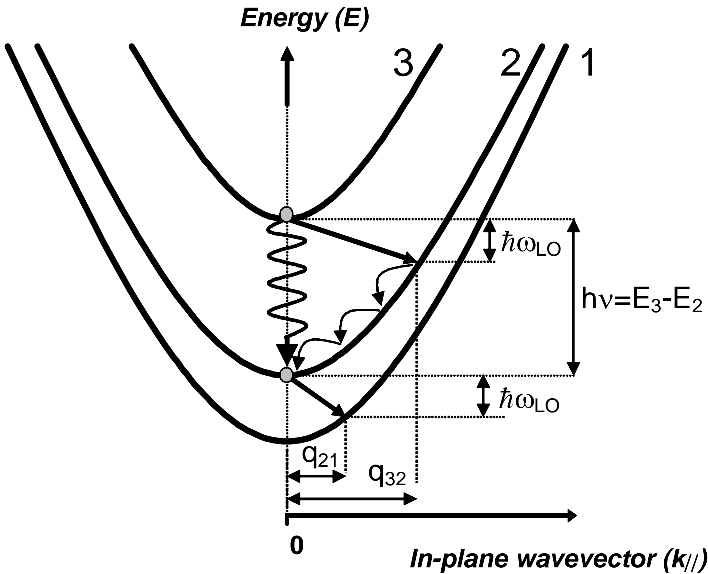


Figure 3. Mechanism for intersubband population inversion in a quantum well. Because (E_3-E_2) >(E_2-E_1), electron scattering from 3 to 2 requires the participation of an LO-phonon of higher in plane wavevector (q_{32}) than is needed for scattering from 2 to 1 (q_{21}). Since $\tau \propto q^2$, $\tau_{32} > \tau_{21}$.

where q and ω are the phonon wavevector and frequency respectively, the scattering potential $V(r) = \int E dr$ is proportional to q^{-1}. Thus, from Equation (1), the LO-phonon mediated intersubband lifetime τ_{if} is expected to be proportional to the square of the in-plane wavevector q_{if} of the phonon involved in the scattering process.

As a result of this phonon wavevector dependence of the electron-LO phonon interaction, the conditions for intersubband population inversion may be automatically fulfilled in any single quantum well containing three or more quasi-confined subbands, where the intersubband separation is greater than $\hbar\omega_{LO}$ [20]. To illustrate this, Figure 3 shows the energy *vs* in plane momentum ($k_{//}$) dispersion relations for a single quantum well in which electrons are injected into the n=3 (E3) quasi-confined level. These injected electrons can rapidly relax within the E3 subband by *intra*subband LO phonon scattering (τ ~100fs), before transferring to the E2 subband, again with the emission of LO phonons. These electrons subsequently undergo LO-phonon mediated intersubband scattering to E1. Since the LO phonon energy is approximately constant, and E3-E2 is intrinsically greater than the E2-E1 energy separation, E3-E2 scattering must involve phonons of higher wavevector (q) than is the case for E2-E1 scattering (Figure 3). Since $q_{32} > q_{21}$, we have that $\tau_{32} > \tau_{21}$, giving the conditions necessary for intersubband population inversion.

The above model is, of course, only valid if electrons are removed sufficiently quickly from E1 (*e.g.* by tunnelling out of the well or scattering to a lower energy state) to prevent inhibition of intersubband scattering due to band filling effects. In addition to this high extraction efficiency (η_{ex}), the QCL active region must have high injection efficiency (η_{in}) into the upper laser level, in order to maximise intersubband population inversion for a given drive current. This means providing optimum conditions for resonant coupling between the injector states and the upper laser level and minimising non-resonant tunnelling to lower levels. As discussed below, it is largely the quest for improvements in η_{ex} and η_{in} that has driven development of QCL active region design in recent years. Another important factor to consider for the optimisation of population inversion is the prevention of thermally-activated escape of electrons from the upper laser level into above-barrier continuum states. This is discussed in more detail in Sections 2.5 and 3.

2.2 Two-well and Three-well Active Regions

Whilst it is relatively straightforward to produce an intersubband population inversion as described above, it is vital to maximise this population inversion for a given injected current in order to ensure sufficient optical gain for lasing. This usually involves exploiting designs based on coupled quantum well structures in which the lowest energy E1 and E2 states have a symmetric/antisymmetric energy splitting close to $\hbar\omega_{LO}$ in order to maximise the E2-E1 intersubband scattering rate by resonant (q_{21}~0) LO phonon emission. Such a design, based on a GaAs/AlGaAs well/barrier system, is illustrated in Figure 4. In this structure, which is designed for emission at $\lambda \approx 11\mu m$ (E3-E2≈112meV), the LO phonon assisted E3-E2 intersubband scattering time τ_{32} is calculated [19] to be ~2ps, while for E2-E1 scattering (E2-

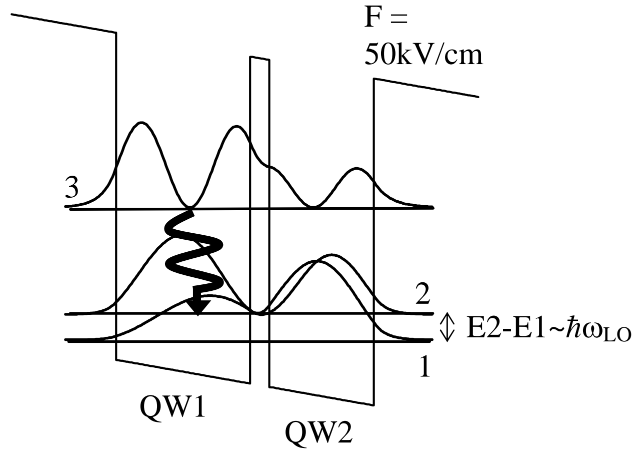


Figure 4. Two-well QCL active region design. The spacing of the anticrossed E1 and E2 levels is chosen to be close to resonance with the LO-phonon energy in the material in order to maximise the E2-E1 scattering rate and hence the E3-E2 intersubband population inversion [4]

E1$\approx\hbar\omega_{LO}$), τ_{21}~0.3ps. It should be noted, however that the close spacing of the E2 and E1 levels means that in addition to E3-E2 scattering, a significant fraction of E3 electrons will scatter directly to E1 by LO phonon emission (since $q_{32}\approx q_{31}$), reducing the overall lifetime of E3 electrons by a factor of ~2. Nevertheless, a significant E3-E2 intersubband population inversion is ensured for such a structure, provided that (a) efficient electron injection into E3 and extraction out of E1 is maintained, and (b) parasitic loss of carriers from E3 (by, for example, tunnelling into the above barrier continuum) is minimised. The latter requirement is addressed in all QCL active region designs by the use of an appropriate materials system with sufficient quantum well depth to provide adequate confinement for the upper laser level electrons in conjunction with a suitably designed superlattice "bridging" region between successive active periods. This produces a minigap in the density of states which blocks tunnelling escape (Figure 1).

The earliest high performance QCL designs were based on the two-well active region described above [21]. Subsequently, however, it was shown that significant improvements in device performance could be achieved by adding an extra, thin quantum well to the active region in order to selectively enhance coupling between the ground state of the injector region and the upper laser level, thereby increasing the injection efficiency [22] and minimising non-resonant transfer of electrons to the lower energy states of the active region. An example of such a three-well active region design is shown in Figure 5.

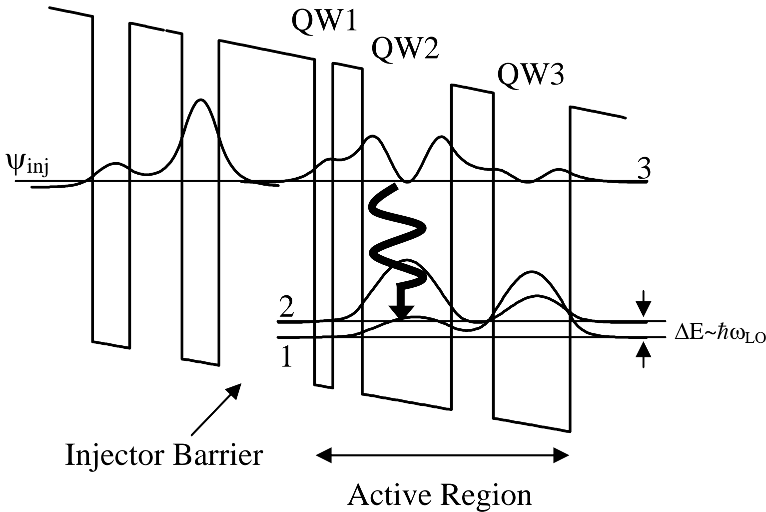


Figure 5. Three-well active region. The addition of the thin well QW1 promotes resonant coupling of the upper laser level with the injector ground state (ψ_{inj}), increasing the injection efficiency compared with the two-well design

2.3 Double Phonon Resonance Active Region

The double phonon resonance design, developed by Hofstetter *et al.* [23] is very similar in concept to the three-well active region discussed above. In these structures, however, an extra quantum well is added to the active region to produce an additional anticrossed level below the lower lasing state. As in the three-well design, well and barrier widths are chosen so that the splitting between these levels is close to the LO phonon energy, ensuring very rapid scattering from the lower laser level (Figure 6). These designs typically display a reduction in the lower state lifetime of about 20% compared with the three-well design, whilst providing essentially identical injection efficiency and optical matrix element for the intersubband transition. A significant increase in population inversion and optical gain is therefore achieved in this design, which has been the basis for some of the highest performance QCLs yet developed [6,24].

2.4 Bound to Continuum Active Region

Whilst the double phonon resonance design is highly effective in reducing the lower laser level lifetime, the ultimate bottleneck is the efficiency with which electrons can be removed from the lowest energy states of the active region into the injection/relaxation region. In both the three-well and double phonon devices this extraction is by resonant tunnelling, with a characteristic lifetime of the order of 1ps [25]. This is much longer then the resonant LO-phonon assisted scattering lifetime between the low energy states in the active region which is of the order of 0.1ps.

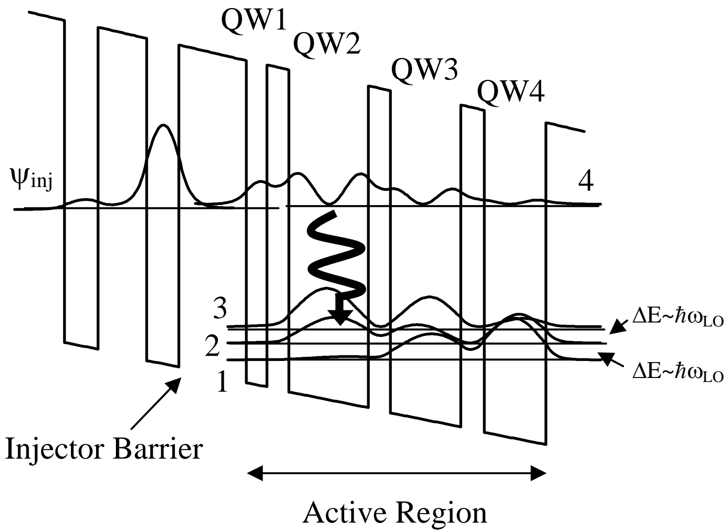


Figure 6. Double phonon resonance design. The addition of QW4 to the three-well design provides an additional anticrossed state for resonant LO-phonon assisted intersubband scattering from the lower laser level (3), increasing the extraction efficiency.

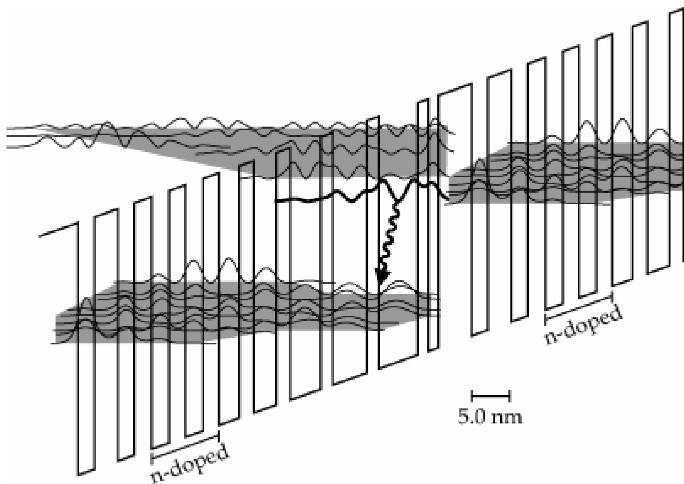


Figure 7. Bound to continuum active region. The design combines the high injection efficiency of the three-well and double phonon resonance designs with the high extraction efficiency of an interminiband superlattice laser. Diagram courtesy of Stéphane Blaser, Alpes Lasers SA. Reprinted with permission from Stéphane Blaser, Dmitri A. Yarekha, Lubos Hvozdara, Yargo Bonetti, Antoine Muller, Marcella Giovannini, and Jérôme Faist, Applied Physics Letters, 86, 041109 (2005). Copyright 2005, American Institute of Physics.

This leads to significant "backfilling" and lower state blocking effects which act to decrease population inversion, particularly at high temperatures and/or drive currents. In order to circumvent this problem, the bound to continuum (BC) design was developed by Faist and co-workers [26]. The BC design can be regarded as a hybrid of the three-well design and the interminiband superlattice QCL designs developed by Scarmarcio *et al.* [27] and Tredicucci *et al.* [28]. A schematic diagram of a BC active region is shown in Figure 7. As in the three-well active region, and in contrast to interminiband superlattice QCLs, injection into the upper laser level occurs via resonant tunnelling from the injector region, providing significantly enhanced injection efficiency compared with interminiband designs. However, the optical transition occurs in this case between this discrete upper level and a tilted lower miniband of states formed within the chirped superlattice active/bridging region. Hence, electron transport from the lower state proceeds via highly efficient intraminiband scattering on a sub-picosecond timescale, minimising the bottleneck effects described above. The BC design has recently been used by Blaser *et al.* [7] to demonstrate very high single mode performance from a $\lambda \approx 5.4\mu m$ device, with CW emission being observed up to a temperature of 30°C.

Because of the quasi-continuum nature of the lower states, the BC design has a homogeneously broadened gain spectrum that can be used as the basis for QCLs that are tunable over a relatively wide mid–IR wavelength range, by grating coupling in an external cavity configuration [29]. As shown in Figure 8, this allows tuning of the laser emission over a range of around $150 cm^{-1}$, which is highly advantageous for sensing of complex hydrocarbons and gas mixtures.

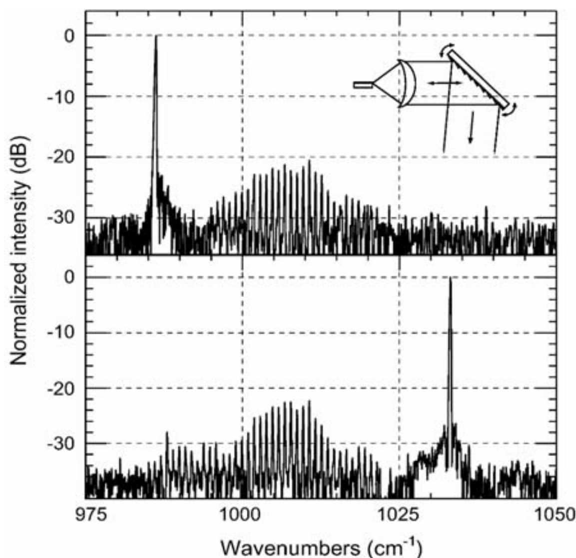


Figure 8. Tuning range of 150 wavenumbers ($\Delta\lambda \approx 9.11\mu m$-$10.56\mu m$), achievable with bound to continuum QCL in grating-coupled external cavity configuration, by variation of grating angle. Reprinted with permission from Richard Maulini, Mattias Beck, Jérôme Faist, and Emilio Gini, Applied Physics Letters, 84, 1659 (2004). Copyright 2004, American Institute of Physics.

2.5 Comparison of InP-based and GaAs-based QCLs

QCLs were originally developed in the $In_{0.53}Ga_{0.47}As/Al_{0.48}In_{0.52}As$ heterostructure system, lattice matched to InP, and from the outset there was considerable interest in the question of whether QCL operation could also be realised in the GaAs/AlGaAs materials system. Aside from natural curiosity, motivations for this included the greater technological maturity of GaAs, increased suitability for advanced device processing by dry etching techniques, and the excellent $GaAs/Al_xGa_{1-x}As$ lattice match for all values of x, which simplifies growth somewhat and provides an extra degree of flexibility in device design.

There are, however, several properties of the GaAs/AlGaAs system that render it less promising for QCL development than InGaAs/AlInAs/InP. For instance, the maximum Γ-point conduction band offset (ΔE_c) obtainable with direct bandgap $Al_xGa_{1-x}As$ barriers is about 390meV ($x = 0.45$), whereas InP-based QCLs have ΔE_c ranging from 520meV to ~740meV, depending on whether lattice-matched or strain compensated InGaAs/AlInAs is used. This means that electrons confined in the upper laser level of a GaAs-based QCL have greater probability of leakage into the above-barrier continuum at high temperatures and electric fields, leading to decreased intersubband population inversion. The Γ-point conduction band offset can be increased by using barrier aluminium fractions beyond $x=0.45$, but problems then arise due to the indirect bandgap nature of the barrier, which can provide carrier leakage paths via lateral X-point states. These barrier height issues are discussed in more detail in Section 3. Another disadvantage for GaAs-based QCLs is the effective mass of electrons in the quantum wells ($m^*=0.067m_0$), which is appreciably higher than the effective mass in $In_{0.53}Ga_{0.47}As$ ($m^*=0.043m_0$). This decreases both the intersubband dipole matrix element and the LO-phonon assisted intersubband scattering lifetime, both of which have a detrimental effect on optical gain. The key factor here is that for a given intersubband transition energy (*i.e.* laser emission wavelength), a greater well width (L) is required for a lower effective mass, and both the intersubband dipole and the form factor for intersubband scattering by the z-component of the LO phonon wavevector are proportional to L [30]. In combination, these two effects lead to the material gain coefficient scaling as $(m^*)^{-3/2}$.

In addition to superior electronic properties, the InGaAs/AlInAs/InP system provides advantages for optical waveguide design and thermal management compared with GaAs/AlGaAs. The InP substrate has lower refractive index than the core region and therefore naturally forms a dielectric waveguide cladding layer, with either AlInAs or InP being deposited on top of the active core to complete the waveguide structure. This makes the implementation of low loss, high confinement waveguides relatively easy to achieve. By contrast, in GaAs-based QCLs, optical confinement must be provided by either high aluminium fraction AlGaAs, or by heavily doped GaAs, exploiting the rapid decrease in refractive index close to the plasma resonance [31]. In general, $Al_xGa_{1-x}As$ has relatively poor electrical and thermal characteristics, while the doped GaAs "plasmon enhanced" claddings suffer from relatively high free-carrier losses and strongly wavelength-dependent confinement. Hence, both optical loss (α) and confinement factor of the optical mode (Γ) are considerably better for InP-based devices, which for $\lambda \approx 9\mu m$ have $\alpha \approx 9 cm^{-1}$ and $\Gamma \approx 40\%$, compared with $\alpha \approx 14 cm^{-1}$ and $\Gamma \approx 28\%$ for GaAs-based

waveguides. An alternative approach to waveguide design in GaAs-based QCLs was demonstrated by Green *et al.* [32] who reported room temperature operation of a QCL with $Ga_{0.51}In_{0.49}P$ cladding layers. $Ga_{0.51}In_{0.49}P$, which is lattice matched to GaAs, provides dielectric optical confinement and superior electrical characteristics to high aluminium fraction AlGaAs. Current flow through the structure is aided by the fact that most of the GaAs-GaInP bandgap offset lies in the valence band, giving a conduction band offset of only ~100meV. Calculations predict optical losses and confinement factors of $11cm^{-1}$ and 54% respectively, although the performance of the QCLs in reference [32] showed little improvement in performance compared with similar devices with GaAs plasmon waveguides. A further advantage of InP-based QCLs is that the thermal conductivity of InP (~$0.7Wcm^{-1}K^{-1}$) is considerably higher than that of GaAs (~$0.5Wcm^{-1}K^{-1}$). This is of key significance, due to the thermal management issues arising from the high electrical power densities required for QCL operation, which are at least $20kWcm^{-2}$ even in “state of the art” devices [6].

GaAs/AlGaAs QCLs ($\lambda \approx 9.5\mu m$) were first demonstrated by Sirtori *et al.* [33] and, consistent with the arguments presented above, the performance of these devices was markedly inferior to that of their InP-based counterparts [34]. Significant improvements were achieved by implementing two key design modifications: replacement of $Al_{0.9}Ga_{0.1}As$ dielectric waveguide cladding layers with plasmon enhanced waveguides based on doped GaAs [35] and increasing the aluminium fraction in the barriers [36]. These design strategies allowed the

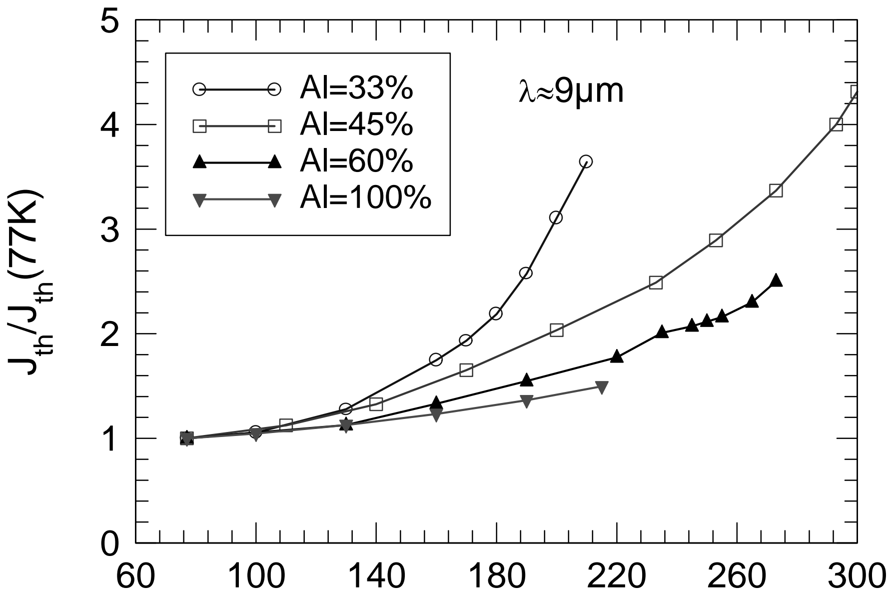


Figure 9. Variation of temperature dependence of threshold current density for a series of GaAs/Al(Ga)As QCLs with increasing aluminium fractions in the barriers. Figure courtesy of Hideaki Page, Thales LCR.

maximum temperature for laser emission to be increased from 200K to 300K, with other parameters such as low temperature threshold current density and maximum peak emission power remaining essentially unchanged. More recently, pulsed laser emission at $\lambda \approx 11\mu m$ has been observed at temperatures up to 400K in structures based on a bound-to-continuum design with $Al_{0.45}Ga_{0.55}As$ barriers [37]. These devices provided over 300mW of peak optical power at room temperature. Although high threshold current densities were observed at low temperature ($J_{th} \sim 8kAcm^{-2}$) a very high characteristic temperature T_0 (defined in the usual way by the empirical relationship $J_{th}(T) = J_0 e^{T/T_0}$) of almost 200K enabled good performance to be achieved at high temperatures.

The influence of barrier composition on device performance has been clearly demonstrated by Sirtori, Page and co-workers [38] in experiments on devices having similar active region designs but different $Al_xGa_{1-x}As$ compositions ranging from x = 0.33 (direct bandgap) to x=1 (indirect bandgap). The temperature dependence of the threshold current densities of these devices is shown in Figure 9. The data clearly show that J_{th} becomes significantly less dependent on T with increasing x, although the best overall performance (highest operating temperature) is achieved with x=0.45. The AlAs barrier devices showed a remarkable insensitivity of J_{th} to changing temperature, although the maximum operating temperature was limited by the onset of NDR in the electrical characteristics of the devices [39]. Room temperature operation of a GaAs/AlAs QCL was achieved by Carder *et al.* [40] for an emission wavelength of 8.5µm, by selective incorporation of InAs monolayers in the active region. This is the shortest wavelength yet achieved for room temperature operation of a GaAs-based QCL.

The aforementioned shortcomings of GaAs/AlGaAs as the basis for mid–IR QCL design make it all the more important to optimise device fabrication and packaging in order to extract maximum performance from the devices. This has been very well illustrated in recent work by Page *et al.* [41], who have followed this approach to demonstrate CW operation of GaAs-based QCLs for temperatures up to 150K. Although CW operation had been demonstrated previously [42] for a GaAs/AlAs superlattice QCL, the maximum CW operating temperature was only 30K. Page *et al.* used the "standard" GaAs/$Al_{0.45}Ga_{0.55}As$, design described in [36] and followed a number of key fabrication steps to enable CW operation to be observed. These included the adoption of narrow ridges, with ~7µm wide current apertures defined by proton implantation to provide separate optical and electrical confinement (minimising optical scattering losses at the ridge walls) and epilayer-down mounting on oxygen-free high-conductivity copper submounts. In this way, threshold currents were reduced to less than 500mA at liquid nitrogen temperatures, with CW optical power of ~80mW.

To summarise this sub-section, significant progress has been made in the development of mid-IR GaAs-based QCLs in recent years, to the extent that the performance of these devices is now perfectly adequate for a number of technological applications. It should also be mentioned that although outside the scope of this review, GaAs/AlGaAs is the materials system of choice for the rapidly developing area of THz QCLs [43]. In the mid-IR, however the performance of InP-based QCLs is significantly better in almost all respects, and these seem likely to dominate optoelectronic applications for the foreseeable future. Comparisons of

some key operational parameters for GaAs and InP-based QCLs are summarised in Table 1.

Table 1. Comparison of some state of the art characteristics for GaAs- and InP-based QCLs. The low temperature J_{th} data for InP-based QCLs are estimated from extrapolation of room temperature measurements, as low temperature data are often no longer quoted for high performance InP devices.

	J_{th} 77K ($kAcm^{-2}$)		J_{th} 300K ($kAcm^{-2}$)		Average power	Minimum λ, 300K (pulsed)
	Pulsed	CW	Pulsed	CW		
GaAs-based	3 [39]	3.6 [39]	15 [36] 7 [a]	-	85mW; T=77K [39] 11mW; T=240K [37]	8.5μm [39]
InP-based	<1 [6]	<1.5 [6]	1.1 [6]	1.75 [6]	600mW; T=300K [6]	3.8μm [53]

[a]Unpublished data, H. Page *et al.*, Thales LCR

3 Quantum Cascade Lasers for 3–5μm Operation

As discussed in previous sections, spectacular levels of performance have been demonstrated for QCLs operating between 5 and 12μm. Many important technological applications such as gas detection, diagnostic breath analysis and military countermeasures would benefit greatly from the availability of such high performance sources in the 3–5μm atmospheric window region. The fundamental requirement here is that the quantum wells within the active region of the device should be deep enough not only to accommodate the laser transition energy of interest, but also to inhibit escape of electrons from the upper lasing level to above-barrier continuum states, which can be particularly detrimental to intersubband population inversion at high temperatures and/or electric fields. The latter problem is exacerbated by the fact that a potential of approximately $\frac{(hc/\lambda)+\hbar\omega_{LO}}{e}$ must be dropped across each active period of the structure under operating conditions. This, combined with the fact that narrower active regions are required to produce the higher energy transitions of interest, means that significantly higher electric fields

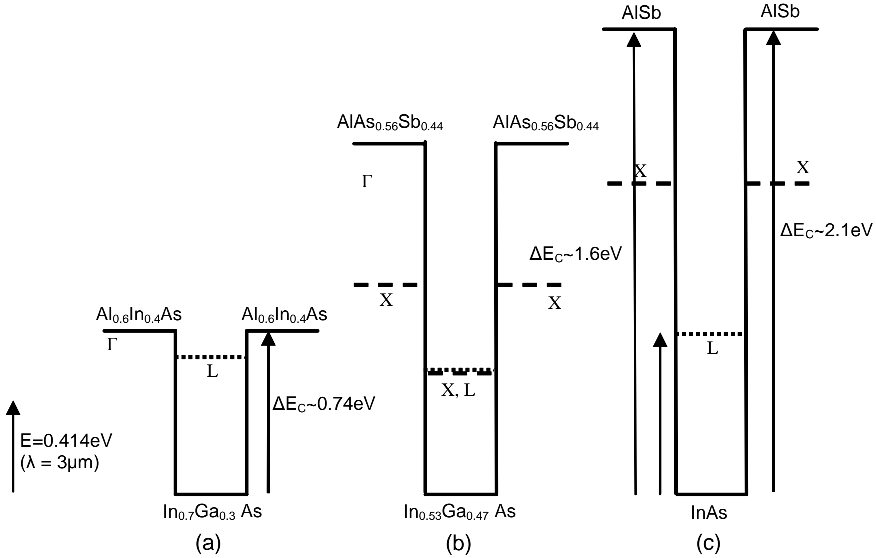


Figure 10. Some heterostructure systems of interest for short-wavelength QCLs: (**a**) Strain compensated $In_{0.7}Ga_{0.3}As/Al_{0.6}In_{0.4}As$ (InP substrate), (**b**) $AlAs_{0.56}Sb_{0.44}/In_{0.53}Ga_{0.47}As$ (InP substrate, lattice matched), (**c**) AlSb/InAs (InAs or GaSb substrate). Estimated energies and nature of the lowest satellite conduction band minima in the quantum wells are labelled, as are the X-point minima in the indirect bandgap barriers in (b) and (c). Note that for $In_{0.53}Ga_{0.47}As$, the energies of the X and L satellite minima are very close in energy. Calculations of satellite valley minima are based on parameters taken from [50].

are required for operation at short wavelengths. For example, an InP-based QCL of the bound to continuum design and an emission wavelength of around 9μm requires an operating electric field of around $35kVcm^{-1}$ [26], compared with the field of ~$67kVcm^{-1}$ required for a similar design emitting at $\lambda \approx 5.4\mu m$ [7]. These high fields produce increasingly triangular tunnelling barriers, significantly enhancing the escape probability of high energy electrons to above-barrier states and placing further demands on barrier height in short wavelength QCLs. Further difficulties can arise with high electric fields in QCLs based on very narrow bandgap materials such as InAs, due to generation of holes by interband tunnelling and/or impact ionization [44]. The narrow quantum wells required for short wavelength operation also place significantly greater demands on crystal growth, since fluctuations and growth mis-calibrations on a monolayer scale have a proportionally greater effect on confinement energies than is the case for wider well devices. In addition, interface scattering has a greater effect in narrow well devices.

From inspection of the literature, a rough guideline emerges that a conduction band offset (ΔE_c) of around twice the laser transition energy is required for room temperature pulsed operation of a QCL at a particular wavelength, with higher multiples desirable for maximum average power, efficiency and operating temperature. This requirement effectively rules out GaAs-based QCLs for short wavelength operation. For $Al_xGa_{1-x}As$ barriers with direct bandgap ($x \leq 0.45$), ΔE_c

has a maximum value of about 390meV, leading to an expected short wavelength limit in the region of 7μm. In fact, the shortest wavelength observed for room temperature pulsed operation of a GaAs/$Al_{0.45}Ga_{0.65}As$ QCL is ~9μm [45], although shorter wavelengths (λ ~ 7.4μm) have been obtained at lower temperatures by selective incorporation of ultrathin InAs layers in the active regions of the device [46]. Although higher aluminium concentrations in the barrier provide a larger ΔE_c at the Γ-point of the Brillouin zone, the indirect nature of the barrier bandgap results in an increasing importance of states confined at the X-point of the barrier [10,47,48] as carrier leakage paths, leading to reduced intersubband population inversion and hence degraded performance. Thus, for example, the shortest wavelength observed for room temperature pulsed operation of a GaAs/AlAs QCL is 8.5μm [40], despite the very large ΔE_c of ~ 1eV at the Γ-point (the Γ-X well-barrier conduction band offset (ΔE_c(Γ-X)) is only ~130meV for GaAs/AlAs [47,49]).

In general, when considering materials systems for QCL design, the possible effects of conduction band satellite valleys on intersubband population dynamics should not be underestimated. In addition to the indirect bandgap barrier states discussed above, scattering of electrons from the upper laser levels to the satellite valleys of the active quantum wells themselves must be considered. Even for devices where performance is not limited by barrier height, the energy of the lowest satellite valley minimum in the well may set an approximate lower limit on the emission wavelength possible within a given heterostructure. Unfortunately, the energies of the well and barrier satellite valleys are often not well-known for the materials systems of interest, particularly when localized strain must be taken into account, or when the conduction band line-up is not precisely understood. This highlights the need for further basic materials research to determine these parameters in short wavelength QCL materials, or at least some degree of empirical optimisation in device development. Estimated positions of well/barrier satellite minima for some 3–5μm QCL heterostructure systems, calculated from the parameters presented in Vurgaftman *et al.* [50] are indicated in Figure 10.

The most promising high ΔE_c materials systems for short wavelength QCL development are (i) strain compensated InGaAs/AlInAs on InP substrates, (ii) lattice matched (or strain compensated) InGaAs/AlAsSb on InP and (iii) InAs/AlSb on either InAs or GaSb. In the following sections, recent developments in QCL technology based on these systems are discussed. The rapid progress in this area in the past couple of years provides strong confidence in the emergence of high performance QCLs for 3–5μm applications in the near future, and demonstrates the flexibility of the QCL concept in its transferability between widely different heterostructure systems.

3.1 Strain Compensated InGaAs-AlInAs-InP

In the $In_xGa_{1-x}As/Al_yIn_{1-y}As$/InP materials system, which has to date provided the highest performance levels for mid–IR QCLs, the ΔE_c for the InP lattice matched composition (x = 0.53, y = 0.48) is around 520meV. Whilst this is too low to provide efficient operation for emission wavelengths in the 3–5μm range, the

situation can be significantly improved by moving to a pseudomorphic strain compensated regime. Here, the InGaAs quantum well layers are compressively strained with respect to InP by addition of extra indium (*i.e.* x > 0.52), which produces a smaller bandgap. This strain is compensated within an individual well/barrier pair by the addition of extra aluminium (*i.e.* y > 0.48) to the AlInAs layers to produce tensile strain and a larger bandgap. The overall result is a net strain-neutral heterojunction which can have a significantly higher ΔE_c than the InP lattice matched system. The extent to which ΔE_c may be increased is determined by the requirement that individual strained InGaAs or AlInAs layers must not exceed the critical thickness for dislocation formation at the composition of interest. For typical layer thicknesses encountered in QCL design (<~5nm), the most highly strained well/barrier combination possible has x and y values of approximately 0.7 and 0.6 respectively, giving a predicted ΔE_c of approximately 740meV [8,51]. A QCL design based on such strain compensated material was used by Faist *et al.* [8] to demonstrate pulsed laser emission at wavelengths as low as 3.49μm (T=10K), increasing to 3.6μm at 280K, albeit at a rather high threshold current density of ~20kAcm^{-2}. At the time of writing, this remains the shortest wavelength QCL operation yet reported.

More recently, Evans *et al.* [6] have demonstrated record performance levels from strain compensated $In_{0.68}Ga_{0.32}As/Al_{0.64}In_{0.36}As$ QCLs employing the double phonon resonance design discussed in Section 2.3. These lasers, which displayed λ~4.8μm CW operation at temperatures of up to 323K employed high reflectivity rear facet coating to minimise threshold current densities and gold electroplating on top of the laser ridge to maximise extraction of heat from the device. Room temperature threshold current densities as low as 1.7kAcm^{-2} were obtained for CW operation (*cf* J_{th} ~ 1.1 kAcm^{-2} for pulsed operation), with average optical powers in excess of 600mW per facet (Table 1). Although it remains to be seen how far wavelengths can be reduced towards 3μm, this important result demonstrates the significant potential for high performance, short wavelength QCLs based on strain compensated InGaAs/AlInAs.

An interesting variation on strain compensated InGaAs/AlInAs/InP has recently been reported by Semtsiv *et al.* [52]. In this QCL design, the AlInAs barriers are selectively replaced by either thin, tensile strained AlAs barriers or thicker bilayer $In_{0.55}Al_{0.45}As/AlAs$ barriers, strain-compensated by the adjacent $In_{0.73}Ga_{0.27}As$ quantum wells (Figure 11). The slightly compressively strained $In_{0.55}Al_{0.45}As$ layers help to strain compensate the AlAs layers in the hybrid barriers, providing an extra degree of freedom in the design of the relative widths of wells and barriers. The $In_{0.73}Ga_{0.27}As/AlAs$ heterojunction possesses a very large Γ-point conduction band offset (ΔE_c ~ 1.2eV). Compared with the situation described above for GaAs/AlAs, a much greater fraction of this offset can be exploited for QCL design, due to the significantly deeper $In_{0.73}Ga_{0.27}As$ quantum wells. Assuming an $In_{0.73}Ga_{0.27}As$ Γ-point bandgap of 620 meV and an AlAs Γ-X valley separation of 860meV [50], and that the conduction band/valence band offset ratio remains unchanged with In fraction, a value for $\Delta E_c(\Gamma\text{-}X)$ of ~700meV is obtained, compared with $\Delta E_c(\Gamma\text{-}X)$ ~ 130meV for GaAs, as discussed in the previous section. Thus, structures can be designed for short emission wavelengths whilst still maintaining the barrier X-point states well above the upper lasing level of the structure. In fact, the limit on short wavelength performance in this system is thought to be set by the position of the L-

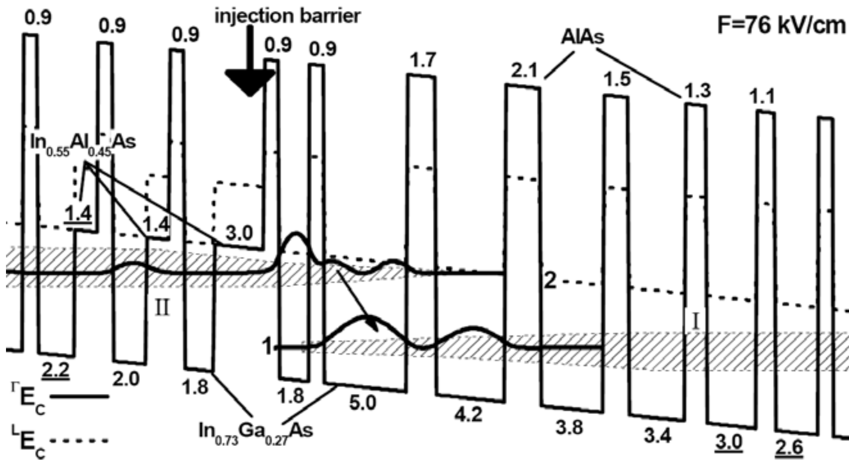


Figure 11. Band-edge diagram of strain-compensated InGaAs/AlInAs/AlAs QCL active region for high performance λ≈3.8μm operation. Reprinted with permission from M. P. Semtsiv, M. Ziegler, S. Dressler, W. T. Masselink, N. Georgiev, T. Dekorsy, and M. Helm, Applied Physics Letters, 85, 1478 (2004). Copyright 2004, American Institute of Physics.

valley conduction band minimum in the $In_{0.73}Ga_{0.27}As$ wells, which lies approximately 90meV above the upper lasing level for the λ ~ 3.8μm design described in [52]. QCLs based on this design have displayed pulsed laser action at wavelengths between 3.7 and 4.2μm with low threshold current densities (J_{th}~ 0.86kAcm^{-2} at 8K, increasing to 6 kAcm^{-2} at room temperature) and peak optical powers in excess of 250mW at room temperature. The performance of these lasers is significantly superior in all respects to that of the slightly shorter wavelength devices (λ ≈ 3.5-3.6μm) reported in [8].

Finally, as a further indication of the potential of the strain compensated InGaAs-AlInAs-AlAs system for short wavelength intersuband devices, intersubband absorption has been observed at wavelengths down to ~2μm [53,54]. However, QCL operation at such short wavelengths is likely to be hampered not only by the proximity of the quantum well L-valley but also by the fact that the lowest AlAs barrier confined X-point state would lie below the upper lasing state [54], leading to carrier leakage problems similar to those encountered in the GaAs/AlAs system described above.

3.2 InGaAs-AlAsSb-InP

As a consequence of its very large Γ-point conduction band offset (ΔE_c~1.6eV [55]), the $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$-InP system has been the subject of considerable interest not only for short wavelength QCLs but also for intersubband optical switches and modulators for λ ≈ 1.3-1.55μm telecommunications wavelengths

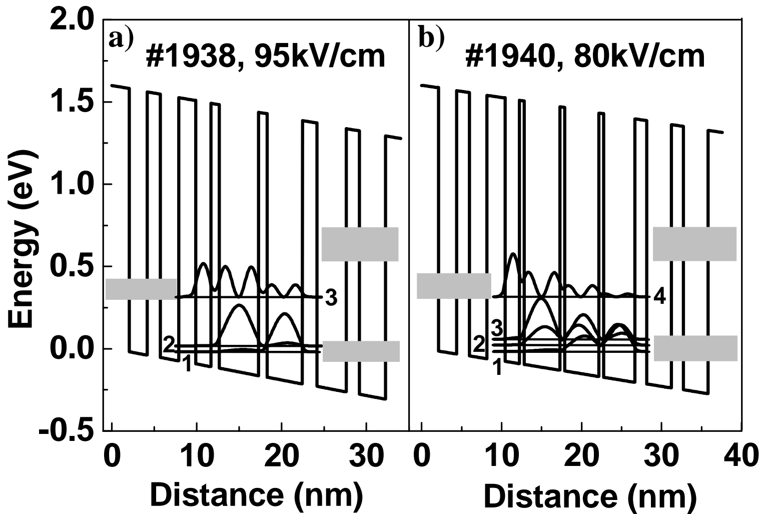


Figure 12. a) Three-well and **b**) double phonon resonance $In_{0.53}Ga_{0.47}As/AlAs_{0.56}Sb_{0.44}$ QCL designs for emission at $\lambda \approx 4.1 \mu m$ and 4.8 μm respectively. The *shaded regions* represent injector minibands. Reprinted with permission from D. G. Revin, L. R. Wilson, E. A. Zibik, R. P. Green, J. W. Cockburn, M. J. Steer, R. J. Airey, and M. Hopkinson, Applied Physics Letters, 84, 1447 (2004). Copyright 2004, American Institute of Physics.

[56,57,58]. $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$ is of particular interest because, in contrast to the InAs/AlSb system described in the next section, it not only provides a large ΔE_c, but also provides lattice-matched compatibility with established InP low-loss waveguide design and device fabrication technology. In addition to its low refractive index, which provides cladding layers with good optical confinement without the need for high doping (which leads to higher optical losses at IR wavelengths due to free carrier absorption), InP has relatively high thermal conductivity (~$0.7 Wcm^{-1}K^{-1}$) which facilitates efficient heat extraction from the operating device, as discussed in Section 2.5. Thermal management can be improved further by epitaxial regrowth of material around the etched laser ridge. This buried heterostructure approach has been found particularly beneficial for high temperature CW operation of QCLs. Again, the use of InP is advantageous here, due to the existence of well established technology for epitaxial regrowth of InP by MOVPE.

Despite the potential advantages of $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$-InP, widespread development of intersubband devices based on this system has been hampered by difficulties with crystal growth of high quality heterostructures. These are mainly associated with the requirement to switch both Group III and Group V species at the well/barrier interface, the large AlAsSb miscibility gap, and the differing, and temperature dependent sticking coefficients of As and Sb [59]. Interface abruptness is difficult to control due to interdiffusion and exchange of Group III/Group V species, and the conduction band offset and the nature of the heterojunction band alignment (Type I or Type II) can be strongly influenced by growth conditions [55]. In addition, even if perfect lattice matching is achieved, the conduction band offset

is influenced by the intrinsic interface strain which is another consequence of the difference in both Group III and Group V species in the well and barrier [60]. Nevertheless, it has been convincingly demonstrated, principally by Mozume and co-workers [55,56,58 and references therein], that $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$ quantum wells of high optical quality with abrupt interfaces and good lattice matching can be produced by careful optimisation and control of MBE growth conditions.

In view of the above, it is clear that development of QCLs based on $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$ presents a considerable challenge, and that much work remains to be done before the considerable potential of this materials system for high performance intersubband devices can be fully realised. Recent work [61,62,63,64] has, however, demonstrated the feasibility of producing QC structures based on $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$-InP. The first observation of intersubband spontaneous emission from $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$ QC devices was reported by Revin *et al.* [61]. Samples with 20 active periods, either of the 3 quantum well or 4 quantum well (double phonon resonance) design were produced (Figure 12), with intended emission wavelengths between 2.9 and 4.8μm. The double phonon resonance design presents a particularly difficult challenge for crystal growth, since it requires very thin barriers (down to two monolayers for the thinnest layers). The samples were grown by MBE, using valved cracker sources for the Group V species, with particular attention being paid to optimisation of the As and Sb fluxes and interface growth interruptions. Prior to growth of the QC structures, $AlAs_{0.56}Sb_{0.44}$ calibration layers were grown to establish growth conditions for the InP lattice matched composition. The Sb flux was monitored using an ion gauge placed at the substrate position and the cracker valve adjusted to obtain a value of $1x10^{-6}$ mbar. An Sb_2:As_2 flux ratio of ~1.5:1 was found to produce material with an acceptable lattice mismatch of $\Delta a/a < 2x10^{-3}$, as measured by high resolution X-ray diffraction. Growth of the QC structures required considerable attention to optimisation of well and barrier growth rates, substrate temperature (480°C) and cracker valve/Group III shutter sequencing and growth interruption at the interfaces. Full details of the growth process are given in [61].

Cross-sectional transmission electron microscopy (TEM) on the samples (Figure 13(a)) shows the growth to have good long range uniformity, and active region thicknesses within less than 3% of the intended values. Well-defined AlAsSb layers are observed even for the 2-monolayer barriers employed in the double phonon resonance design. After processing the wafers into cleaved-edge circular mesa structures, low temperature (T = 20K) electroluminescence measurements revealed clear intersubband spontaneous emission peaks, obeying the expected polarization selection rules, at wavelengths between 3.15 and 5.3μm (Figure 13(b)). Optical powers in the region of tens of nanowatts were observed, similar to the spontaneous emission powers obtained from GaAs-based and "conventional" InP-based QC structures. Following optimisation of active region design and doping density, subsequent work led to the first observation of intersubband lasing in the $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$-InP system [62]. Several active region designs were investigated, paying particular attention to the thicknesses of the injection and exit barriers. The best lasing performance was obtained from samples with a very narrow (~1.1nm) exit barrier, with the active region approaching a bound-to-continuum design. Laser emission was observed from these structures at wavelengths of around

4.3μm (Figure 13(c)), with the best devices displaying laser action up to T = 240K with a characteristic temperature (T_0) of 150K. Clearly, this maximum operating emperature and the rather large threshold current densities observed (J_{th} ~ 6kAcm^{-2} at low temperatures, rising to ~15kAcm^{-2} at 240K) represent performance levels which are significantly below the current state of the art for InGaAs-AlInAs-InP QCLs (Table 1). However, in view of the aforementioned difficulties with crystal growth, it is clear that there is considerable scope for further work to optimise material quality and produce $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$ QCLs with greatly improved performance that fully exploit the highly favourable electronic properties of this materials system.

As a final point, it is worthwhile remembering that $AlAs_xSb_{1-x}$ has an indirect bandgap (X-minimum lowest) for all alloy compositions. However, in contrast to the GaAs/AlAs system, where, as discussed previously, problems arise due to the

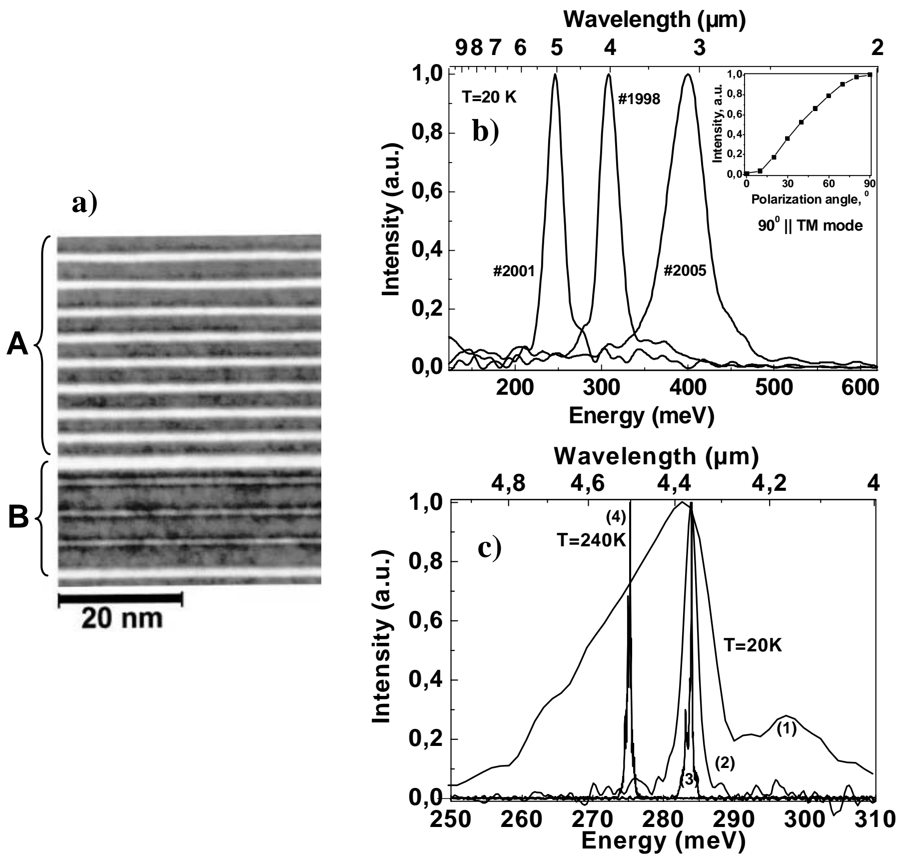


Figure 13. Summary of data obtained from $In_{0.53}Ga_{0.47}As/AlAs_{0.56}Sb_{0.44}$ QC structures: **a**) cross-sectional TEM image from double phonon resonance design (A=injector, B=active region), **b**) spontaneous intersubband emission from QC devices, **c**) emission spectra for λ≈4.5μm bound to continuum QCL, showing sub-threshold spontaneous emission (1,2) and laser emission (3,4).

relatively low Γ-X conduction band offset, $In_{0.53}Ga_{0.47}As$-$AlAs_{0.56}Sb_{0.44}$ has a much larger $\Delta E_c(\Gamma\text{-}X)$ of around 1eV (Figure 9, estimated from [50]) which is highly unlikely to cause difficulties for short wavelength QCL operation. In fact, the short wavelength limit in this system will most likely be set by the positions of the X and L minima in $In_{0.53}Ga_{0.47}As$, which lie about 550meV above the bottom of the quantum well. At present, however, the possible influence of satellite conduction band minima on QCL performance is not fully understood, and as QCL technology moves towards shorter wavelengths and less mature materials systems more work is clearly required to understand these effects.

3.3 InAs-AlSb

Of the III-V materials systems that have successfully been used for QCL development to date, InAs/AlSb is the only one that lies outside mature GaAs or InP-based technology, being based on either InAs or GaSb substrates. Notwithstanding the technological difficulties that this inevitably introduces, InAs/AlSb can in many ways be regarded as the ideal QCL materials system for short wavelength operation, due to its highly favourable electronic properties [65] . As indicated in Figure 10, InAs/AlSb has $\Delta E_c(\Gamma\text{-}\Gamma)$ ~2.1eV, $\Delta E_c(\Gamma\text{-}X)$~1.4 eV, and the lowest satellite valley minimum (L-point) lies approximately 750meV above the bottom of the InAs quantum well layers. All these parameters, which are of great significance particularly for short wavelength operation, are much higher than for any other III-V heterostructure system. In addition, the band-edge effective mass of InAs has a very low value of $m^* \approx 0.023m_0$, which might be expected to provide enhanced intersubband optical gain due to the long non-radiative LO phonon scattering lifetime and large optical dipole matrix element (Section 2.5). It should be remembered, however, that this advantage is somewhat diminished by the strong conduction band non-parabolicity at high electron energies, due to the very narrow bandgap of InAs. This non-parabolicity is usually accounted for in QCL design by introducing an energy-dependent effective mass of the form [66,67]

$$m^*(E) = m^*\left(1 + \frac{2m^*\gamma E}{\hbar^2}\right) \tag{2}$$

where m* is the effective mass at the conduction band minimum, E is the energy relative to the conduction band minimum, and γ is a non-parabolicity parameter obtained from **k.p** calculations [66,67] or from empirical fits of calculated quantum well transition energies to experimental data.

To illustrate the significance of non-parabolicity for QCL design, Figure 14 compares the calculated energy dependent effective mass for InP lattice matched $In_{0.53}Ga_{0.47}As$ ($\gamma=1.13 \times 10^{18}$ m^2), strained $In_{0.7}Ga_{0.3}As$ ($\gamma=1.53 \times 10^{18}$ m^2) and InAs ($\gamma=5 \times 10^{18}$ m^2). The values of γ for the two InGaAs compositions are obtained from the three-band model of Nelson *et al.* [66], and have been shown to work well for a wide range of transition energies in this materials system. The InAs value is taken from 18-band **k.p** calculations [68] combined with empirical fitting to measured

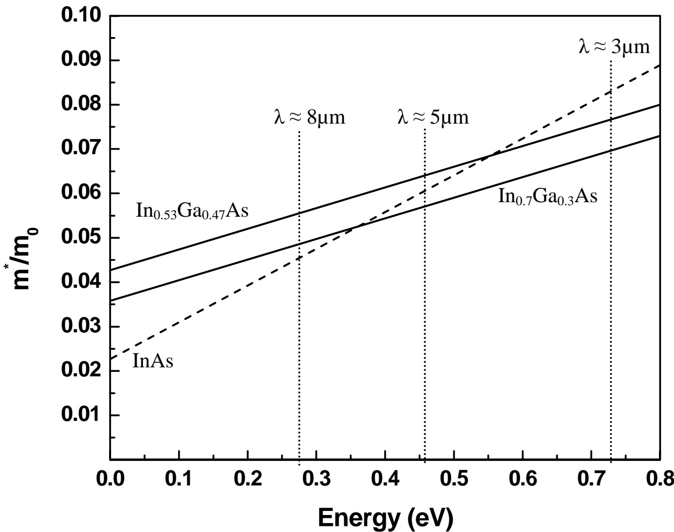


Figure 14. Energy dependence of electron effective mass for various short wavelength QCL materials

transition energies as reported by Becker *et al.* [69]. In addition to the energy dependence of the effective masses, Figure 14 shows the approximate energies of the upper laser level in a three-well QCL design for λ≈3, 5 and 8μm. On the basis of this model, there would seem to be very little effective mass advantage for InAs based QCLs for emission wavelengths below about 7μm, and at shorter wavelengths the InAs effective mass in fact becomes greater than that of both $In_{0.53}Ga_{0.47}As$ and $In_{0.7}Ga_{0.3}As$. Although it would be unwise to place too much trust in this relatively crude treatment of the conduction band structure of InAs at high energies, it seems safe to assume that the advantages of InAs/AlSb QCLs in this respect are not as significant as the low band edge mass in InAs first suggests.

The idea of exploiting the InAs/AlSb system for QCL design was first proposed theoretically by Vurgaftman *et al.* [65]. In fact, these authors suggest the use of $AlAs_{0.16}Sb_{0.84}$ barriers for perfect lattice matching, although arsenic incorporation has subsequently been found to be unnecessary in practice. In this work, a λ≈5.3μm design is simulated, and significant performance improvements predicted relative to InGaAs/AlInAs QCLs. For a 30-period device, with graded AlGaAsSb cladding layers, a theoretical 300K threshold current density of around $800Acm^{-2}$ is calculated for pulsed operation, an improvement of about 40% on the lowest value for InGaAs/AlInAs QCLs reported to date. These authors also predict the possibility of intersubband lasing wavelengths as short as 1.9μm in this system, although the possible effects of satellite conduction band minima are apparently not taken into account in reaching this conclusion.

While intersubband electroluminescence from a Type II "broken-gap" InAs/GaSb/AlSb cascade structure was demonstrated by Ohtani and Ohno [70], the first observation of spontaneous intersubband emission from conventional InAs/AlSb QC structures was reported by Becker *et al.* [69]. In this work, room

temperature spontaneous emission was observed at wavelengths down to 3.7μm with optical powers of the order of tens of nanowatts, although some problems were reported with the electrical stability of the devices. The first QC laser action in this materials system was obtained at the much longer wavelength of λ≈10μm [71], this long wavelength being chosen with the intention of avoiding possible hole generation effects associated with the high electric fields required for short wavelength devices, as discussed above [44]. These devices, which avoided the materials problems associated with quaternary AlGaAsSb dielectric cladding layers by exploiting plasmon enhanced n^+ InAs waveguides, displayed pulsed laser emission up to a maximum temperature of 140K, with low temperature threshold current densities of around $5kAcm^{-2}$.

Progress towards higher performance InAs/AlSb QCLs at shorter wavelengths was made by Teissier *et al.* [72]. A bound to continuum design was used, again with InAs plasmon enhanced waveguides, and pulsed laser emission was achieved at around 6.7μm for temperatures up to 220K. The limitation on maximum operating temperature in this case was the onset of negative differential resistance (NDR) in the electrical characteristics, associated with the breakdown in resonant tunnelling conditions between the injector and upper laser level. In common with the longer

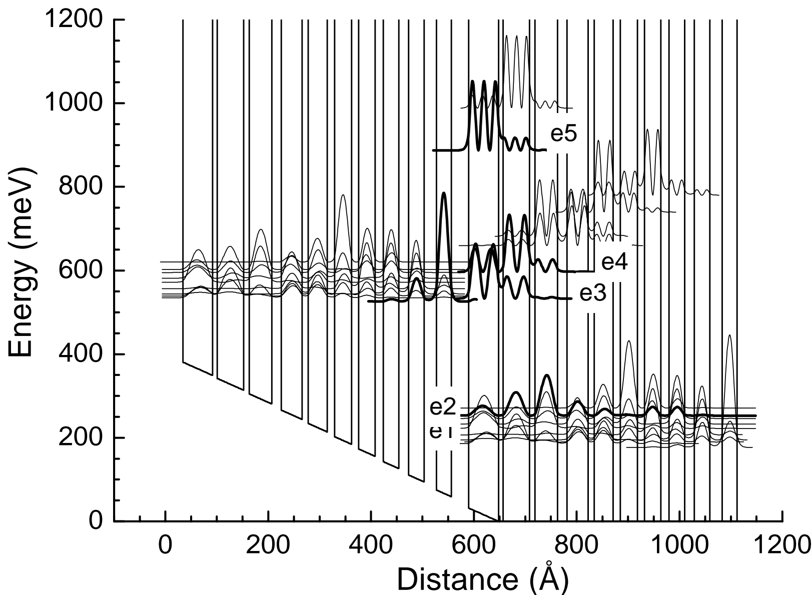


Figure 15. "Two-well" active region InAs/AlSb QCL design for room temperature laser emission at λ≈4.5μm. The confined levels above the upper laser level (e3) provide possible absorption transitions for thermally activated electrons, a potential pitfall for high barrier materials systems. Reprinted with permission from R. Teissier, D. Barate, A. Vicet, C. Alibert, A. N. Baranov, X. Marcadet, C. Renard, M. Garcia, C. Sirtori, D. Revin, and J. Cockburn, Applied Physics Letters, 85, 167 (2004). Copyright 2004, American Institute of Physics.

wavelength device reported in [71], a very high characteristic temperature (T_0) in excess of 200K was observed up to the maximum lasing temperature.

Although the very high barriers provided by the InAs/AlSb system are beneficial in most respects, the problem arises that very thin barriers become necessary for efficient tunnel coupling between electronic states in adjacent wells. Thus, in the devices described in [72], injection barriers as thin as 0.6nm were required. As pointed out in subsequent work by Teissier *et al.* [73], fluctuations in the thickness of this barrier can lead to unpredictable variations in the position of the upper laser level, which causes problems not only for the laser emission wavelength, but also for resonant alignment with the injector (the possible origin of the NDR observed in [72]). Consequently, a new, two-well design was developed for $\lambda\approx4.5\mu m$ operation (Figure 15), in which barriers of 0.9nm minimum width were used. These devices displayed much lower threshold current densities ($J_{th}\approx1.5kAcm^{-2}$ at T = 85K) than previously obtained in this materials system, and provided the first demonstration of pulsed laser emission from InAs/AlSb QCLs at room temperature.

A notable feature of the temperature dependence of this device is that the high T_0 value observed at low operating temperatures is not maintained at T>~200K, where the threshold current begins to increase rapidly, leading to a room temperature threshold current density of around $9kAcm^{-2}$. This was attributed to the onset of intersubband absorption between the E4 and E5 states (Figure 15), as E4 becomes populated with thermally excited carriers from the upper laser level (E3). Such behaviour illustrates another potential pitfall which can arise in the development of intersubband lasers in high ΔE_c materials systems, *i.e.* the existence of confined states above the upper laser level that can absorb the emitted radiation and contribute strongly to optical losses.

Due to the onset of interband absorption, the use of InAs plasmon enhanced cladding layers in InAs/AlSb QCLs limits the minimum emission wavelength to be about 3.7μm. Consequently, quaternary dielectric cladding layers must be developed in order for the lower wavelength limit to be reduced in order to fully realise the potential of this materials system. However, attempts to use such claddings have proved unsuccessful to date, due to poor electrical characteristics associated with deep centres in n-Type tellurium doped AlGaAsSb. Whilst attempts to improve material quality are ongoing, including the investigation of beryllium p-doped claddings which have been found to display lower resistance [74], InAs/AlSb QCLs with dielectric waveguide confinement have yet to be demonstrated at the time of writing. Nevertheless, progress continues with active region design and recent devices have, for the first time, enabled room temperature intersubband spontaneous emission below 3μm ($\lambda\approx2.5\mu m$ [75]).

4 Quantum Cascade Lasers Grown by Metal Organic Vapour Phase Epitaxy

One of the obstacles to the widespread commercial exploitation of QCL technology is the reliance on MBE, which, until very recently, has been used for the growth of all QCL wafers. For many commercial optoelectronic applications, metal-organic vapour phase epitaxy (MOVPE) is the preferred growth method, since it offers the

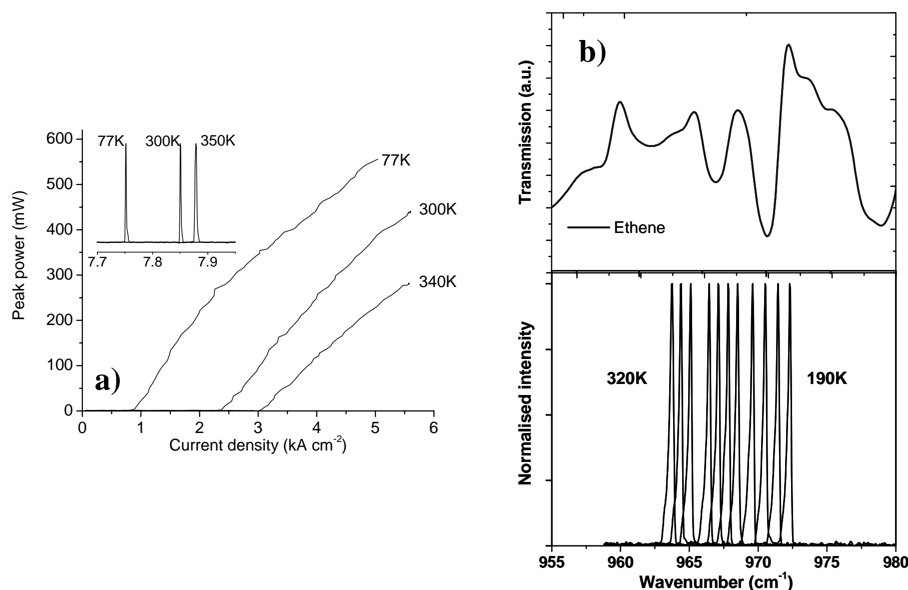


Figure 16. **a**) Temperature dependent L-I characteristics obtained from $\lambda \approx 7.8\mu m$ DFB MOVPE QCL, **b**) Temperature tuning characteristics of $\lambda \approx 10.3\mu m$ DFB MOVPE QCL, showing ability to scan through ethene absorption lines

key advantages of low reactor downtime, ease of maintenance, easy scalability for multi-wafer deposition, low defect densities and higher, more stable and more controllable growth rates. In addition, MOVPE is more suitable for the growth of high quality InP buffer and cladding layers due to lower arsenic contamination during growth. However, for many years it was regarded as doubtful that MOVPE could provide the high interface quality and monolayer thickness control required for the growth of high quality QCLs. In this section, recent work on MOVPE growth will be discussed that should dispel such doubts.

The first demonstration of QCL growth by MOVPE was reported by Roberts *et al.* [76]. In this work, 35-period GaAs-AlGaAs QCLs of a 3-well active region design were grown in an atmospheric pressure MOVPE reactor, at a constant growth rate of 2.7μm per hour. Wafers were grown using continuous deposition, with trimethyl-aluminum (TMA) being periodically switched into the reactor for the growth of each $Al_{0.45}Ga_{0.55}As$ barrier layer. The response time of this reactor was ~1.5s and consequently the $Al_{0.45}Ga_{0.55}As$ barriers were expected be completely resolved within two monolayers. However, in practice it can be difficult to achieve such abrupt changes in composition due to alkyl adsorption/desorption effects within the critical stainless steel pipework of the gas handling sections of the reactor. In order to minimize the retention of reagents, each source had an individually purged vent/run pneumatic valve, which also carried a proportion of the carrier gas. The design was configured to minimize the surface area to length ratio of each connecting pipe and injected the reagent after a 1.5 l min^{-1} hydrogen dilution. Conventional pressure balancing between the vent line and reactor further optimized

the abrupt transfer of reagents. In this growth configuration, interface control depended only on the rapid switching of TMA.

Whilst these MOVPE-grown QCLs displayed laser operation, their performance was significantly inferior to that of structures of identical design grown by MBE [38]. The MOVPE lasers had low-temperature threshold current densities ~10kAcm^{-2} and a maximum operating temperature (T_{max}) of 140K, compared with J_{th} ~ 6kAcm^{-2} and T_{max} ~305K for the MBE devices. In addition the emission wavelength of the MOVPE QCLs ($\lambda\approx$11.8μm) was significantly longer than the designed wavelength ($\lambda\approx$9.8μm), even though cross-sectional transmission electron microscopy revealed the thickness of each active period to be in excellent agreement with the required value. This wavelength variation was attributed to non-optimum growth conditions leading to the occurrence of compositional grading at the GaAs-AlGaAs interfaces. This wavelength shift was thought to be largely responsible for the relatively poor performance of these devices, since the plasmon enhanced doped GaAs waveguide [35] employed in these structures was optimised for the 9.8μm design wavelength, with calculated waveguide losses increasing by a factor of 2 for the 11.8μm emission wavelength observed. In addition the graded interfaces present a lower effective tunnelling barrier, leading to enhanced thermally-activated loss of carriers from the upper laser level at high temperatures.

Subsequent work on MOVPE growth, focussing on InGaAs-AlInAs-InP QCLs has produced lasers with much higher performance than those described above [77]. In this work, interface control was improved by employing a lower growth pressure (150 torr) to decrease the reactor transit time, combined with a lower growth rate. Wafers were grown using conventional trimethyl reagents, arsine, and phosphine. Growth control was found to be very good, with all wafers showing excellent homogeneity and agreement with expected superlattice period as measured by X-ray diffraction and cross sectional TEM. X-ray diffraction wafer mapping measurements showed excellent cross-wafer uniformity, with the measured superlattice periods varying by only ±2% across a 2 inch diameter wafer. Studies were carried out on the dependence of the performance of QCLs fabricated from these wafers on active region growth rate, with the best devices using a rate of around 1μm h^{-1}. These lasers, which displayed laser emission at ~8.6μm (12K) to 8.9μm (290K) displayed low temperature threshold current densities of ~1.5kAcm^{-2}, increasing to 6kAcm^{-2} at the maximum lasing temperature of 290K. In this case, the performance of these three-well, vertical transition MOVPE QCLs was comparable to lasers of the same design grown by MBE [78]. In addition the laser emission wavelength was in good agreement with the design value of 8.5μm Whilst the growth rate used for these active regions (0.8μm h^{-1}) was similar to that used in MBE, the ability to increase to a much higher growth rate for the cladding regions (typically 3μm h^{-1}) means that an entire QCL wafer can typically be grown in less than four hours, compared with the ~12 h required for the MBE growth of a similar structure. Further growth runs revealed a high degree of repeatability, with a series of seven laser wafers designed to emit at 9.0μm, all displaying emission wavelengths in the range 8.75 to 9.1μm.

In order to assess the suitability of MOVPE-grown QCLs for gas sensing applications, further wafers were grown with intended laser emission wavelengths of 7.8 and 10.3μm, which correspond with absorption lines in the infrared spectrum of several molecular species including methane, ethene and nitrous oxide [79]. The active regions of these structures were based on a double phonon resonance design,

with barriers as thin as 3ML, placing high demands on crystal growth. Fabry Perot cavity lasers fabricated from these wafers displayed room temperature threshold current densities of 3.3 and 2.9kAcm^{-2} for the 7.8 and 10.3µm devices respectively, with lasing observed up to T = 320K (the highest temperature obtainable with the measurement system). The wafers were subsequently processed into distributed feedback (DFB) grating devices by G. Strasser and co-workers at the Technical University of Vienna, with the aim of producing single mode lasers suitable for gas sensing. For both samples, a first-order, index-coupled grating of suitable period was used, defined by optical contact lithography and etched to a depth of 1µm in the InP upper cladding layer by reactive ion etching. The processed lasers displayed single mode operation with room temperature wavelengths of 7.88 and 10.38µm.

The room temperature threshold current densities of the DFB lasers were ~3.0kAcm^{-2} and 2.4kAcm^{-2} for the long and short wavelength devices respectively, comparable with the best values obtained from MBE-grown QCLs of similar design. Both devices displayed a temperature tuning coefficient of the single mode emission of around 0.08cm^{-1}K^{-1}, providing a spectroscopically useful tuning range for moderate changes in temperature (Figure 16). In contrast to our GaAs-based MOVPE QCLs, where significant interfacial grading occured, the interfaces in our InP-based devices are found to be at least as abrupt as in those grown by MBE. This is apparent from comparative cross sectional scanning tunnelling microscopy studies (Figure 17) carried out by P Offermans *et al.* (Eindhoven University of Technology) on our QCLs and similar MBE-grown structures [80].

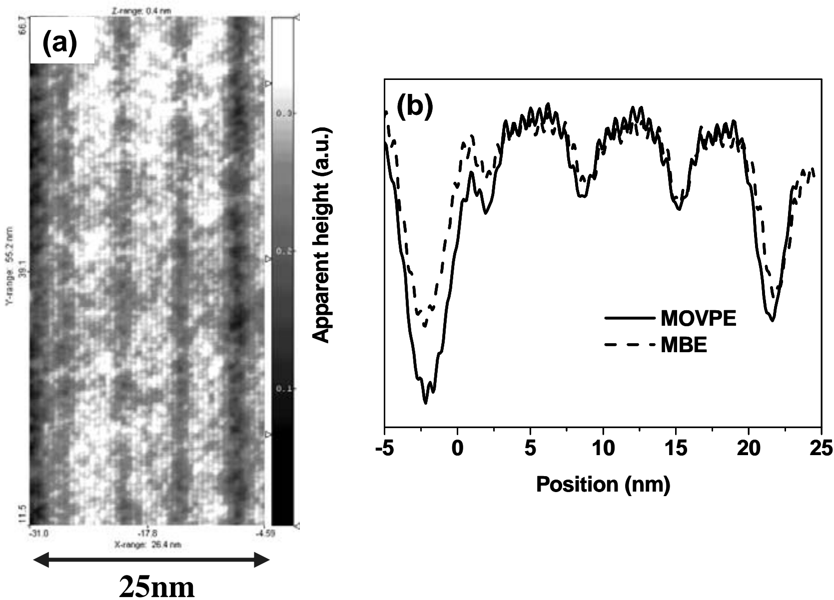


Figure 17. a) Cross sectional STM image of the active region of a λ≈9µm MOVPE QCL. **b**) Comparison of the brightness profile from this image with that obtained from an MBE-grown sample of the same design, as reported in [80].

Very recently, MOVPE growth of InP-based QCLs has also been reported by Troccoli *et al.* [81]. For a λ≈7.6μm device, low temperature and room temperature pulsed threshold current densities were achieved which were similar to those of the λ≈7.8μm QCL reported in [79]. However, the devices reported in [81] also displayed CW operation at temperatures up to ~190K.

5 Conclusions

As a result of continual evolution of active region design, quantum cascade lasers have achieved unprecedented levels of performance for solid state sources in the 5-12μm region. QCLs also offer unique possibilities for the development of a new class of devices exploiting strong intersubband optical nonlinearities. Although the highest device performance is still obtained with the original InGaAs/AlInAs/InP materials system, QCL operation has now been demonstrated for a wide range of III-V materials, and continued optimisation of crystal growth and device design is expected to lead to significantly enhanced operation for high bandgap offset QCLs in the 3–5μm region. Although MBE growth still dominates, the emergence of MOVPE as a viable growth route for high performance QCLs offers the potential for accelerated technological exploitation of QCLs, as well as opening the field to a wider range of research groups. As QCLs enter their second decade, innovative new areas of research continue to emerge, consolidating the position of these devices at the forefront of mid-IR optoelectronic technology.

Acknowledgements

The author wishes to express his thanks to close colleagues in the Sheffield Mid-IR group who have contributed to the work described here, particularly Luke Wilson, Damian Carder, Richard Green, Dmitry Revin and Evgeny Zibik. Thanks are also due to Matthew Steer, Andrey Krysa, John Roberts, Mark Hopkinson and Robert Airey from the EPSRC National Centre for III-V Technologies at Sheffield, for their invaluable contributions in sample growth and processing. Financial support has been provided by the Engineering and Physical Sciences Research Council, UK, and by the European Union SUPERSMILE and ANSWER projects.

References

1. J. Faist, F. Capasso, D. L. Sivco *et al.*. Quantum cascade laser. Science 1994; **264**: 553-556.
2. J. Faist, F. Capasso, C. Sirtori, *et al.* in Intersubband Transition in Quantum Wells: Physics and Device Applications II, Semiconductors and Semimetals Vol. 66, edited by H. C. Liu and F. Capasso ;Academic, San Diego: 1–83 (2000)
3. C. Gmachl, F. Capasso, D. L. Sivco *et al.*. Recent progress in quantum cascade lasers and applications. Rep. Prog. Phys. 2001; **64**: 1533-1601
4. D. A. Carder. PhD thesis, University of Sheffield, 2003

5. M. Beck, D. Hofstetter, T. Aellen *et al.*. Continuous-wave operation of a mid- infrared semiconductor laser at room temperature, Science 2002; **295**: 301-305
6. A. Evans, J. S. Yu, S. Slivken *et al.*. Continuous-wave operation of λ ~ 4.8μm quantum- cascade lasers at room temperature. Appl. Phys. Lett. 2004; **85**: 2166- 2168
7. S. Blaser, D. A. Yarekha, L. Hvozdara *et al.*. Room-temperature, continuous-wave, single- mode quantum-cascade lasers at λ≈5.4μm. Appl. Phys. Lett. 2005; **86**: 041109-041111
8. J. Faist, F. Capasso, D. L. Sivco *et al.*. Short wavelength (λ ~ 3.4μm) quantum cascade laser based on strained compensated InGaAs/AlInAs. Appl. Phys. Lett. 1998; **72**: 680-682
9. R. Colombelli, F. Capasso, C. Gmachl *et al.*. Far-infrared surface-plasmon quantum-cascade lasers at 21.5μm and 24μm wavelengths. Appl. Phys. Lett. 2001; **78**: 2620-2622
10. L. R. Wilson, D. A. Carder, J. W. Cockburn *et al.*. Intervalley scattering in GaAs-AlAs quantum cascade lasers. Appl. Phys. Lett. 2002; **81**: 1378-1380
11. L. Ajili, G. Scalari, J. Faist *et al.*. High power quantum cascade lasers operating at λ≈87 and 130μm. Appl. Phys. Lett. 2004; **85**: 3986-3988
12. J. Faist, F. Capasso, C. Sirtori *et al.*. Laser action by tuning the oscillator strength. Nature 1997; **387**: 777- 782
13. A. Tredicucci, C. Gmachl, F. Capasso *et al.*. A multiwavelength semiconductor laser. Nature 1998;**396**: 350-353
14. C. Gmachl, D. L. Sivco, J. N. Baillargeon *et al.*. Quantum cascade lasers with a heterogeneous cascade : Two-wavelength operation. Appl. Phys. Lett. 2001; **79**: 572-574
15. C. Gmachl, D. L. Sivco, R. Colombelli *et al.*. Ultra-broadband semiconductor laser. Nature 2002; **415**: 883 - 887
16. F. Capasso, C. Sirtori, A. Y. Cho. Coupled quantum well semiconductors with giant electric field tunable nonlinear optical properties in the infrared. IEEE J. Quantum Electron. 1994; **30**: 1313-1326
17. N. Owschimikow, C. Gmachl, A. Belyanin *et al.*. Resonant second-order nonlinear optical processes in quantum cascade lasers. Phys. Rev. Lett. 2003; **90**: 043902-043905
18. J.-Y. Bengloan, A. De Rossi, V. Ortiz *et al.*. Intracavity sum-frequency generation in GaAs quantum cascade lasers. Appl. Phys. Lett. 2004; **84**: 2019-2021
19. R. Ferreira and G. Bastard. Evaluation of some scattering times for electrons in unbiased and biased single-and multiple-quantum-well structures. Phys. Rev. 1985; **B40**: 1074-1086
20. J. W. Cockburn, M. S. Skolnick, D. M. Whittaker *et al.*. Evidence for population inversion in excited electron states of a double barrier resonant tunnelling structure. Appl. Phys. Lett. 1994; **64**: 2400-2402
21. J. Faist, F. Capasso, C. Sirtori, *et al.*. Continuous wave operation of a vertical transition quantum cascade laser above T = 80K. Appl. Phys. Lett. 1995; **67**: 3057-3059.
22. J. Faist, F. Capasso, C. Sirtori *et al.*.. High power mid-infrared (λ~5μm quantum cascade lasers operating above room temperature. Appl . Phys. Lett. 1996; **68**: 3680-3682.
23. D. Hofstetter, M. Beck, T. Aellen *et al.*. High temperature operation of distributed feedback quantum cascade lasers at 5.3μm, Appl. Phys. Lett. 2001; **86**: 396-398
24. T. Aellen, S. Blaser, M. Beck *et al.*. Continuous-wave distributed-feedback quantum-cascade lasers on a Peltier cooler. Appl. Phys. Lett. 2003; **83**: 1929-1921
25. C. Sirtori, F. Capasso, J. Faist *et al.*. Resonant tunneling in quantum cascade lasers. IEEE J. Quantum Electron. 1998; **34**: 1722-1729
26. J. Faist, M. Beck, T. Aellen *et al.*. Quantum-cascade lasers based on a bound-to-continuum transition. Appl. Phys. Lett. 2002; **78**: 147-149

27 G. Scarmarcio, F. Capasso, C. Sirtori *et al.*. High-power infrared (8 micron wavelength) superlattice lasers. Science 1997; **276:** 773-776
28. A. Tredicucci, F. Capasso, C. Gmachl *et al.*. High performance interminiband quantum cascade lasers with graded superlattices. Appl. Phys. Lett. 1998; **73:** 2101-2103.
29. R. Maulini, M. Beck, J. Faist *et al.*. Broadband tuning of external cavity bound-to-continuum quantum-cascade lasers. Appl. Phys. Lett. 2004; **84**: 1659-1661
30. B. N. Murdin, Key issues for mid-infrared emission. Phil. Trans. R. Soc. Lond. A 2001; **359:** 459-473
31. C. Sirtori, J. Faist, F. Capasso *et al.*. Quantum cascade laser with plasmon-enhanced waveguide operating at 8.4μm wavelength. Appl. Phys. Lett. 1995; **66**: 242-3244
32. R. P. Green, L. R. Wilson, D. A. Carder *et al.*. Room temperature GaAs-based quantum-cascade laser with GaInP waveguide cladding. Electron. Lett. 2002; **38**: 1539-1541
33. C. Sirtori, P. Kruck, S. Barbieri, *et al.*. GaAs/$Al_xGa_{1-x}As$ quantum cascade lasers. Appl. Phys. Lett. 1998; **73**: 3486-3488
34. L. R. Wilson, P. T. Keithley, J. W Cockburn *et al.*. A comparison of the performance of GaAs-AlGaAs and InGaAs-AlInAs quantum cascade lasers. Electron. Lett. 1999; **35**: 2034-2035
35. C. Sirtori, P. Kruck, S. Barbieri, *et al.*. Low-loss Al-free waveguide for unipolar semiconductor lasers. Appl. Phys. Lett. 1999; **75**: 3911-3913
36. H. Page, C. Becker, A. Robertson *et al.*. 300K operation of a GaAs-based quantum cascade laser at $\lambda \approx 9\mu m$. Appl. Phys. Lett. 2001; **78:** 3529-3531.
37. C. Pflügl. W. Schrenk, S. Anders *et al.*. Hight-temperature performance of GaAs-based bound-to-continuum quantum-cascade lasers. Appl. Phys. Lett. 2003; **83**: 4698-4700
38. C. Sirtori, H. Page, C. Becker *et al.*. GaAs-AlGaAs quantum cascade lasers: Physics, technology, and prospects", IEEE J.Quantum Electron 2002; **38**: 547-558
39. C. Becker, C. Sirtori, H. Page *et al.*. AlAs/GaAs quantum cascade lasers based on large conduction band discontinuity. Appl. Phys. Lett. 2000; **77**: 463-465
40. D. A. Carder. L. R. Wilson, R. P. Green *et al.*. Room-temperature operation of an InAs-GaAs-AlAs quantum cascade laser. Appl. Phys. Lett. 2003; **82**: 3409-3411
41. H. Page, S. Dhillon, M. Calligaro *et al.*. Improved CW operation of GaAs-Based QC lasers: T_{max} = 150K. IEEE J. Quantum Electron. 2004; **40**; 665-672
42. W. Schrenk, N. Finger, S. Gianordoli *et al.*. Continuous-wave operation of distributed feedback AlAs/GaAs superlattice quantum cascade lasers. Appl. Phys. Lett. 2000; **77**: 3328-3330
43. R. Köhler, A. Tredicucci, F. Beltram et *et al.*. Terahertz semiconductor-heterostructure laser. Nature 2002; **417**: 156–159
44. K. Ohtani, H. Sakuma, H. Ohno: InAs-based quantum cascade light emitting structures containing a double plasmon waveguide, J. Cryst. Growth 2003; **251**: 718-722.
45. H. Page, C. Becker, A. Robertson, G. Glastre, V. Ortiz, C. Sirtori: 300K operation of a GaAs-based quantum-cascade laser at $\lambda \sim 9\mu m$, Appl. Phys. Lett. 2001; **78**: 3529-3531
46. L. R. Wilson, J. W. Cockburn, M. J. Steer *et al.*. Decreasing the emission wavelength of GaAs-AlGaAs quantum cascade lasers by the incorporation of ultrathin InGaAs layers. Appl. Phys. Lett. 2001; **78**: 413-415
47. J. J. Finley, R. Teissier, M. S. Skolnick *et al.*. Role of the X minimum in transport through AlAs single barrier structures. Phys. Rev. B 1998; **58**: 10619-10628
48. J. W. Cockburn. L. R. Wilson, D. A. Carder *et al.*. The effect of intervalley scattering on the performance of GaAs-AlAs quantum cascade lasers. Proceedings of the 26th International Conference on the Physics of Semiconductors, eds. J. R. Long and J. H Davies, 117-123 (2003)

49. E. E. Mendez, W. I. Wang, E. Calleja *et al.*. Resonant tunneling via *X*-point states in AlAs-GaAs-AlAs heterostructures. Appl. Phys. Lett. 1987; **50**: 1263-1265
50. I. Vurgaftman, J. R. Meyer, L. R. Ram-Mohan. Band parameters for III-V compound semiconductors and their alloys. J. Appl. Phys. 2001; **89**: 5815-5875
51. C. G. Van de Walle: Band lineups and deformation potentials in the model solid theory, Phys. Rev. B 1989; **39**:1871-1883
52. M. P. Semtsiv, M. Ziegler, S. Dressler *et al.*. Above room temperature operation of short wavelength (λ = 3.8μm) strain-compensated $In_{0.73}Ga_{0.27}As$-AlAs quantum-cascade lasers, Appl. Phys. Lett 2004; **85**: 1478-1480
53. N. Georgiev, T. Dekorsky, F. Eichhorn *et al.*. Short-wavelength intersubband absorption in strain-compensated InGaAs/AlAs quantum well structures grown on InP. Appl. Phys. Lett. 2003; **83**: 210-212
54. K. T. Lai, R. Gupta, M. Missous, *et al.*. Intersubband absorption from 2 to7μm in strain-compensated double-barrier $In_xGa_{1-x}As$ multiquantum wells. Semicond. Sci. Technol. 2004); **19:** 1263-1267
55. N Georgiev, T. Mozume. Photoluminescence study of InGaAs/AlAsSb heterostructure, J. Appl. Phys. 2001; **89**, 1064-1069
56. H. Yoshida, T. Mozume, A. Neogi *et al.*. Ultrafast all-optical switching at 1.3μm/1.5μm using InGaAs/AlAsSb/InP coupled double quantum well structure for intersubband transitions, Electron. Lett. 1999; **35**: 1103-1105
57. T. Asano, M. Tamura, S. Yoshizawa *et al.*. Pump-probe measurement of ultrafast all-optical modulation based on intersubband transition in *n*-doped quantum wells. Appl. Phys. Lett. 2000; **77**:19-21
58. T. Mozume, J. Kasai, N. Georgiev*et al.*. Ultralow intersubband absorption saturation intensity at communication wavelength achieved in novel strain compensated InGaAs/AlAs/AlAsSb quantum wells grown by molecular beam epitaxy, Jpn. J. Appl. Phys. 2003; **42**: 5500-5507
59. G. Almuneau, E. Hall, S. Mathis *et al.*. Accurate control of Sb composition in AlGaAsSb alloys on InP substrates by molecular beam epitaxy, J. Cryst. Growth 2000; **208**: 113-116
60. M. S. Hybertsen: Role of interface strain in a lattice-matched heterostructure, Phys. Rev. Lett. 1990; **64**: 555-558
61. D. G. Revin, L. R. Wilson, E. A. Zibik *et al.*. λ~4-5.3μm intersubband emission from InGaAs-AlAsSb quantum cascade structures. Appl. Phys. Lett. 2004; **84**: 1447-1449
62. D. G. Revin, L. R. Wilson, E. A. Zibik, *et al.*: InGaAs-AlAsSb quantum cascade lasers, Appl. Phys. Lett. 2004; **85**: 3992-3994
63. D. G. Revin, M. J. Steer, L. R. Wilson *et al.*. InGaAs-AlAsSb quantum cascade structures emitting at 3.1 μm. Electron. Lett. 2004; **40**: 874-875
64. Q.K.Yang, C.Manz, W. Bronner *et al.*. Room-temperature intersubband emission from GaInAs-AlAsSb quantum cascade structure. Electron. Lett. 2004; **40**: 1339-1340
65. I. Vurgaftman, J. R. Meyer, F. H. Julien *et al.*. Design and simulation of low-threshold antimonide intersubband lasers. Appl. Phys. Lett. 1998; **73**: 711-713
66 D. F. Nelson, R. C. Miller, D. A. Kleinman: Band non-parabolicity effects in semiconductor quantum wells, Phys. Rev. B 1987; **35**: 7770-7773
67 C. Sirtori, F. Capasso, J. Faist *et al.*. Nonparabolicity and a sum rule associated with bound-to-bound and bound-to-continuum intersubband transitions in quantum wells: Phys. Rev. B 1994; **50**: 8663-8674

68. I. Prevot , B. Vinter, F. H. Julien *et al.*. Experimental and theoretical investigation of interband and intersubband transitions in Type-II InAs/AlSb superlattices, Phys. Rev. B, 2001; **64**: 195318-195323
69 C. Becker, I. Provot, X. Marcadet *et al.*. InAs/AlSb quantum-cascade light-emitting devices in the 3–5 μm wavelength region. Appl. Phys. Lett. 2001; **78**: 1029-1031
70 K. Ohtani, H. Ohno: Intersubband electroluminescence in InAs/GaSb/AlSb Type-II cascade structures, Appl. Phys. Lett. 1999; **74**: 1409-1411
71. K. Ohtani, H. Ohno: InAs/AlSb quantum cascade lasers operating at 10μm, Appl. Phys. Lett. 2003; **82**: 1003-1005
72. R. Teissier, D. Barate, A. Vicet *et al.*. InAs/AlSb quantum cascade lasers operating at 6.7μm, Electron. Lett. 2003; **39**: 1253-1254
73. R. Teissier, D. Barate, A. Vicet *et al.*. Room temperature operation of InAs/AlSb quantum cascade lasers, Appl. Phys. Lett. 2004; **85**:167-169
74 A. Baranov, private communication.
75. D. Barate, R. Teissier, Y. Wang *et al.*. Short wavelength intersubband emission from InAs/AlSb quantum cascade structures, Appl. Phys. Lett., 2005; in press.
76. J. S. Roberts, R. P. Green, L. R. Wilson *et al.*. Quantum cascade lasers grown by metalorganic vapor phase epitaxy. Appl. Phys. Lett. 2003; **82**: 4221-4223
77. R. P. Green, A. Krysa, J. S. Roberts, D. G. Revin, L. R. Wilson, E. A. Zibik,, W. H. Ng., J. W. Cockburn: Room-temperature operation of InGaAs/AlInAs quantum cascade lasers grown by metalorganic vapor phase epitaxy, Appl. Phys. Lett. 2003; **83**: 1921-1923
78. C. Gmachl, F. Capasso, J. Faist *et al.*. Continuous-wave and high-power pulsed operation of index-coupled distributed feedback quantum cascade laser at $\lambda \approx 8.5$ μm. Appl. Phys. Lett. 1998; **72**, 1430-1432
79. R. P. Green, L. R. Wilson, E. A. Zibik *et al.*. High-performance distributed feedback quantum-cascade lasers grown by metalorganic vapor phase epitaxy, Appl. Phys. Lett. 2004; **85**: 5529-5531
80. P. Offermans, P. M. Koenraad, J. H. Wolter *et al.*. Digital alloy interface grading of an InAlAs/InGaAs quantum cascade laser studied by cross-sectional scanning tunneling microscopy. Appl. Phys. Lett. 2003; **83**: 4131-4133
81. M. Troccoli, D. Bour, S. Corzine *et al.*. Low-threshold continuous-wave operation of quantum-cascade lasers grown by metalorganic vapour phase epitaxy. Appl. Phys. Lett. 2004; **85**: 5842-5844

Part III

LEDs and Detectors

Mid-infrared Electroluminescence in LEDs Based on InAs and Related Alloys

A. Krier, X.L. Huang and V.V. Sherstnev*
Lancaster University, Lancaster, LA1 4YB, UK.

* Visiting research Fellow from The Ioffe Institute, St. Petersburg, Russia

1 Introduction

Mid-infrared light emitting diodes which operate in the (2-5μm) spectral region have been extensively studied in recent years [1-3] due to their increasing potential for applications in a number of areas including; optical gas sensors for pollution monitoring, remote sensing, industrial process control, spectroscopy, optical alignment and infrared countermeasures[4,5]. Consequently there has been much research interest aimed at providing monochromatic sources, such that a number of mid-infrared light-emitting diodes (LEDs) operating at different target wavelengths have now been realised [6-9]. Such sources are particularly attractive as key components in infrared gas sensor instrumentation based on optical absorption because many of the target gases, such as CH_4 CO_2 and CO, have their fundamental absorption bands in the mid-infrared spectral region. There are well established markets for infrared gas sensors, as they are recognised as being truly gas specific and therefore suitable for reliable gas detection in a wide variety of applications. Although some instruments which make use of thermal infrared sources already exist, the recent progress in developing room temperature LEDs [10-14] for the mid-infrared range holds considerable promise for extending the use of infrared methods still further and to provide superior instrumentation. Currently, pulsed filament sources suffer from inherently slow modulation, whereas incandescent bulbs are limited to < 4μm by the glass envelope of the source and require burning in before they become stable. From an engineering viewpoint, LEDs do not have these drawbacks and offer the additional advantages of high emittance, room-temperature operation, high modulation rate, small size, simplicity, low power consumption, and are likely to be more cost-effective than diode lasers for deployment in practical gas sensors.

In order to obtain low detection thresholds (<10ppm), the optical output power of light sources used for such applications should be as high as possible [15]. Most current sensors use low cost, simple optical arrangements which often collect only a small fraction of the radiation emitted from the source. Estimates suggest that mid-infrared LEDs must be able to produce sufficient output power (~ 10μW) within the absorption band of interest for them to be used in a practical gas sensor. The aim of this chapter is to give an overview of the physics and technology associated with the realisation of efficient mid-infrared LEDs. The physical processes which limit device performance will be considered with respect to different materials and operating conditions, to give an insight into the problems facing the design engineer.

1.1 Background

One of the first reported mid-infrared LEDs was made by Melngailis and Rediker in 1966 who produced a room temperature LED by Zn diffusion into InAs [16] and observed recombination into Zn acceptor states at 3.7μm as part of an investigation of InAs diode lasers which they were also the first to realise and report [17]. Although surprisingly effective, these devices were only homostructures and it was not long before heterojunction LEDs were fabricated using liquid phase epitaxy (LPE). The advent of epitaxial growth technology has had the effect of increasing the range of materials available which extended the wavelength range and also resulted in more efficient LEDs and lasers based on single or double heterostructures (DH). Much attention has been given to InAs and GaSb based alloys (*e.g.,* InAsSbP, InGaAs, InGaAsSb and InAsSb) [18] as an attractive materials system for the fabrication of optoelectronic devices in the near to middle infrared wavelength region [19,20]. A number of LEDs have now been developed operating at various wavelengths covering the mid-infrared spectral region from 1-5μm. The literature is too extensive to review here in detail, but recent results include, the fabrication of LEDs tuned to the absorption bands of a number of target gases, including for moisture measurement at 1.94 μm [21], CH_4 at 3.3 μm [10], CO_2 at 4.2 μm [22], CO at 4.6 μm [2,13], HF at 2.5 μm [12] and the detection of H_2S at 3.8 μm[23].

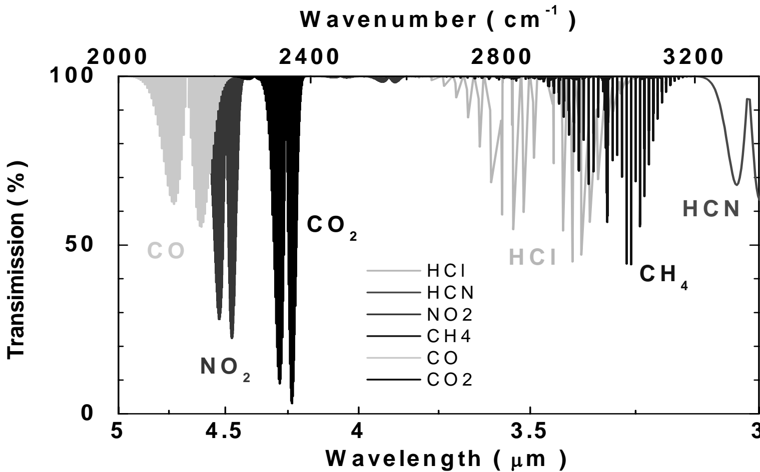


Figure 1. Some target gas absorption spectra in the mid-infrared spectral range

With the advent of MBE and MOVPE a variety of quantum structures have been employed to access longer wavelengths in III-V systems. These include InAs/InAsSb strained quantum well LEDs which benefit from Auger suppression in Type II structures to yield emission at 5 μm and 8 μm [24,25] which is 64% lower than the alloy band gap. Unipolar quantum cascade LEDs based on InGaAs/AlInAs have also been able to access the 8-13 μm wavelength region but so far have only

low power output (~6nW, using 20% duty cycle at 600mA) near room temperature [26-28]. More recently, a strained layer superlattice (SLS) LED has been demonstrated consisting of a p-i-n doped SLS grown on a mismatched InAs substrate, with a strained AlSb barrier to prevent electron migration [25]. The LED operates at 4.2μm with an internal efficiency of 2.8%. Without the constraint of having to grow pseudomorphically on the substrate, it is envisaged that the operating wavelength may be further extended by increasing the Sb composition in the superlattice.

Room temperature electroluminescence has also been observed from InSb based sources using a thin $In_{1-x}Al_xSb$ barrier to provide carrier confinement (& exclusion)[29]. The devices have a $p^+\underline{p}^{\pm}$ л n^+ structure, with confinement of the electrons and holes in the л region provided by the wider gap $\underline{p}^{\pm}$ and degenerate n^+ regions respectively. Because of the high lattice mismatch a composition of $x = 0.17$ is used and the barrier thickness is kept to 20 nm consistent with strain requirements. The result is a room temperature LED which operates at 5.5 μm with an internal efficiency of 5% and an output power of 10 μW (2.5% duty cycle). The performance can be significantly enhanced by incorporation of a Winston cone into the structure to improve optical extraction [30]. Under reverse bias these $p^+\underline{p}^{\pm}$ л n^+ devices have the interesting property that they undergo minority carrier extraction at the pn junction and exclusion at the isotype junction, leading to Auger suppression [31] and negative luminescence associated with a lower than equilibrium carrier concentration in the active (л) region [32,33].

Whereas there has been a huge effort to develop HgCdTe for infrared detectors, there has been little investigation of the electroluminescence emission properties of this material [34]. However, improvements in growth by MBE and MOVPE have now led to the observation of laser action [35,36]. LPE technology has been used to grow LEDs on sapphire [37] and on CdZnTe (CZT) substrates [38] with 300K internal efficiencies as high as 6% at 3.5μm. The latter using backside emission and optical coupling with a CZT lens to increase optical extraction efficiency. More recently, similar types of $p^+\underline{p}^{\pm}$ л n^+ structures to those implemented above for InSb have been used in the HgCdTe material system resulting in emission at longer wavelengths (8-12μm) [39]. The internal quantum efficiency obtained at low injection was $4x10^{-4}$ with a maximum output power of 6nW at 295K (50% duty cycle) from an 80μm diameter LED. This low efficiency value is indicative of a large concentration of SRH centres based on a theoretical efficiency of $4x10^{-2}$ assuming an 'Auger-limited ' (CHCC) recombination mechanism.

By comparison, research on Pb-salt LEDs is very sparse indeed. Although PbSSe LEDs operating at 4.6μm with an output power of ~ 100nW at room temperature were reported some time ago [40], most research effort has been dedicated towards the fabrication of tunable diode lasers for molecular spectroscopy. The result is that Pb-salt lasers have now been developed to a high art [41-43], but there are surprisingly few reports concerning LEDs. Many of the problems in this material system relate to poor substrate quality and low thermal conductivity of the IV-VI material. However, the 'mirror-band' structure is less favourable for Auger recombination than in III-Vs and the Auger rate in PbSe has been measured to be almost 100 times lower than that in InSb & related alloys [44].

Shi [45] has shown that there may be some advantages in hybrid III-V/IV-VI LED and laser structures if they can be successfully grown.

2 Limitations to LED Performance

2.1 Internal Quantum Efficiency

It is important to consider the relative importance of the different non-radiative recombination mechanisms with respect to the choice of III-V materials, the conductivity type and low or high injection conditions. Compared with GaAs LEDs which have now been developed to a very high standard, the narrow gap III-Vs such as InAs have longer diffusion lengths (~25μm for electrons), lower radiative efficiency and stronger non-radiative recombination. Because the injected minority carriers are able to travel further within the material there are important consequences. Interface/surface recombination becomes more important and should be avoided, but increasing the active region width leads to increased re-absorption of the emitted light. For low injection Shockley-Read-Hall recombination (SRH) dominates but at higher injection levels Auger recombination becomes more significant. Although the injected carrier density in operational InAs-based LEDs is higher than the background doping ($\sim 10^{16}cm^{-3}$), the situation is very different to that in a laser. LEDs typically operate at relatively low injected carrier densities ($\sim 10^{16}cm^{-3}$) and still provide a useful total output. We conclude that SRH recombination via impurities and defects can be a significant non-radiative recombination path in InAs LEDs at room temperature under typical device operating conditions. So, it is apparent that minimising SRH recombination should be very effective in improving LED performance. Similar arguments also hold for 300K InAs detectors, and certainly for reverse-biased detectors operating in the carrier exclusion mode [46], such that reduction of SRH centres should yield performance near the theoretical limit.

The situation is rather different in InSb however, because the material contains fewer traps such that Auger recombination limits LED efficiency and is even more severe for InSb-based lasers which necessarily operate under high injection conditions. Kane *et al.*, [47], have compared the radiative recombination efficiency of bulk p-type materials by measuring electroluminescence emission from InAs and InAsSb LEDs. Results indicate that GaAs is superior to InAs whose radiative efficiency rolls off around $3x10^{17}cm^{-3}$, while $InAs_{0.9}Sb_{0.1}$ is worse still rolling off at $3x10^{16}cm^{-3}$. This behaviour is consistent with the increasing importance of Auger recombination. At low injection the radiative recombination rate is proportional to np and the non-radiative recombination rate is proportional to p^2n, assuming that CHSH and CHLH Auger processes dominate in p-type material [48]. This yields for the efficiency

$$\eta = [\ 1 + 1/Bp\tau_{nr} + C_p\, p/B\ ]^{-1} \qquad (1)$$

where τ_{nr} is the non-radiative lifetime associated with Shockley-Read processes, B is the radiative rate constant and C_p is the Auger rate constant. Experimental

measurements [49,50] have produced a value of $C_p = 2x10^{-28}$ cm^6s^{-1} for low injection and indicate that original estimates for p-InAs by Takeshima [51] could easily be too large by a factor of 10. A similar consideration of efficiency in n-type III-Vs reveals that Auger recombination is more severe leading to an efficiency roll off near 6×10^{15} cm^{-3} in n-InAs. This is because the CHCC mechanism is some 50 times stronger than the CHSH and CHLH processes. The result is that low p-type (л) material would be the primary choice for the LED active region at low injection. But, as injection increases the CHCC processes become progressively more significant and the efficiency decreases such that InAs LEDs with either low n or p-type active regions show similar efficiency.

As a result, there is much interest in the use of band structure engineering to produce Auger suppression in InAs alloys. The use of W- structures and strained quantum wells to modify band structure and remove holes at large k is finding successful application particularly in lasers where the injection levels are high. However, for LEDs one could argue that the minimisation of SRH recombination is more important and the approach must be to eliminate non-radiative recombination centres by removing unwanted impurities and native lattice defects. This essentially requires that one should further purify the epitaxial material in the LED active region. This has been the approach used in our laboratory to realise improved InAs and InAsSb LEDs and is described in the next sections.

2.2 Purification Using Rare Earth Gettering and Pb Neutral Solvent Epitaxy

InAs has a room temperature energy bandgap of 0.354eV ($\lambda_g = 3.5$ μm) and so is a good choice as a starting point for the active region in the fabrication of LEDs spectrally tuned for methane detection near 3.3μm. However, although there have been a number of investigations into the electrical properties of epitaxial InAs grown from the liquid phase [52], the corresponding studies of the associated optical properties have received little attention. One approach to obtain a reduction of SRH recombination and to improve the quantum efficiency of InAs has been to use different melt purification techniques during LPE growth. These include both rare earth gettering and neutral solvent epitaxy.

The purity and quality of epitaxial InAs depends principally on the number of residual impurities and the number of structure defects. Various methods to prepare device quality InAs epitaxial layers with a low residual carrier density and a high electron mobility have been investigated. For example, purification of the starting materials (InAs and In) by prolonged annealing can reduce the impurity concentration and the corresponding residual electron density by an order of magnitude (from 10^{17} to 10^{16} cm^{-3}). The concentration of volatile impurities which result in 1-3 meV shallow donors can be further reduced by the addition of small quantities of rare earth ions. The common impurities (for example, S, Se, Si, C, Te, O, *etc.*) can readily form stable compounds with reactive rare earth elements such as Gd, Yb, and Er. These compounds are insoluble in indium solution and are not thought to be incorporated into the grown solid epitaxial layer and therefore rare

earth gettering is an attractive means of removing unwanted impurities from LPE grown InAs solid solutions and reducing the background doping level towards 10^{15} cm^{-3} [53].

2.3 Purification of Epitaxial InAs Using Gd Gettering

As mentioned earlier, room temperature quantum efficiency in narrow gap LEDs is low and is thought to be limited by band-to-band Auger recombination. However, the lower the injected carrier concentration the more likely it is that SRH recombination will be significant compared with Auger recombination and particularly if the devices are not actually operating at the Auger limit. Undoped melt grown InAs which is the basic substrate material for detectors and LEDs in the 2-5μm range is naturally n-type and usually contains residual donors of comparably high concentration (> $2x10^{16}$ cm^{-3}). Unintentionally doped InAs alloy layers grown by liquid phase epitaxy are also n-type with a background concentration of around 10^{16}-10^{17} cm^{-3}, whereas lower residual carrier concentrations (~10^{15} cm^{-3}) are obtainable using MBE or MOVPE. However, because LPE growth occurs at thermodynamic equilibrium the crystalline perfection of the resulting material is superior, making LPE material an attractive and inexpensive choice for mid-infrared optoelectronics provided the residual carrier concentration can be reduced. By adding small quantities of Dy or Gd to the melt it has been shown that the background carrier concentration in InGaAs (and InAs) can be reduced to ~ $3x10^{15}cm^{-3}$ [54-57]. This is possible because the rare earths form stable R.E. silicides/sulphides in the melt (*e.g.* DySi or Gd_2S_3) and these compounds are insoluble in indium which effectively reduces the incorporated shallow donor concentration in the solid layer. (Gettering of acceptors also occurs but at a lower rate because the intermetallic compounds Dy_2Zn_3 and Dy_2Mg_3 form less easily). Also it has been shown that rare earth elements themselves do not incorporate readily into the epitaxial layer at concentrations below ~0.8at% [58]. This has important consequences for LEDs (and detectors) based on these materials. Rare earth gettering can be expected to improve the quantum efficiency of InAs by reduction of trap/recombination centres associated with impurities and reduction of background carrier concentration by removal of unintentional donors.

High quality epitaxial InAs has been grown by LPE with the addition of small amounts of the rare earth element Gd as an impurity gettering agent added to the growth melt [59,60]. The InAs(Gd) epitaxial layers exhibited n-type conduction with Gd mole fraction lower than 3×10^{-4}, increasing above this value the epitaxial layers became p-type. When a Gd mole fraction of 1.84×10^{-5} was used then a minimum residual electron concentration of $6\times10^{15}cm^{-3}$ was obtained at liquid nitrogen temperature in good agreement with previous work [61]. The carrier mobility was measured to be $75000cm^2V^{-1}s^{-1}$. The photoluminescence emission intensity also increased by at least a factor of 10, due to a reduction in Shockley-Read-Hall centres. As shown in Figure 2 the purified InAs(Gd) layers exhibited four peaks in the 4K photoluminescence emission spectrum at (A) 411meV, (B) 397meV, (C) 376meV and (D) 367meV respectively Peaks A and B were identified as originating from bound exciton and donor-acceptor transitions respectively [62]

whereas C and D were much weaker and were attributed to defects or impurity related transitions reported previously in the literature [63].

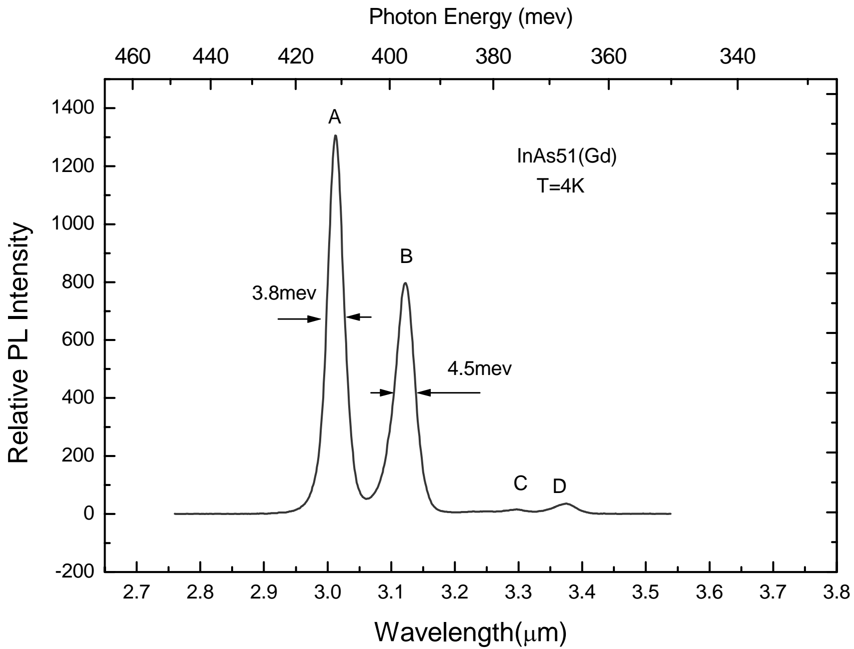


Figure 2. Low temperature photoluminescence spectrum of InAs prepared using Gd gettering of the melt during LPE growth (after Gao *et al.* [60])

The strong increase in photoluminescence intensity and the spectral linewidth narrowing arise due to a reduction in the residual impurity concentration (and corresponding number of non-radiative recombination centres). This is directly attributable to the gettering effects of the (Gd) rare-earth ions which were intentionally added to the melt. Compared with untreated InAs layers, the PL emission intensity of the InAs(Gd) epitaxial layers is much stronger and sharper. The linewidth (FWHM) of the bound exciton peak was measured to be 3.8meV, which is narrower than in much MBE and MOVPE grown InAs epitaxial material [64-66] with carrier concentration lower than $5\times10^{15}cm^{-3}$. These results confirm that both the optical and electrical properties of LPE grown InAs and its III-V alloys can be substantially improved by introducing small amounts of Gd into the growth solutions as a gettering agent and that as a direct consequence there should be corresponding significant device improvements.

2.4 Fabrication of InAsSb LEDs at 4.6μm

Melt purification techniques can be effectively extended and applied to other alloys as well as binary InAs. For example, the InAsSb ternary alloy has been purified using the rare earth ion Yb as the gettering agent and LEDs have been fabricated containing a purified ternary InAsSb active region. The 300K electroluminescence emission spectrum from one of these LEDs is shown in Figure 3, where the absorption from atmospheric CO_2 at 4.2μm is clearly visible and the LED emission has a peak wavelength of 4.6μm.

The LED structure, which has been described in detail previously elsewhere [67], consisted of two cladding layers of n^+ $InAs_{0.65}Sb_{0.15}P_{0.30}$ and p^+ $InAs_{0.65}Sb_{0.15}P_{0.30}$ and an undoped active region of $InAs_{0.89}Sb_{0.11}$ sandwiched between them. Because the $InAs_{0.89}Sb_{0.11}$ ternary active region material has a large lattice mismatch with respect to the InAs substrate layer, a buffer layer with an intermediate composition ($InAs_{0.94}Sb_{0.06}$) having a 0.41% positive mismatch to the InAs substrate layer was introduced in the structure between the substrate and the n-type $InAs_{0.65}Sb_{0.15}P_{0.30}$ layer to reduce dislocations and to relieve strain caused by the lattice mismatch.

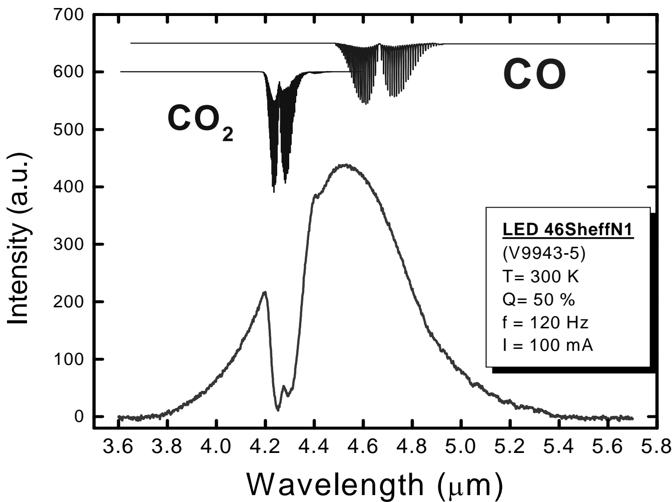


Figure 3. The electroluminescence emission spectrum of an $InAs_{0.89}Sb_{0.11}$ LED at 300 K

The growth melts for the InAsSb alloy epitaxy were prepared in a similar manner to those for the growth of InAs above. However, in order to investigate the effect of material purity directly on the quantum efficiency of the active region, four different types of devices were fabricated in which the same epitaxial growth was used for each, but the nature of the antimony source material was different. The pre-

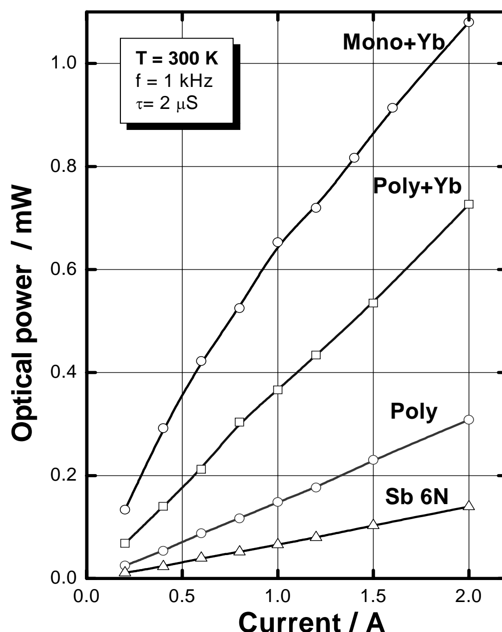


Figure 4. The optical output power characteristics measured from 4.6 μm LEDs prepared with different purification treatments of the InAsSb active region (after Krier *et al.* [69])

cursors used in the growth melt of the active region were as follows: 6Ns Sb metal, polycrystalline InSb, polycrystalline InSb +Yb and monocrystal InSb+Yb. In this experiment, we expect the improvement in precursor material purity (in going from metal to polycrystalline to monocrystal material) to result in increasingly pure epitaxial layers of InAsSb. The purpose of the Yb was to act as a gettering agent in the growth solution in order to remove S, Si impurities and to further improve the quantum efficiency of the layer. Figure 4 shows the optical output power from each of these four types of devices. (The current used for the comparison was 2μs pulse duration at a frequency of 1kHz.) From these experimental curves, it can be seen that as the purity of the precursors used in the preparation of the active region goes up, the output power exhibits a large improvement. In fact, the output power of the device grown from monocrystal InSb and Yb was eight times larger than that grown from Sb metal at 2A drive current. This is clear evidence that improving the purity of the precursor materials used in the LPE growth melt results in a reduction of the residual carrier concentration and defects which reduce Shockley-Read-Hall recombination (SRH) in the active region. This is entirely consistent with the improvements obtained in the PL investigation of InAs discussed above, where the photoluminescence intensity of Gd-treated InAs was improved by 10–100 times compared to the undoped epitaxial InAs at low temperature. The results also indicate that SRH recombination is more significant than Auger recombination in

InAsSb LEDs at room temperature. At first sight this is inconsistent with other reports concerning mid-infrared lasers in the literature which are predominantly Auger limited near room temperature. However, the injected carrier concentration in an operational LED is much lower ($<10^{16}$ cm^{-3}) than in a laser ($>10^{18}$ cm^{-3}) and therefore SRH recombination (which is proportional to n and not n^3) can become the dominant non-radiative mechanism. Consequently, reducing the concentration of SRH centres using rare earth gettering can be used to major advantage in LEDs.

2.5 InAs LEDs for Methane Detection at 3.3μm

The rare earth gettering technique can also be applied to produce powerful InAs LEDs by following the same procedures used for InAsSb above. However there are some important differences that arise due to the Type II nature of the InAs/InAsSbP heterointerface of the DH. LEDs have been fabricated in which the unintentionally doped n-InAs active layer is enclosed between p- and n- InAsSbP confinement layers. The P content in the confinement layers was 0.40 (Eg = 570 meV, T = 300 K) to provide a higher bandgap energy and large interface band offsets for good electron confinement. The InAs active region was 0.7μm thick (Eg = 354 meV, T = 300 K), and the InAsSbP layers were iso-periodic with InAs and each 3.0μm in thickness. The cladding layers were doped with Sn up to a concentration of $5x10^{18}$ cm^{-3} and with Zn up to $1x10^{18}$ cm^{-3} for n- and p- sides respectively. By using Yb rare earth gettering the residual carrier concentration in the active layer was reduced to $< 5x10^{15}$ cm^{-3}.

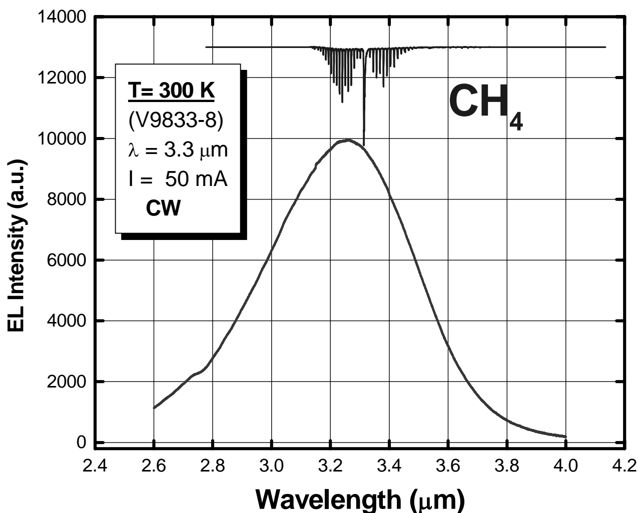


Figure 5. The electroluminescence emission spectrum of an InAs double heterostructure LED at 300 K

The quasi-continuous electroluminescence emission spectrum of one such InAs LED at room temperature is shown in Figure 5. A single peak is obtained at 3.28μm which is coincident with the Q-branch absorption of methane gas. Furthermore, the LED peak wavelength remains constant over the whole range of drive currents used, unlike the situation in bulk InAs homojunction LEDs. The full width at half maximum (FWHM) of the spectrum increases only slightly with drive current and was measured to be 500 nm for these 3.28μm LEDs. The InAs/InAsSbP heterostructure is Type II in nature and although the structure contains no quantum wells, the band bending at the interface leads to some confinement of electrons and holes on opposite sides, which affords the possibility of reduced Auger recombination and increased electroluminescence at room temperature. The peak wavelength of 3.28μm is not the same as that expected from the InAs energy gap, *i.e.* $\lambda_g = 3.50\mu m$ at 300K. The emission is also observed to be more intense than in similar epitaxial homo-junction InAs LEDs. Electroluminescence at higher energy than the bandgap (*i.e.* 3.28μm) and which is insensitive to drive current is consistent with emission originating from Type II recombination at the InAs/InAsSbP heterojunction interface.

We can account for the improvements in output power in these LEDs by simply considering the carrier recombination mechanisms which are operating in the diode. At equilibrium, the total recombination rate, R_{tot} per unit area is given by [68]

$$R_{tot} = e\,(A_{SRH}n + B_{Rad}n^2 + C_{Auger}n^3)\,d \tag{2}$$

where n is the carrier density in the active layer, e is the electronic charge, and d is the thickness of the active layer. The total recombination rate R_{tot} includes band-to-trap non-radiative recombination $A_{SRH}.n$, band-to-band radiative recombination $B_{Rad}.n^2$, and Auger recombination $C_{Auger}\,n^3$. The coefficient C has an exponential temperature dependence and takes the following form:

$$C_{Auger} = Co \exp^{-Ea/kT}, \quad \text{where} \quad E_a = [m_e/(2m_{hh})]E_g \tag{3}$$

where E_a is the activation energy for the Auger process (and m_e and m_{hh} are the respective electron and heavy hole effective masses).

The total output power has been increased by minimising the non-radiative recombination in two ways. First, because the dominant radiative transition occurs across the interface and is Type II in nature, since the electrons and holes are localised on opposite sides of the heterojunction, then Auger recombination is directly suppressed and the coefficient C is reduced in the above equation. Because these levels are localised there is little wavelength shift with increasing drive current. Second, using Yb ions as gettering agents for the melt purification during LPE growth of the LED active region reduces the number of non-radiative recombination centres and so directly reduces the coefficient A. Because reducing the number of such centres and shallow donors also reduces the residual carrier concentration this indirectly reduces the last term $C_{Auger}n^3$ [69]. The relative importance of the two terms $A_{SRH}.n$, and $C_{Auger}\,n^3$ depends on the temperature and also on the injection level in the device when it is operating. In this respect the

$A_{SRH}.n$ term is thought to dominate at low temperatures, while $C_{Auger}\,n^3$ dominates at room temperature. This is certainly the case under high injection conditions. But, for LEDs where the injection level is lower (typically < $5x10^{16}cm^{-3}$), SRH recombination is also important, which is why an improvement due to the Yb gettering is obtained at 300K.

2.6 Neutral Solvent Epitaxy

The LPE growth of InAs from In-rich solutions has been investigated and it has been established that because epitaxy occurs below the binary melting point, As deficiency results in the formation of arsenic vacancies V_{As} in the solid, which can lead to structure defects or complex defects (V_{As}-impurity) which manifest themselves as donor or acceptor levels in the InAs band gap. There is evidence that arsenic vacancies (V_{As}) produce donor levels 43 meV below the conduction band and that complex defects result in deep acceptors [70]. Instead of using indium solution, the use of lead as a neutral solvent during growth of n-type InAs by liquid phase epitaxy makes it possible to adjust the ratio of In to As in the molten solution by varying the concentration of Pb and thus to prepare materials with reduced concentrations of structure defects. Electrical measurements on such samples have

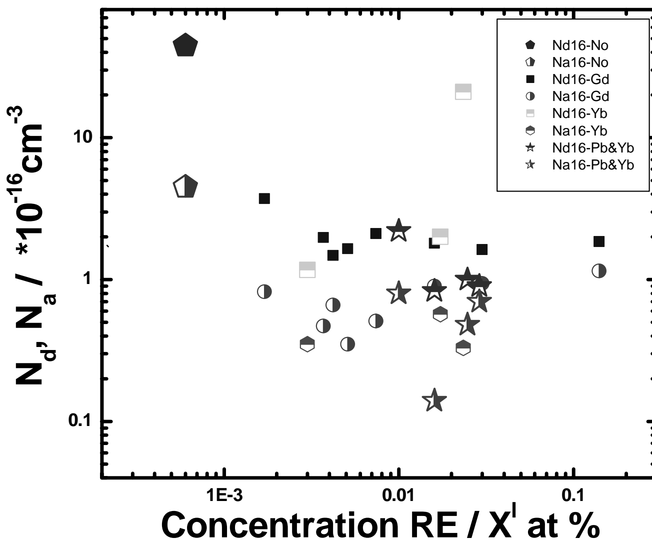


Figure 6. The dependence of residual impurity concentration on the rare earth gettering or melt purification methods used

so far confirmed that epitaxial InAs with a low residual carrier density and a high mobility (n = 6x10^{15} cm^{-3} and μ_{77} ~54000 cm^{2}V^{-1}s^{-1}) can be obtained at a growth temperature of ~650ºC using lead concentration of ~63 mole% in the molten solution.

Figure 6 shows the dependence of the residual impurity concentration of the active region in response to the different melt purification techniques, based on Hall effect measurements. The results indicate that electrically the better material was obtained when Pb neutral solvent solution was used together with an additional amount of Yb. Using In solution and Yb alone produced superior results compared with Gd, which gave some improvements over the untreated samples, although the results were not entirely reproducible due in part to the small amounts of Gd or Yb being added. In order to investigate the effect of material purification on the quantum efficiency of the active region, a number of lattice-matched InAsSbP/InAs/InAsSbP double heterojunction devices were fabricated in which the same epitaxial growth procedure was used for each, except for the production of the active region. Devices with InAs active regions grown from In-rich solutions in the normal way were compared with others in which the InAs was grown using a neutral solvent - *i.e.* Pb as indicated above. In other experiments small quantities of either Yb or Gd were added to the In-rich growth solution for comparison with rare earth gettering effects. The mole fractions in the growth solution were typically X^l_{In} = 0.918 and X^l_{As} =

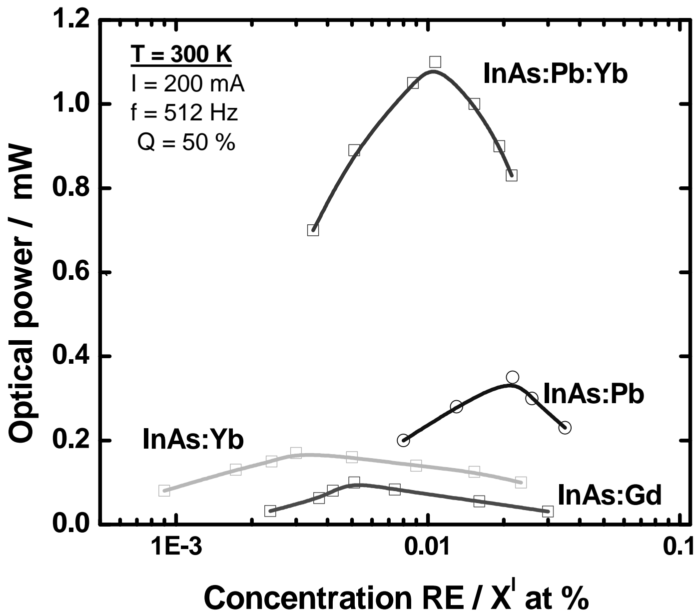


Figure 7. The dependence of the LED output power on concentration for the different purification methods used

0.082 ($X^{l}_{R.E}$ = 0.001–0.02 mol %) respectively, corresponding to a liquidus temperature of ~610 °C. The LEDs containing InAs grown from Pb neutral solvent solution were grown using the mole fractions X^{l}_{Pb} = 0.664, X^{l}_{In} = 0.318 and X^{l}_{As} = 0.018 which gave a liquidus temperature of 610 °C. These values correspond to those which gave the highest values for 77 K mobility and minimum residual electron concentration in the investigation of electrical and transport properties carried out in previous work [52]. In all cases LPE growth was carried out using the step cooled technique with a supercooling of 7ºC.

Figure 7 shows the results obtained from a comparison of the LEDs fabricated from samples with the corresponding active region treatments. A similar trend in the electroluminescence efficiency as found in the electrical results is observed, with the Pb neutral solvent solution and Yb together giving by far the best performance. We conclude that using the Pb neutral solvent solution to adjust the stoichiometry and reduce the As vacancies can be more effective than using Gd or Yb gettering to remove residual impurities. But, clearly the addition of Yb to the Pb neutral solvent gives a further improvement.

Figure 8 shows the light-current density characteristics measured from LEDs prepared with different purification treatments of the active region and using samples chosen from Figure 7 having the optimum rare earth concentrations in each case [71]. The LEDs show excellent output with saturation occurring only at higher current densities. These results are re-plotted in Figure 9 where the slope z

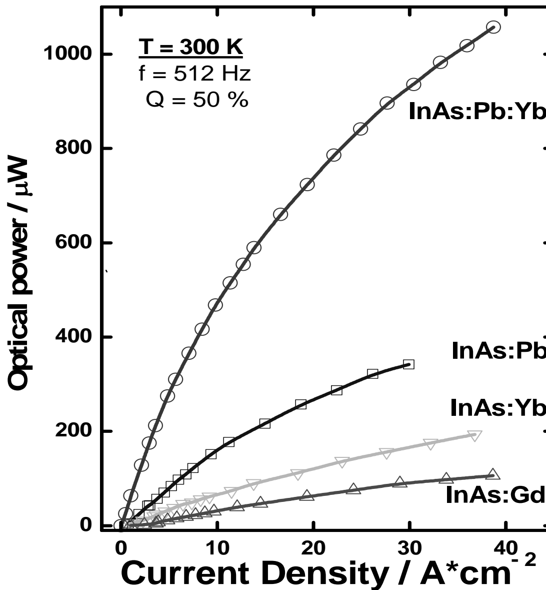


Figure 8. Light-current density characteristics measured from LEDs prepared with different purification treatments of the active region, *circles* - InAs: Pb: Yb, *squares* - InAs: Pb, *triangles down* - InAs: Yb, *triangles up* - InAs: Gd

identifies the dominant recombination mechanism. In the low current range SRH recombination dominates and the slopes are all ~1, followed by radiative recombination in the mid current region until Auger recombination begins at higher injected carrier concentrations. Clearly the InAs:Pb LED appears best throughout the entire range due to the higher purity and structural perfection of the active region.

Although Auger recombination is thought to be a limitation to the operation of room temperature mid-infrared sources these results indicate that the elimination of Shockley-Read-Hall recombination centres is important at low injection. Auger recombination is a three carrier process and the overall rate depends on n as ~ Cn^3 where C is the Auger coefficient. Consequently, to avoid Auger recombination it is necessary to produce as much radiative output as possible by operating the LED at low injected carrier concentration. This in turn means that the residual carrier concentration must be kept low and therefore it is essential to remove un-intentional impurities and native lattice defects. It would appear that the use of Pb as a neutral solvent is highly effective in this respect. By using Pb neutral solvent together with Yb in p-i-n heterojunction LEDs it is possible to achieve an increase in 300 K output power of about a factor of 10.

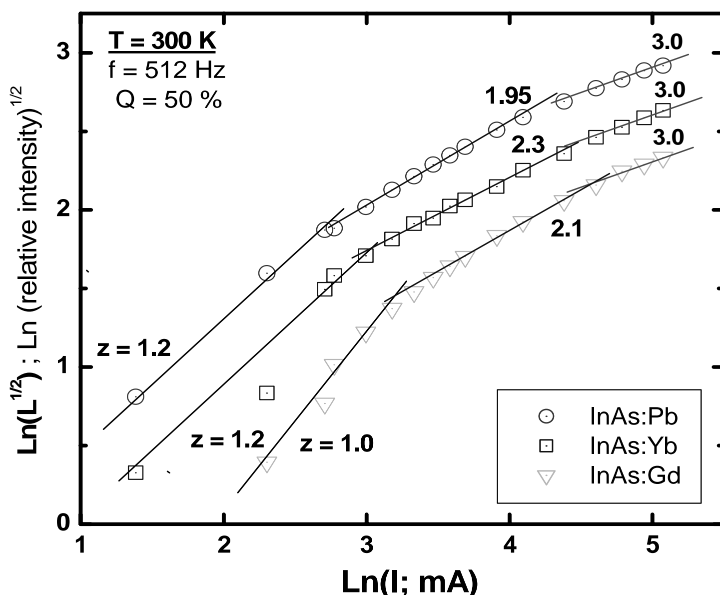


Figure 9. The relative electroluminescence intensity (L) – current (I) characteristics measured from the various LEDs plotted as ln ($L^{1/2}$) against ln (I): *circles* – InAs: Pb, *squares* - InAs: Yb, *triangles* - InAs: Gd. The *slope* represents the index z in the equation $I = An+Bn^2+Cn^3$ and reveals the different recombination mechanisms operating as the current injection increases [71].

3 High-pressure Measurements

Hydrostatic pressure has been extensively used to study the recombination mechanisms that occur in semiconductor LEDs and laser structures [72]. But, although the technique has been widely used to study visible and near infrared emitters, its application to sources operating at wavelengths beyond 2 μm has only recently been investigated. The application of hydrostatic pressure can give valuable information about the relative significance of the different recombination mechanisms that are sensitive to changes in the bandgap. To obtain information about the relative importance of SRH and Auger recombination in the mid-infrared, the behaviour of 3.3 μm and 3.8 μm LEDs, which had InAs or $InAs_{0.94}Sb_{0.06}$ in the active region, (giving Type II and Type I band line-ups respectively) has been investigated [72]. The devices were operated using a low duty cycle (typically 5%) in order to avoid Joule heating effects. The measurements were made at room temperature using a He hydrostatic pressure cell and 77 K InSb photodetector. Figure 10 shows a plot of the integrated spontaneous emission, corrected for detector response, against current in the pressure range 0 to 9.4 kbar, for one of the 3.3 μm LED samples after Yb purification of the active region. A significant increase in radiative efficiency of x4 was observed for these Type II diodes. Since

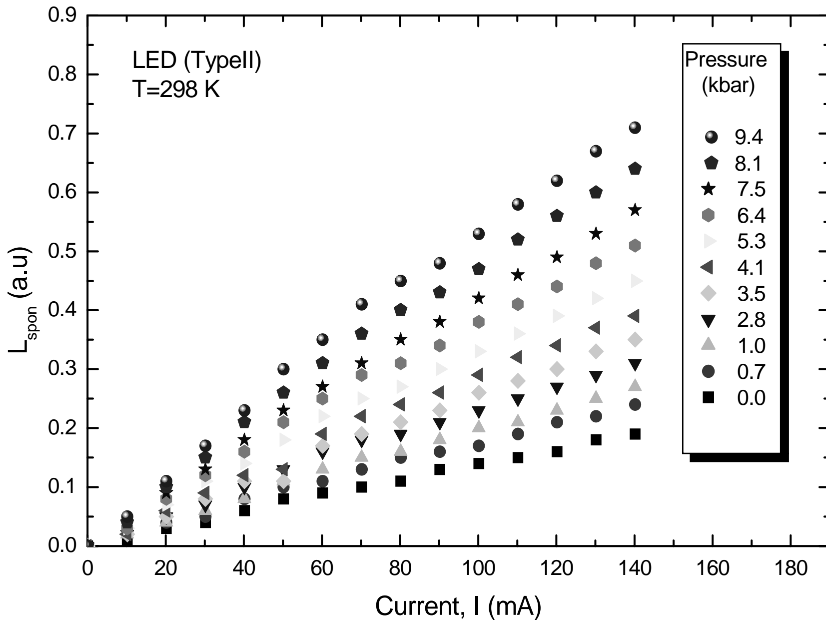


Figure 10. Measured integrated spontaneous emission intensity as a function of current for 0 kbar$<P<$9.4 kbar at room temperature for an InAs/InAsSbP 3.3 μm LED [72]

the band offsets and the capture cross-section for SRH are almost independent of pressure it is clear that the increase in electroluminescence is associated with a reduction in Auger scattering.

One can account for the improvement in light output with hydrostatic pressure in these LEDs by considering the effect of pressure on the CHCC and CHSH Auger recombination processes. Simple analytical expressions can help to understand the variation of these processes with band-gap. The band-to-band Auger process can be expressed as a thermally activated process given by [73]

$$C = C^0 \exp\left(\frac{-\Delta E_\alpha}{\kappa T}\right) \Leftrightarrow \ln C = \ln C^0 - \frac{\Delta E_\alpha}{\kappa T}. \quad (4)$$

where C^0 is approximately independent of temperature and pressure, but includes the terms due to the coulomb interaction matrix between initial and final Auger states. The main pressure dependence of the Auger activation energy (ΔE_α) for the direct Auger recombination processes CHCC and CHSH is due to the bandgap dependence of ΔE_α and is given by [74]

$$\Delta E_\alpha(CHCC) = \frac{m_e E_g}{(m_e + m_h)}, \quad \Delta E_\alpha(CHSH) = \frac{m_{so}(E_g - \Delta_{so})}{(2m_h + m_e - m_{so})}, \quad (5)$$

where m_e ($0.023m_o$), m_h ($0.41m_o$) and m_{so} ($0.16m_o$) are the electron, hole and spin-orbit effective mass respectively (the values within the parenthesis are the effective masses of InAs at ambient conditions). Electron and spin-orbit effective masses, increase approximately in proportion with increasing pressure whilst the hole effective mass is relatively independent of pressure [72]. Δ_{so} is the spin-orbit split-off energy. It can be seen that the Auger coefficient reduces in significance with increasing E_g and hence with increasing pressure ($\frac{\partial E_g}{\partial P}$ for InAs is approximately 10 meV/kbar[75]). For longer wavelength devices the phonon-assisted Auger processes have activation energies much larger than the phonon energy and they are far less sensitive to the band structure and temperature compared to the band-to-band case[72]. Although in the InAs active region, the band gap and the spin–orbit band gap are nearly equal, experimentally no resonant behaviour in the light output as a function of applied hydrostatic pressure was observed.

Using this simplified model the bandgap dependence of ΔE_α and the bandgap dependence for the direct Auger recombination processes CHCC and CHSH can be calculated. Although both processes are reduced as a function of pressure the relative change for the CHSH process (a factor of 2.55 (2.55/3.74=68 %)) is much closer to the experimental change in light output (a factor of 3.74) than the change due to the CHCC process (a factor of 1.21 (1.21/3.74=32 %)). Note that the total change due to CHSC+CHCC (a factor 3.76) is in excellent agreement with the experimental change in light output (a factor of 3.74). This suggests that CHSH is the dominant process in these devices from ambient up to the pressure range studied here. (According to these calculations the activation energies for the two processes became comparable at the highest pressure around 9.4 kbar). Note that inter-valence band absorption (IVBA) would also be reduced by de-tuning E_g above Δ_{so}, but this does not affect these measurements which were made viewing through the n-type

side of the sample. It is worth mentioning that if radiative recombination dominates the device response, the output would reduce with increasing pressure due to reduction of the matrix element with increasing bandgap.

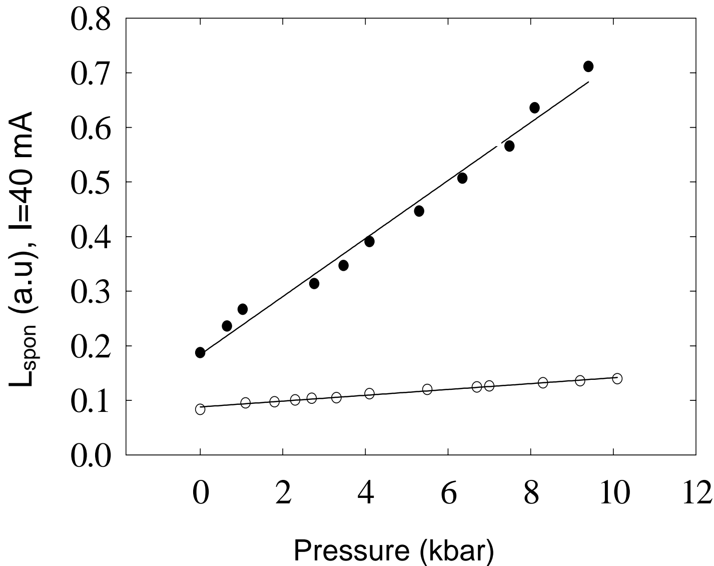


Figure 11. A comparison of the pressure dependence of the electroluminescence emission intensity for Type-I (3.8 μm, *open points*), and Type-II (3.3 μm, *solid points*) double heterostructure LEDs at 300 K. Reprinted with permission Copyright 2003 AIP [76].

Figure 11 shows a plot of the light output at constant current (140 mA) as a function of applied hydrostatic pressure for the two samples together [76]. A factor of 1.7 increase in the light output for the Type I, 3.8 μm LED (open points), and a factor of 3.96 for the Type II, 3.3 μm LED (solid points) in the pressure range up to 10 kbar was obtained. There is clearly a much larger increase in electroluminescence from the Type II LED compared with the Type I LED. The 3.7 increase in the radiative efficiency of the Type II LED compared to 1.7 in the case of the Type I is an experimental indication that Auger processes are more important for the Type II LED compared to the Type I LED. However, it is well known that the Type II structure reduces the overlap between electron and hole wavefunctions since the carriers are delocalized. This should result in a reduction of the Auger coefficient *C* in Type II structures compared to Type I. But, these results indicate stronger Auger suppression with applied pressure for Type II rather than Type I structures. At first this seems anomalous, but the effective width of the recombination zone within the active region of the Type II LED, (*d*eff) is significantly smaller, since the Type II band edge alignment results in band bending and the injected carriers become localized near the interface. Therefore, the value of *d*eff is essentially determined by the carrier tunnelling distances, estimated as *d*eff

~150 Å. We suppose that, although Type II recombination across the InAs/InAsSbP interface reduced the Auger coefficient *C*, the injected carrier concentration is actually greater than that found in the Type I LED due to the small recombination volume localized near the interface. As Auger processes are strongly dependent on injected carrier concentration (~increase as n^3), the net result is an overall increase in the Auger rate (Cn^3). We may then interpret the larger increase in the light output as a function of pressure for the Type-II LED compared to the Type I LED as being associated with Auger suppression. In type II quantum wells the radiative and Auger recombination coefficients depend on the band gap as $B \sim Eg^2$ and $C \sim Eg^{-5.5}$ respectively, and at zero applied pressure, $B \sim 10^{-10}$ cm^3s^{-1} and $C \sim 8 \times 10^{-27}$ cm^6s^{-1}. This gives the radiative contribution to the current density as 14% at P=0 and 40% at P =9.4 kbar. Simple estimates predict an increase of about 2.85 in light output when the pressure is increased from 0 to 9.4 kbar at constant current (I=140 mA). This is approximately in agreement with the value of 3.74 obtained experimentally.

4 Optical Extraction

In addition to the problems associated with radiative efficiency, mid-infrared LEDs also suffer from the problem of low optical extraction efficiency which is common to most LEDs as a result of the high refractive index (n~3.5) of the semiconductor material. This results in a small critical angle (~15°) and a narrow escape cone such that the ratio of external to internal efficiency is $1/n(n+1)^2 \sim 1/70$. The LED structure employed must ensure that the majority of generated photons meet the emission surface within the critical angle and that the that these photons are transmitted to the emitting surface within a minimal path length to prevent absorption losses. In contrast to the work on improving internal radiative efficiency there has been surprisingly little research into solving this fundamental problem. The principal approaches have been to use immersion lenses, optical concentrators [7,78], surface texturing [79] or resonant cavity structures [80] to extract more light from the LED.

Although Bragg mirror technology has been extensively implemented in visible and near-IR resonant cavity LEDs (RC-LED)[81] and vertical cavity surface-emitting lasers (VCSELs), the technology has been slow to transfer out to the longer wavelengths. In general RC-LEDs exploit the properties of planar micro-cavities to redirect a larger fraction of the spontaneous emission into the escape cone, thereby increasing the extraction efficiency and improving the spectral performance. Whereas VCSELs need highly reflective mirrors (*i.e.* >99%) the requirements for LED enhancement are much less stringent. In a resonant cavity LED (RCLED) the micro-cavity essentially clamps the linewidth and emission wavelength with respect to the LED drive current. The cavity mode and the semiconductor bandgap have opposite temperature sensitivity resulting in a much more stable emission wavelength. This is an important consideration for a practical gas sensor where the LED is tuned to the target gas absorption resonance and should remain fixed over the ambient operating range (-40ºC to +50ºC). RCLEDs are somewhat easier to work with than VCSELs because they need only a small Bragg period number (5-6) rather than 20-30 typically required for VCSELs. These LEDs can in principle

provide some of the most important advantages of the VCSEL but without the corresponding stringent requirements on epitaxy. Potential performance advantages include; increased optical extraction (higher output power), improved spectral purity (narrow linewidth), directed emission and improved temperature stability. Modelling by De Neve *et al.* [82] has shown that it is possible to obtain high external quantum efficiencies using low reflectivity mirrors. Some mid-infrared resonant cavity light emitting diodes (RCLEDs) based on CdHgTe alloys and grown by MBE have now been reported [80]. These devices contain a single Bragg mirror reflector and emit at 3.18 μm and 3.35μm. The reported spectral bandwidth (FWHM) is less than half that of the bulk active layer emission bandwidth and the device also has improved beam divergence characteristics. Application to III-V systems has been slow mainly due to epitaxial growth problems associated either with mismatch between DBR layers or between active region and DBR layers. However, powerful (>20mW CW) VCSELs operating at 2.2μm have been demonstrated [83].

InSb optical concentrators have been modelled and applied in a "Winston cone" configuration to effectively make the volume of the emitting material smaller while keeping the emission constant. Micro machining of a transparent n^+ - InSb substrate using sequential reactive ion etching with CH_4 : H_2 and O_2 was employed to increase the output of InSb/InAlSb LEDs [78]. Optical gains of 3.5 over conventional mesas were obtained from the resulting large area ($3cm^2$) LED arrays. In the case of near IR GaAs LEDs, the problem is alleviated by depositing a lens shaped epoxy resin to aid index matching and increase the escape cone. Unfortunately this is inefficient for mid-IR LEDs since most inert epoxy resins have strong absorption in the 2-5μm region and fabrication of such index matching lenses in silicon is impractical. Silicon immersion micro lenses have been used for detector arrays but with limited effect due to contact and alignment problems [84]. This technique has however been successfully implemented for HgCdTe LEDs using a CdZnTe coupling lens, such that 20% of the available light could be extracted [38]. Matveev *et al.* [85] used a similar approach for InAsSb, but employed selective chemical etching to produce a device with a 30μm deep inverted mesa and a silicon sub-mount to produce an emitting surface free of contact pads and bond wires, giving a peak output of 60μW at 5.5μm using 1A pulses (1.5% duty cycle).

Another potentially useful approach for mid-infrared optoelectronic devices is an InAs micro lens, integrated onto the radiation emitting chip. The LED is fabricated and mounted in the epi-side down configuration with the lens etched into the substrate on the opposite side. To achieve the necessary diameter the micro lens production process involves first etching a ring into the substrate and then shaping the surface using an isotropic etchant. The final shape is strongly affected by the original ring diameter and depth. Using a ring diameter of 650μm and a deep groove, it is possible to produce an integral lens as shown in Figure 12. Optical modeling using ray-tracing software has shown that even though the refractive index is high, the lens needs to be quite thick before it has a significant effect on extraction. The usual 500 μm substrate thickness of readily available InAs effectively places an upper limit on this and reduces the benefits of these lenses. Further increases can be obtained by defining a small spot for the emission region and etching a mesa to act as a "corner reflector". However, in practice this has so far

not resulted in more than a factor of 2 improvement in extraction efficiency compared with a conventional surface emitting mesa.

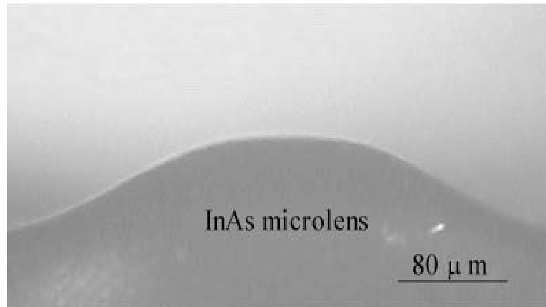


Figure 12. Cross-section of an integral micro lens etched into the substrate on the back of an InAs LED

5 Comparison of Devices

Comparing the performance of LEDs at different wavelengths in the 2-5 μm spectral range reveals that the output power and associated quantum efficiency decreases with increasing wavelength. This observation is generally consistent with a decreasing spontaneous emission rate which decreases as the alloy bandgap decreases. The temperature dependence of emission intensity is determined by the competing non-radiative recombination processes, and in InAs alloys primarily through the Auger mechanisms [86,87] CHCC and CHSH, which have a stronger (exponential) temperature dependence than the spontaneous recombination coefficient. A plot of the temperature quenching of the electroluminescence emission from different LEDs made in our laboratory and normalised to 1.0 at 77 K is shown in Figure 13.

The devices all showed generally similar behaviour, but the 3.8μm LED exhibited the least severe reduction in electroluminescence on approaching room temperature, while the 3.3μm and 4.6μm LEDs were almost the same. We may estimate the Auger coefficient in the active region of these LEDs and its temperature dependence by fitting our experimental data. Because in each case the active region has been purified, we are able to make the assumption that non-radiative band-to-trap recombination can be neglected, to a first approximation. The temperature dependence of B is known to be of the form $B(T)=B_0(T_r/T)$ [88] and T_r was set to 30K, with $B_0 = 2\times10^{-10}$ cm^{3}s^{-1}. The temperature dependence of C was assumed to be exponential and of the form $C(T)=C_0\exp(-E_\alpha/kT)$ [89], where E_α is the activation energy for Auger recombination and C_0 is the Auger coefficient. A reasonably good fit to the data was obtained when $C_0=1.5\times10^{-26}$ cm^{6}s^{-1} and E_α=31 meV, implying that $C = 1.35\times10^{-28}$ cm^{6}s^{-1} at 77K and 4.5×10^{-27} cm^{6}s^{-1} at 300K respectively. The values of Auger coefficient, $C_0=1.5\times10^{-26}$ cm^{6}s^{-1} and activation energy, E_α=31meV are generally in good agreement with those obtained previously by other workers [49,86].

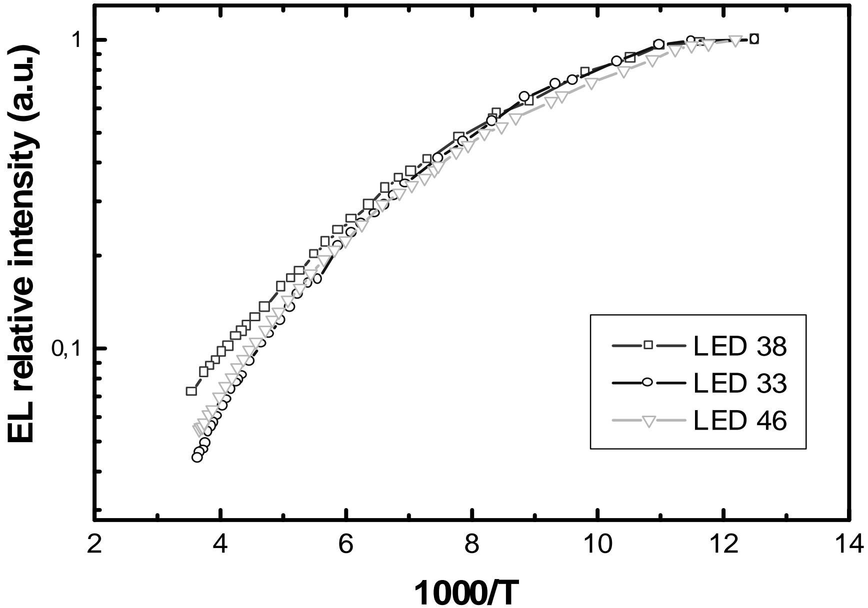


Figure 13. A comparison of the normalised temperature dependence of electroluminescence emission intensity of 3.3, 3.8 and 4.6 μm LEDs over the temperature range 77-300 K (after Krier *et al.,* [23])

In the 3.8μm LED the bandgap energy (E_g = 0.314eV) is less than the split-off band gap energy (Δ_{so} = 0.347eV) at room temperature and the detuning may reduce non-radiative CHSH Auger recombination compared with the 3.3μm LED in which there is a near resonance. However, the 3.3μm LED is overall brighter because of the higher spontaneous emission rate and because the recombination is Type II across the InAs/InAsSbP interface as described earlier. We also consider that the Type II nature of the recombination in the 3.3μm LED is producing some Auger suppression. In the 3.8μm LED, which has a Type I heterointerface, the temperature at which $E_g = \Delta_{so}$ resonance occurs is estimated to be near 160K. But as is evident from Figure 13, we failed to observe any pronounced reduction in EL emission intensity at this temperature in our experiment. This indicates that CHSH may not be the dominant non-radiative recombination mechanism in these LEDs and CHCC may be more important instead, which is consistent with recombination in an n-type active region. Indeed it may be that CHCC alone is responsible for the temperature quenching which we observed in each of the LEDs, (since they all have n-type active regions). In the 4.6μm LED there is an even larger difference between E_g (= 0.267eV) and Δ_{so} (= 0.311eV), yet the output power is lower than for the 3.8μm LED and in this case we attribute this partly to the higher net strain (~0.5%) present in the active region, which is also type I. This latter LED also has a strain-relaxed buffer layer to help accommodate the total strain of ~1% from the InAs substrate

and although partly effective we consider that the 4.6μm LED has the highest concentration of Shockley-Read non-radiative recombination centres of the devices which we have fabricated to date.

6 InAsSb Quantum Dot Light Emitting Diodes Grown by Liquid Phase Epitaxy

There is currently much interest in the epitaxial growth and optical properties of quantum dots (QDs) for optoelectronic device applications because the zero dimensional properties of QDs could provide a means to suppress Auger recombination. In recent years, InSb, GaSb, AlSb and InAs QDs have been grown on GaSb and GaAs substrates by molecular beam epitaxy (MBE) or metal-organic vapour phase epitaxy (MOVPE). The resulting photoluminescence (PL) and electroluminescence (EL) from these QDs were however observed in the near infrared [90-98]. By comparison there have been very few, reports of InSb or $InAs_{1-x}Sb_x$ QDs exhibiting emission in the spectral range beyond 2 μm. Recently, photoluminescence from $InAs_{1-x}Sb_x$ QD in the mid-infrared near 4 μm has been obtained [99] leading to the first observation of EL from $InAs_{0.75}Sb_{0.25}$ QD light emitting diodes. The samples were grown from the liquid phase using a modified (rapid slider) liquid phase epitaxy technique, as described previously elsewhere, and which is capable of producing quantum wells (QWs) and QDs [99]. Essentially a linear motor is employed to rapidly wipe a thin slit of supersaturated melt over the substrate resulting in melt-substrate contact times as short as 0.5 ms. Under lattice matched growth conditions, continuous ultra thin layers or QWs are obtained [93], but when highly mismatched, self-assembled QDs are produced instead.

Figure 14 shows an AFM image of $InAs_{1-x}Sb_x$ QD with x~0.25 grown sandwiched between undoped (n-type) InAs encapsulation layers which were each 1 μm in thickness. The $InAs_{1-x}Sb_x$ QD layer was grown using a 0.6 ms contact time and 20 ^oC melt supercooling. After growth of the active region the sample was removed from the reactor and a PL characterisation of the encapsulated dots was made. The sample was reloaded and a 1 μm thick p^+ (Zn doped $2x10^{18}$ cm^{-3}) contact layer was re-grown at 455 ^oC. The sample was then processed into 375 μm diameter mesa-etched LEDs for EL measurements. A pulse current with 50 % duty cycle at 315 Hz was used to excite EL in the LED.

Figure 15 shows the 4 K EL spectrum obtained from an InAsSb QD LED using 250 mA injection current. The broad peak in the region near 4 μm is due to QD related transitions, the large FWHM being related to the inhomogeneous broadening from their size distribution. By using Gaussian line shapes for fitting, 3 peaks (P6, P7, P8) can be de-convoluted from the spectrum, which are attributed to recombination involving confined hole states in the QD [100]. The peaks P1(2.98 μm), P2(3.08μm) and P4(3.24μm) can be assigned to characteristic recombination in the InAs encapsulation based on previously observed results in the literature[13,14].

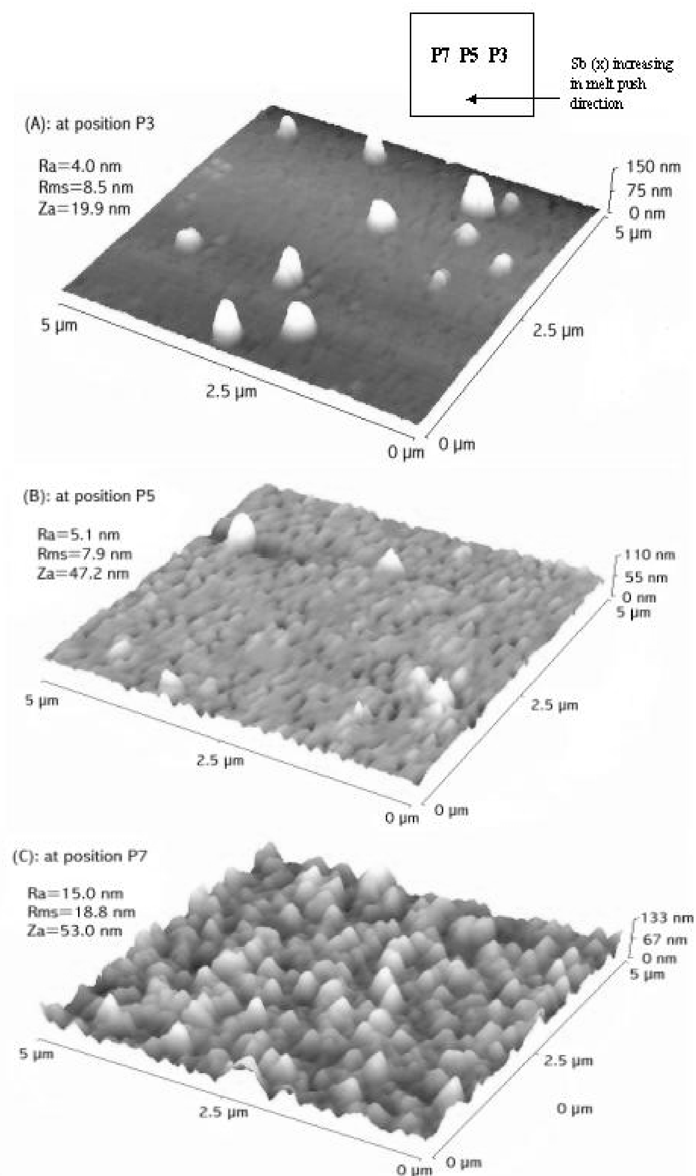


Figure 14. The AFM images obtained from un-encapsulated $InAs_{1-x}Sb_x$ quantum dots with increasing Sb composition grown by LPE and corresponding to the passing direction of the melt across the substrate. Approximate compositions are A (x=0.20); B (x=0.25); and C (x=0.30) Reprinted with permission. Copyright 2000, [99] American Institute of Physics.

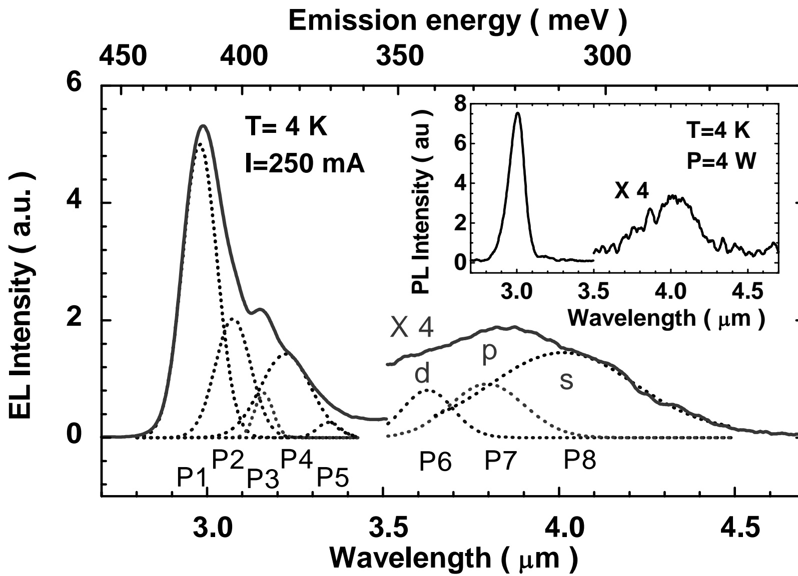


Figure 15. The electroluminescence emission spectrum obtained from an InAsSb quantum dot LED measured at 4 K using 250 mA injection current. De-convolution reveals the s,p,d confined state transitions of the QD. The *inset* shows the PL spectrum at 4 K for comparison. Reprinted from A. Krier, X-L Huang, Physica E, 15, 159 (2002) with permission from Elsevier [100]

A schematic energy diagram of the type II $InAs_{1-x}Sb_x$ /InAs QD system is shown in Figure 16. In bulk or epitaxial InAs, transitions involving a shallow donor level E_S and acceptors, E_{A1}, E_{A2} are normally observed in PL spectroscopy [101-105]. Additional deep states are sometimes also observed in non-stoichiometric material or in samples containing strain. The broad band can readily be assigned to transitions originating within the quantum dot in good agreement with PL spectroscopy [99], as shown in the inset of Figure 15, where the 4 K PL peak related to the $InAs_{0.75}Sb_{0.25}$ QD in the active region of the sample was also centred at 4 μm with a full width at half maximum (FWHM) of 400nm. We note that the QD emission is wider in EL than in PL due to the higher carrier injection level, which suggests that more QD sub-levels are populated. As the injection level is increased, Pauli blocking causes a saturation of the carrier number in lower energy states inside the QD and the relaxation rate from higher states decreases leading to observable emission from the excited levels. Two additional peaks P3 (3.16 μm) and P5 (3.33 μm) are associated with a thin wetting layer or a strained InAs layer covering the InAsSb QD. These peaks are more prominent in the EL than in the PL spectra and indicate the onset of the phonon bottleneck effect in these structures.

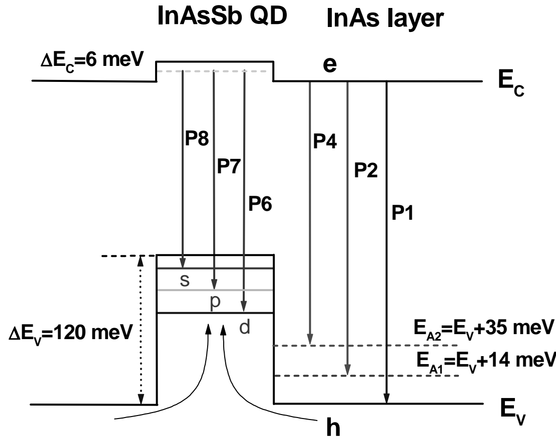


Figure 16. A schematic energy diagram of the Type II $InAs_{1-x}Sb_x/InAs$ encapsulated quantum dot

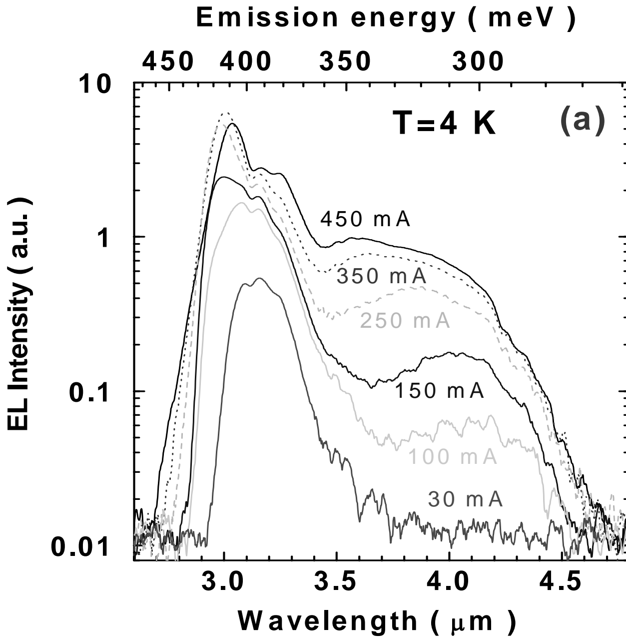


Figure 17. The 4 K EL spectrum of the QD LED at various injection currents Reprinted from A. Krier, X-L Huang, Physica E, **15**, 159 (2002) with permission from Elsevier [100].

Figure 17 shows the quantum dot LED emission spectra and their dependence on injection current at 4K. It is evident that at 30mA there is essentially no QD emission and the injection has to be increased to 100mA before any significant QD emission is observable. This is evidence of the phonon bottleneck where there are insufficient phonons available to scatter holes into the deep QD potential and the dots can become populated essentially only by Auger scattering which becomes significant at high carrier densities. The intensity of the QD electroluminescence at 4K was found to increase rapidly with increasing current relative to the InAs encapsulation. A blue shift was observed for the QD emission band as the current increased from 100 to 450 mA, due to the small density of states in the QD. In contrast, the InAs encapsulation shows no such shift. As the injection current increases, the intensity of the QD emission increases relative to the InAs-related

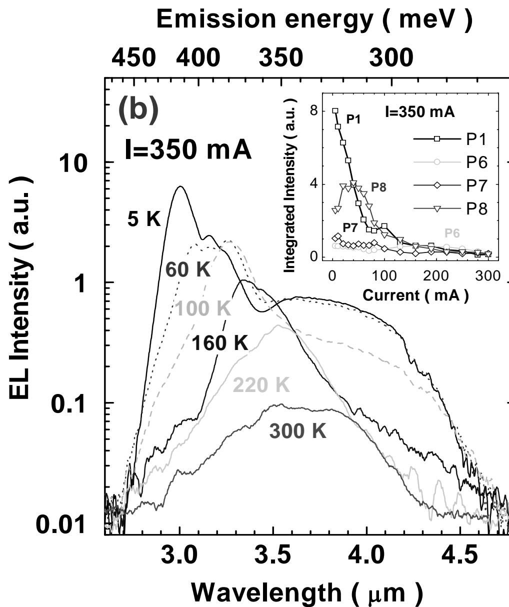


Figure 18. The integrated EL intensity dependence on temperature at 350 mA injection current. Reprinted from A. Krier, X-L Huang, Physica E, **15**, 159 (2002) with permission from Elsevier [100].

transitions (P1,P2). Figure 18 shows the EL spectra at various temperatures in the range from 5 to 300 K. The intensity of the QD emission band increases as the temperature is increased from 5 to 30 K, before decreasing with the band-band recombination (P1). It was also observed that as the temperature was increased the minimum injection current for observable QD emission decreased, and for a fixed injection level the optimum QD EL intensity was obtained near 50K as shown in the inset of Figure 18. For each of the individual QD levels there exists a competition between thermal activation, carrier relaxation and radiative recombination. The details of the observed EL emission spectra depend on the balance between these

processes. Assuming that the LED is operating in the phonon bottleneck regime at low temperature, and that the QD ground state is populated mainly by Auger recombination, then it is conceivable that as the temperature is increased the QD electroluminescence will increase in intensity as the phonon population increases. Then, as the temperature approaches 100 K, holes are released by thermal activation or perhaps via tunnelling. Nevertheless, from examination of Figure 18 it is evident that at 300 K the InAs related transitions are mostly quenched whereas the QD EL persists up to room temperature. This result is encouraging since although these LEDs are not optimised and consist only of a p-i-n structure with a single layer of QDs in the active region, they compare quite favourably with previous work on 4.6 μm InAsSb LEDs [69].

7 Superluminescence and Ring Lasers

The increase in optical output power in the above investigations has come about largely through a careful reduction of the residual carrier concentration in the active region of mesa-etched LEDs with a central dot contact. By employing a ring contact it is possible to increase output further and to observe superluminescence from mid-infrared LEDs. A Nomarski microscope photograph of a typical mesa-etched LED chip produced with such a ring electrode is shown in Figure 19. The corresponding back contact was deposited over the entire rear surface of the chip.

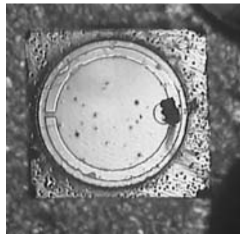


Figure 19. Nomarski microscope photograph of a typical mesa etched LED (350 μm diameter) with ring contact electrode (30 μm wide)

The full-width at half-maximum (FWHM) of the resulting output from the top and from the edge of the LED is compared in Figure 20. At 300K the electro-luminescence peaks near 4.6μm, making these emitters suitable for use in the environmental monitoring of carbon monoxide. By using similar ring electrodes on LED structures with InAs active regions (as described earlier) it is possible to obtain superluminescence at 3.3 μm.

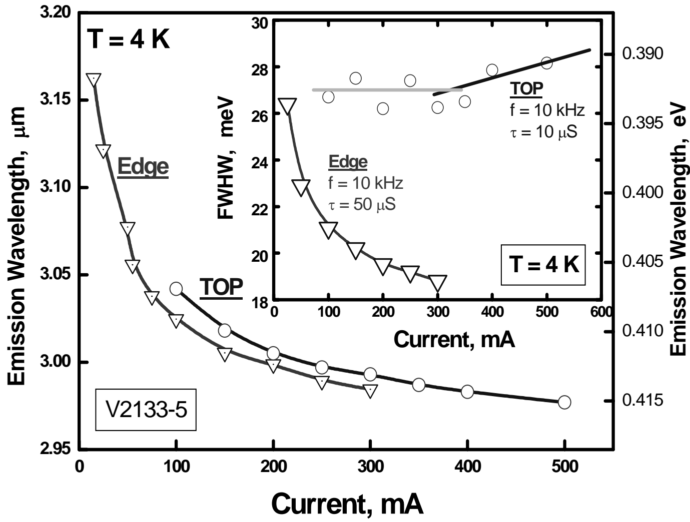


Figure 20. The current tuning behaviour of peak emission wavelength for edge and surface (t*op*) emission from the LED measured at 4K at a frequency of 10 kHz using 50 % duty cycle (50 µs pulses). The *inset* shows the corresponding FWHM variation with current, (duty cycle reduced to 10% at highest current to control Joule heating) [106].

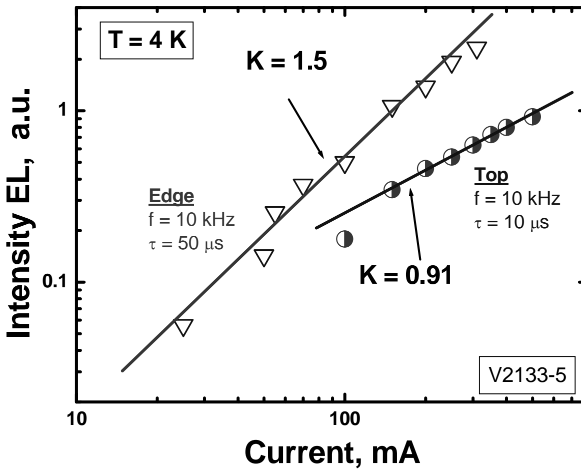


Figure 21. The light–current characteristics measured from the surface and edge of the ring geometry LED at 4K [106]

Due to current crowding effects the ring contact produces a whispering gallery mode as injection current is increased, resulting in superluminescence at room temperature and increased optical output. The electroluminescence emission spectra measured from the edge of the ring contact LED exhibit spectral line narrowing with increasing injection current [106]. The halfwidth (FWHM) at 15 mA is 880 nm whereas at 100 mA it is only 580 nm. The ratio of these halfwidths is 1.51 and shows that the device is a superluminescent diode. The light–current characteristics for surface and edge emission are compared in Figure 21. Both the curves have a linear portion in the current range from 0.015 to 0.3 A. The corresponding slopes were determined to be 1.5 for the edge radiation and 0.91 for the top radiation. A value of the slope of the light–current characteristic which is greater than 1 is further evidence for superluminescence (for edge radiation) in these LEDs.

Following the observation of superluminescence at 3.3 and 4.6μm due to the high Q arising from total internal reflection in LEDs with a ring contact configuration, further optimisation led to the demonstration of the first mid-infrared ring laser operating near 3 μm at 80 K as shown in the electroluminescence spectrum of Figure 22 [107].

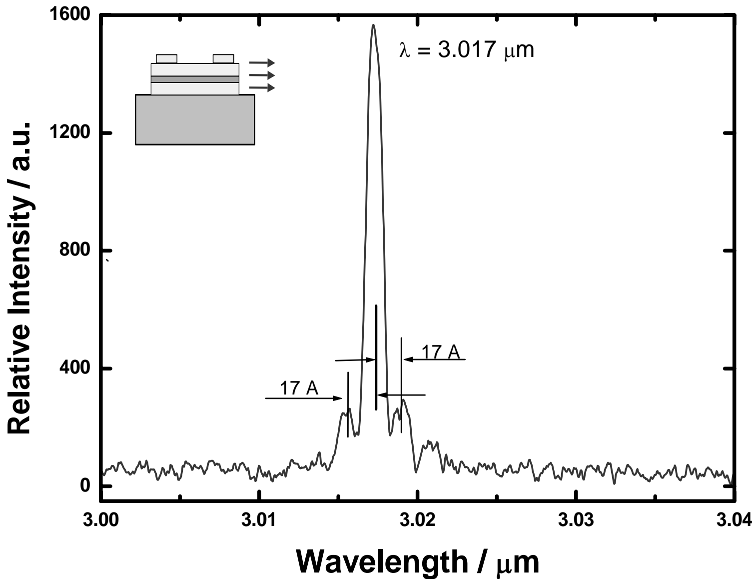


Figure 22. The mode spectrum obtained at 80 K from a mid-infrared whispering gallery mode ring laser [107]

Similar optical resonant mode behavior has been observed in microdisk lasers [108,109] and light emitting diodes operating at much shorter wavelengths [110,111]. It has been shown that a microdisk cavity may support two different

resonant mode types, radial and whispering gallery modes (WG) [112]. Radial modes are dominated by photon wave motion along the radial direction of the disk, the equivalent cavity being formed between the edge and the centre of the disk. These radial oscillations are unlikely in this case due to the ring contact and local current crowding which prevents such propagation. The WG mode, on the other hand, may be thought of as in-plane propagation around the inside perimeter of the mesa, which is facilitated by total internal reflection. The modes are solutions of the three-dimensional Maxwell equations. However, assuming that the radiation is effectively confined vertically within the active region of the mesa, then we may approximate this situation using a two-dimensional solution which yields Bessel functions for the radial field distribution. The effective cavity length of $2\pi R$ imposed by the periodic boundary condition on the circulating wave results in the WG eigenmode condition: $2\pi Rn = m\lambda$ for large integer m, where R is the mesa radius, n is the refractive index, and the mode spacing is given by $\Delta\lambda_{WG} = \lambda^2/2\pi Rn$. This enables us to estimate a WG mode spacing of 2.1 nm, which agrees reasonably with the value of 1.7nm measured experimentally from the spectra of Figure 22. Although the device had no intentional output coupler, coherent output was obtained, most likely due to light scattering from imperfections on the mesa edge of the ring waveguide. The laser exhibited a characteristic temperature of T_0 = 25 K with a maximum operating temperature of 125 K. Consequently, with proper design of the output coupler, further improvements in the laser operating temperature and threshold current can be expected.

8 Conclusion

The physics and technology associated with the fabrication and operation of mid-infrared LEDs has been briefly reviewed. The fundamental difficulties associated with realising practical sources at different target wavelengths for gas analysis have been discussed. The principal limitations to performance in determining high output power at room temperature were identified to be quantum efficiency and optical extraction. Some of the different device designs and approaches used for the suppression of non-radiative Auger recombination and the reduction of Shockley-Read-Hall centres were described to establish an overview of the "state-of-the-art". Even though measurements at longer wavelengths are more challenging, the use of high applied pressure as an investigative tool has been shown to be very useful in identifying and understanding recombination mechanisms in mid-infrared LEDs. Liquid phase epitaxy continues to hold a strong position in mid-infrared LED technology and many of the current best LEDs have been fabricated using this technique. Furthermore, the use of rare earth gettering in LPE growth of InAs(Sb) LEDs has been established as an important technique. LED performance improvements were realised (a) by using rare earth gettering as a melt purification technique to reduce residual carrier concentration and (b) by use of Type II interface heterostructures to obtain Auger suppression. LEDs which match some of the target wavelengths of special interest for gas detection, including 3.3μm (CH_4) and 4.6μm (CO) have been realised [113]. These emitters are poised to become useful components in the realisation of cost-effective practical instruments for the

environmental monitoring of methane and carbon monoxide. The recent development of quantum dot LEDs and whispering gallery mode ring lasers for the mid-infrared also have interesting possibilities in this rapidly expanding field.

Acknowledgements

The authors wish to thank all those colleagues who have made significant contributions to the work summarized in this chapter, especially H. Gao, Z. Labadi and S. Choulis. We are grateful to the Engineering and Physical Sciences Research Council (EPSRC) for providing the funding associated with this research. We also wish to thank The Royal Society and Kidde plc for supporting various aspects of the work.

References

1. See for example: InAsSbP/InAs LEDs for the 3.3-5.5μm spectral range, B. Matveev *et al.*, IEE Proc. Optoelectronics, **145,** Special Issue on Mid-IR devices & materials and papers therein – (see also vol **144** for additional related papers)
2. High Power 4.6 μm LEDs for CO detection grown by LPE using rare earth gettering. A.Krier, H.Gao, V.Sherstnev, Yu.Yakovlev. Electronics Letters, (USA), **35**, 1665-7(1999)
3. A.N.Baranov, A.N.Imenkov, O.P.Kapranchik, V.V.Negreskul, A.G.Chernyavskii, V.V.Sherstnev Yu.P.Yakovlev. Long-wavelength light-emitting diodes, based on InAsSbP/InAs heterostructures (λ=3.0-4.8μm) at 300K with wide-gap window. Pisma v Zhurnal Tekh.Fiziki (USSR), **16** (16), p.42-47, (1990).
4. S. McCabe and B.D. MacCraith , *Electron.Lett.* **29** 1719–21(1993)
5. S.D. Smith, A. Vass, P Bramley, J.G. Crowder and C.H. Wang, IEE proc-Optoelectronics, **144**, No 5, 266 (1997)
6. M.K. Parry and A. Krier, *Electron. Lett.* **30** 1968–69 (1994)
7. A.A. Popov, V.V.Sherstnev, Y.P. Yakovlev, A.N. Baranov and C. Alibert, *Electron. Lett.* **33** 86–88 (1997)
8. A. Krier and Y. Mao, *IEE Proc.Optoelectron.* **144** 355–59 (1997).
9. A.A. Popov , M.V Stepanov, V.V. Sherstnev and Y.P. Yakovlev, *Tech. Phys. Lett.* **24** 596–98 (1997)
10. A. Krier & V.V. Sherstnev, J. Phys. D: Appl. Phys. **33**, 101-106 (2000) Powerful interface light emitting diodes for methane gas detection.
11. A.A Popov, V.V. Sherstnev, Y.P. Yakovlev, A.N. Baranov and C.Alibert, 'Powerful mid-infrared light emitting diodes for pollution monitoring' Electron. Lett., **33** (1997) 86-88
12. A.Krier and Y.Mao, '2.5μm light emitting diodes in $InAs_{0.36}Sb_{0.20}P_{0.44}$ /InAs f or HF detection, IEE Proc Optoelectronics **144**, 355-359 (1997)
13. Powerful 4.6μm light emitting diodes for CO detection, A. Krier, H.H. Gao, V.V. Sherstnev & Y. Yakovlev, - J. Phys D, **32** , 3117-3121 (1999)

14. A.A Popov, M.V. Stepanov, V.V Sherstnev, Y. P. Yakovlev 'InAsSb light emitting diodes for the detection of CO_2 (λ=4.3μm)' Technical Phys. Lett., Vol **24**, No.8, 1998 p596-598
15. A. Verdin, Gas Analysis Instrumentation, Wiley, New York, 1975
16. I. Melngailis & R.H. Rediker, J. Appl. Phys, **37**,899(1966)
17. I. Melngailis & R.H. Rediker, Appl. Phys. Lett., **2**,202 (1963)
18. O. Madelung, Physics of III-V compounds, Wiley, New York, 1964
19. A. Andreev *et al.*, Sov. Techn. Phys Lett., **16**, 135-137(1990), Semicond. **27**, 236-240 (1993)
20. N. P. Esina *et al.*, J . Appl. Spectrosc**.** **42**, 1985, 7465
21. A.A Popov, V.V. Sherstnev, Y.P. Yakovlev, 1.94μm LEDs for moisture content measurements, Tech. Phys. Lett., **23**, 783 (1997)
22. H.H. Gao, A. Krier and V.V. Shertsnev, J. Phys. D – Appl. Phys. **32** (1999) 1768-1772
23. A. Krier, V.V. Sherstnev & H.H. Gao, A novel LED module for the detection of H_2S at 3.8 μm, J. Phys. D.-Applied Phyiscs, **33**, 1656, (2000)
24. PJP Tang *et al.*, Appl. Phys. Lett., **72**, 3473 (1998) efficient 300 K light emitting diodes at λ = 5 μm and 8 μm from InAs/InAsSb single quantum wells.
25. M.J. Pullin *et al.*., Appl. Phys. Lett., **74**, 2384 (1999) Room temperature InAsSb strained-layer superlattice light emitting diodes at 4.2μm with AlSb barriers for improved carrier confinement
26. C. Sirtori *et al.*., Appl. Phys. Lett., **66**, 4, (1995) Quantum cascade unipolar intersubband light emitting diodes in the 8-13μm wavelength region
27. J. Faist *et al.*, Appl. Phys. Lett.,**64**,1144 (1994)
28. J. Faist *et al.*, Appl. Phys. Lett., **65**,94(1994)
29. T. Ashley, C.T. Elliott, N.T. Gordon, R.S. Hall, A.D. Johnson & G.J. Pryce, Appl. Phys. Lett.,**64**, 2433(1994)
30. T. Ashley, N.T. Gordon & T.J. Phillips, J. Modern Optics, **46**, 1677 (1999) Optical modelling of cone concentrators for positive and negative IR emitters.
31. T. Ashley & C.T. Elliott, Electron. Lett., **21**, 451 (1985)
32. T. Ashley, C.T. Elliott, N.T. Gordon, R.S. Hall, A.D. Johnson & G.J. Pryce, Infrared Phys & Technol., **36**, 1037 (1996)
33. T. Ashley, C.T. Elliott, N.T. Gordon, R.S. Hall, A.D. Johnson & G.J. Pryce, J. Crystal growth, **159**, 1100 (1996)
34. C. Verie & C.R. Granger, Acad. Sci. Paris, **261**, 3349 (1965)
35. K.K. Mahavadi *et al.*, J. Vac. Sci. & Tecjhnol., **A 8**, 1210 (1990)
36. A. Ravid &A. Zussman , G. Cinader & A. Oron, Appl. Phys. Lett., **55**, 2704 (1989)
37. R. Zucca *et al.*, J. Vac. Sci. & Technol., **A 6**, 2725 (1988).
38. P. Bouchut *et al.*., J. Vac. Sci. & Technol.,**B9**, 1794 (1991)
39. T. Ashley, C.T. Elliott, N.T. Gordon, R.S. Hall, C.D. Maxey & B.E. Matthews, Appl. Phys. Lett. **65**, 2314 (1994)
40. W. Lo & D.E. Swets, Appl. Phys. Lett., **36**, 450 (1980)
41. Z. Shi *et al.*, Appl. Phys Lett., **66**, 2573 (1995)
42. Xu J, Lambrecht A, Tacke M**,** IEEE Photonic Tech. Lett. **10**: (2) 206-208 (1998)
43. Feit Z, Mak P, Woods R, *et al.*. Spectrochimica Acta A **52**: (8) 851-855 (1996)
44. C.R. Pidgeon *et al.*, Phys Rev. B. **58**, 12908 (1998)
45. Z. Shi, Appl. Phys. Lett., **72**, 1272 (1998)
46. T. Ashley *et al.*, Appl. Phys. Lett., **64,** 2433 (1994) and T. Ashley & C.T. Elliott, Semicond. Sci & Technol. **6**, C99 (1991)
47. M.J. Kane *et al.*, Mater. Res. Soc. Symp.Proc. **450**, 129 (1997)
48. A.R. Beattie, J. Phys. Chem. Solids, **23**, 1049 (1962).

49. M.Kane *et al.*, "Emission efficiency in InAs LEDs controlled by surface recombination" Presented at Boston MRS conference (Dec 1996)
50. M.J. Kane *et al.*, Appl. Phys. Lett., **76**, 943 (2000)
51. M. Takeshima, Auger recombination in InAs, Jap. J. Appl. Physics **22**, 491, (1983)
52. See for example; A.N.Baranov, T.I.Voronina, T.S.Logunova, M.A.Sipovskaya, V.V. Sherstnev, Yu.P.Yakovlev. Semiconductors, **27**, (1993). Zotova NV, Karandashev SA, Matveev BA, *et al.*. Semiconductors, ***33,*** 920, (1999). A.Krier, H.H.Gao, V.V.Sherstnev. Journal of Applied Physics, 85, 8419, (1999).
53. A.Krier, V.V. Sherstnev, Yu.Yakovlev. Journal of Physics D, **33**, 101, (2000).
54. A.Kumar, D.Pal & D.N. Bose, J. Electronic Mater., **24**, (1995) 833
55. M. C. Wu *et al.*., J. Appl. Phys. **71**(1) 1992, 456
56. W. Gao *et al.*., J. Appl. Phys. **80**(12) 1996, 7094
57. T. Bagraev, L.S. Vlasenko, K.A. Gatsoev, A.T. Gorelenok, A.V. Kamanin, V.V. Mamutin, B.V. Pushnyul, V.K. Tibilov, Y.P. Tolparov and A.E. Shubin, Sov. Phys.-Semiconductors **18** (49) 1984
58. F. Bantien, E. Bauser & J. Weber, "Incorporation of erbium in GaAs by liquid phase epitaxy", J. Appl. Phys.,**61**,(1987)2803
59. Purification of epitaxial InAs grown by liquid phase epitaxy using gadolinium gettering, A. Krier, H.H. Gao & V. Sherstnev, J. Appl. Phys., **85**,(12)8419-8422 (1999).
60. High quality InAs grown by liquid phase epitaxy using gadolinium gettering, H.Gao, A. Krier & V. Sherstnev, Semicond. Sci & Technol. **14**, 441-5 (1999), IOP Publishing
61. N.V. Zotova *et al.*., The first intenational conference on mid-optoeletronics -materials and devices, MIOMD Lancaster, 1996
62. R. D. Grober , *et al.*., Physics Review B, vol. **43**, No. 14, 1991, 11732
63. Z.M.Fang, *et al.*., J.Appl. Phys. **67**(11), 1990, 7034
64. Y.Lacroix, *et al.*., Appl. Phys. Lett. **66**(9) 1995, 1101
65. R. D. Grober *et al.*., J. Appl. Phys**. 65**(10) 1989, 4079
66. P.J.P Tang, *et al.*., Semicond. Sci. Technol**. 8**, 1983, 2135
67. H.H. Gao, A. Krier and V.V. Shertsnev, J. Phys. D – Appl. Phys. **32** (1999) 1768-1772
68. S.Kim, M.Erdtmann, D.Wu, E.Kass, H.Yi,J.Diaz, and M.Razeghi, Photoluminescence study of InAsSb/InAsSbP heterostructures grown by low-pressure metalorganic chemical vapor deposition. Appl.Phys.Letters **69** (11), pp. 1614-1616 (199b).
69. A. Krier, H.H. Gao, V.V. Sherstnev & Y. Yakovlev, High Power 4.6μm light emitting diodes for CO detection, J. Phys D, **32** , 3117-3121 (1999), IOP Publishing
70. A.N.Baranov, T.I.Voronina, A.A.Gorelenok, T.S.Logunova, A.M.Litvak, M.A.Sipovskaya, S.P.Starosel'tseva, V.A.Tikhomirova, V.V.Sherstnev, Yu.P.Yakovlev. Semiconductors, **26**, 905, (1992).
71. A.Krier, V.V. Sherstnev, The influence of melt purification and structure defects on mid-infrared light emitting diodes, J. Phys. D-Applied Physics, **36** (13), 1484, 2003, IOP Publishing
72. S.A. Choulis, A. Andreev, M. Merrick, S. Jin, D.G. Clarke, B.N. Murdin, A.R. Adams, A. Krier, V.V. Sherstnev, Phys. Stat. Sol. B- Basic Research **235**(2), 312 (2003), John Wiley
73. A. R. Adams, M. Silver, and J. Allam, *Semiconductors and Semimetals*,Vol. **55**, Academic, London, (1998), pp. 301–331.
74. G. G. Zegrya and A. D. Andreev, JETP **82**, 328, (1996).
75. G. G. Zegrya and A. D. Andreev, Appl. Phys. Lett. **67**, 2681, (1995).
76. S.A. Choulis, A. Andreev, M. Merrick, , A.R. Adams, B.N. Murdin, A. Krier, V.V. Sherstnev, Appl. Phys. Lett, **82**(8), 1149 (2003), American Institue of Physics

77. T. Ashley, D.T. Dutton, C.T. Elliott, N.T. Gordon & T.J. Phillips, SPIE conference 3289 "Micro-optics Integration and assemblies" part of Photonics West meeting, San Jose, CA, USA, 26-30 Jan 1998
78. T. Ashley, N.T. Gordon & T.J. Phillips, J. modern Optics, **46** (11), 1677 (1999)
79. I. Schnitzer *et al.*, Appl. Phys. Lett., **63**, 2174 (1993)
80. Hadji E. et.al.: Resonant cavity light emitting diodes for the 3-5μm range, *Solid-State Electronics*, Vol. **40**, pp.473-6, 1996
81. P.N. Stavrinou *et al.*, J. Appl. Phys., **86**,3475 (1999)
82. H. De Neve *et al.*, Impact of planar microcavity effects on light extraction - Part I: Basic concepts and analytical trends , IEEE J Quantum Elect **34**: (9) 1612-1631 (1998)
83. Baranov AN, Rouillard Y, Boissier G, *et al.*. Electron Lett **34**: (3) 281-282 (1998)
84. N.T. Gordon, C.L. Jones & D.J. Purdy, Infrared Physics., **31**, 599 (1991)
85. B. Matveev *et al.*, IEE Proc. Optoelectronics **145**, 254 (1998)
86. A. Sugimura, "Band to band Auger effect in GaSb and InAs lasers", J. Appl. Phys. Vol. **51**, No.8, 4405-4411(1980).
87. M. Takeshima, "Disorder-enhanced Auger recombination in III-V alloys" J. Appl. Phys. Vol.**49**, No12, pp6118-6124.(1978)
88. S. Chinn, P. Zory, and A. Reisinger, IEEE J. Quantum Electron., **QE-24** 2191 (1988)
89. B. Gel'mont, Z.Sokolova, and I. Yassievich, Sov. Phys. Semicond. **16**, 382 (1982)
90. E. R. Glaser, B. R. Bennett, B. V. Shanabrook, and R. Magno, Appl. Phys. Lett. **68**, 3614 (1996).
91. G. Park, O. B. Shchekin, d. L. Huffaker, and D. G. Deppe, Appl. Phys. Lett. **73**, 3351 (1998).
92. E. Alphandery, R. J. Nicholas, N. J. Mason, B. Zhang, P. Mock, and G. R. Booker, Appl. Phys. Lett. **74**, 2041 (1999).
93. A. Krier, Z. Labadi, and A. Hammiche, J Phys D: Appl. Phys. **32**, 2587 (1999).
94. K. Mukai, N. Ohtsuka, H. Shoji, and M. Sugawara, Appl. Phys. Lett. **68**, 3013 (1996).
95. S. Raymond, P. Hawrylak, C. Gould, S. Fafard, A. Sachrajda, M. Potemski, Wojs, S. Charbonneau, D. Leonard, P. M. Petroff, and J. L. Merz, Solid State Commun. **101**, 883 (1997).
96. Raymond, X. Guo, J. K. Merz, and S. Fafard, Phys. Rev. B, **59**, 7624 (1999).
97. S. Fafard, Appl. Phys. Lett. **76**, 2707 (2000).
98. J. Oswald, K. Kuldová, J. Zeman, E. Hulicius, S. Jullian, and M. Potemski, Materials Science and Engineering B, **69-70**, 318 (2000).
99. A. Krier, X. L. Huang, and A. Hammiche, Appl. Phys. Lett. **77**, 3791 (2000)
100. A. Krier, X-L Huang, Physica E, **15**, 159 (2002)
101. P. J. P. Tang, C. C. Philips, and R. A. Stradling, Semicond. Sci. Technol. **8**, 2135 (1993).
102. P. J. P. Tang, M. J. Pullin, S. J. Chung, C. C. Phillips, R. A. Stradling, A. G. Norman, Y. B. Li and L. Hart, Semicond. Sci. Technol. **10**, 1177 (1995).
103. R. Grober & H.D. Drew, J. Appl. Phys. **65**, 4079 (1989)
104. S. H. Wei, and A. Zunger, Phys. Rev. B, **52**, 12039 (1995).
105. Y. B. Li, D. J. Bain, L. Hart, M. Livingstone, C. M. Ciesla, M. J. Pullin, P. J. Tang, W. T. Yuen, I. Galbraith, C. C. Phillips, C. R. Pidgeon, and R. A. Stradling, Phys. Rev. B, **55**, 4589 (1997).
106. V.V. Sherstnev, A. Krier, G. Hill, High tunability and superluminescence in InAs mid infrared light emitting diodes, J. Appl. Phys. D - Applied Physics, **35**, 196, (2002) IOP Publishing
107. A. Krier, V.V. Sherstnev, D. Wright, A.M. Monakhov, G. Hill, Electron. Lett., **39** (12), 916 (2003) IEE

108. S.A. Backes, J.R.A. Cleaver, A.P. Heberle , J.J. Baumberg and K. Kohler, Appl. Phys Lett., **74**, 176 (2000)
109. S. Chang, B. N. B. Rex, R.K. Chang, G. Chong and L.J. Guido, Appl. Phys. Lett., **75**, 166 (2000)
110. S.X. Jin, J.Li, J.Z. Li, J.Y. Lin and H.X. Jiang, Appl. Phys. Lett., **76**, 631 (2000)
111. R.A. Mair, K.C. Zeng, J.Y. Lin , H.X. Jiang, B. Zhang, L. Dai, A. Botchkarev, W. Kim, H. Morkoc, and M. A, Khan, Appl. Phys. Lett., **72**, 1530 (1998).
112. R.P. Wang and MM. Dumitrescu, J. Appl. Phys., **81**, 3391 (1997)
113. Powerful interface light emitting diodes for methane gas detection, A. Krier & V.V. Sherstnev, J. Phys. D: Appl. Phys. **33**, 101-106 (2000)

LED-Photodiode Opto-pairs

Boris A. Matveev
IOFFE Physico-Technical Institute, Russian Academy of Sciences, Polytekhnicheskaya ul. 26, St. Petersburg, 194021 Russia

1 Introduction

Optical sensing, that is chemical composition determination through the use of spectroscopic/transmission measurements, is becoming more and more attractive due to fascinating improvement of the performance of diode lasers, light emitting diodes (LEDs) and photodiodes (PDs). The latter are of particular importance since photovoltaic operation mode leads to a minimum noise that is beyond the capabilities of photoresistors, thermopiles and bolometers.

Recent years have seen extensive research of the mid-IR (3–14 μm) diode optopair, *i.e.* an optically coupled source and detector assembly that contains p-n junctions which can effect each other *via* electrically generated or/and absorbed photons. Diode optopairs in this wavelength range which allow the introduction of a gas or a liquid sample under investigation into the optical path, or in other words "diode optopair sensors" are of particular interest for environmental monitoring. This is because most industrial gases/liquids have characteristic absorption bands in the above spectral range and because the atmospheric transparency in two bands ("windows") at 3–5 (I) and 8–14 (II) μm allow the use of long paths for transmission measurements and hence the detection of trace gases.

The laser is without doubt the most powerful instrument for "spectroscopic" applications because of its high power, small size and narrow linewidth [1]. Measurements made using a tunable single mode laser produce the most advanced data that allow the measurement of gas concentrations in the ppb range. The technique has been extensively investigated, however according to [1] "the common experience of these studies is that no "sweet spot" in the multi-dimensional space of performance/cost/field use capability has ever been achieved". In addition to this, near room and room temperature lasing in the long wavelength region (> 3 μm) has not been achieved in the "classical" long lifetime double heterostructure devices. Quantum cascade lasers exhibit excellent performance in terms of output power, spectral purity and tunability, however, because of impurity diffusion through thin heterostructure layers there are certain doubts about their capability for long lifetime operation especially at elevated temperatures (>100 ºC).

Extensive research has been carried out on p-n junction devices with active layers made from A^2B^6 and A^4B^6 materials for the first (I) and second (II) atmospheric windows. These materials, however, suffer from low thermal conductivity which together with their metallurgical instability and low resistance to moisture drives the search for the alternative materials. A^3B^5 semiconductors, namely InAs, InSb and the alloys of these two are good candidates for the task and

they have already been used for fabrication of powerful LEDs (see, *e.g.* the overview by Smith *et al.* [2]) and efficient detectors.

This review covers recent results on the 3–8 μm mid-IR LEDs and 3–4.3 μm photodiodes operating in the 0-300 ºC temperature range and grown onto InAs substrates by the LPE method at the IOFFE Physico-Technical Institute.

2 Device Configuration and Fabrication

The temperature of the activated p-n junction is among most important factors influencing mid-IR diode performance. Under extreme pumping conditions, the Joule-heating together with strong non-radiative recombination remarkably effects the current voltage (I-U) [3], light-current (L-I) and emission spectrum [4] of a conventional diode. Presented below is an example of the latter measured in an InGaAs p-n homojunction LED for several pumping conditions.

InGaAs layers with nearly constant composition along the growth direction were grown at 650-720 °C by the LPE method on a 350 μm-thick p-InAs (111) substrate.

The growth of the 50-60 μm thick p-InGaAs buffer layer was followed by the deposition of a 3-5 μm thick n-type InGaAs active layer. Zn or Mn was used as a p-dopant for p-n junction formation during the growth. N-type layers were not intentionally doped. Hall measurements indicate that the free carrier concentration is $n \approx p \approx 1 - 2 \times 10^{17}\ cm^{-3}$.

The resulting heterostructures were processed by wet photolithography into circular mesas with a diameter, D_m, of 430 μm and top p-contact with a diameter, d_c, of 150 μm. Both n- and p-contacts were vacuum sputtered and finally covered by Au using galvanic deposition. Finally LED chips of $0.06 \times 0.9 \times 1.0\ mm^3$ dimensions were mounted substrate side down onto a TO-18 header with an Au wire soldered to the top p- contact area as shown in Figure 1.

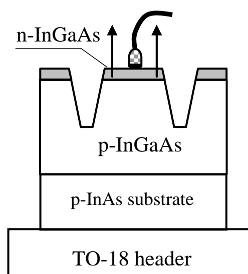


Figure 1. Schematic of the conventional InGaAs LED with top wire bonding

Figure 2 (left) presents the variation in emission spectrum with pulse duration at a pumping current of 1A. As seen from Figure 2 (left) the emission peak shifts to long wavelength with increasing pulse duration as if the band gap is decreasing with increasing temperature/pulse width.

Ignoring the self absorption of the emitted light and taking into account an exponential distribution of thermalized carriers in the valence band, we can estimate the p-n junction temperature by using the short wavelength shoulder spectrum

approximation as I~exp((hυ$_0$-hυ)/kT). The corresponding data together with temperature estimations based on electroluminescence peak position in the 20-180 ºC range taken from Aidaraliev *et al.* [5] are presented in Figure 2 (right). In the latter case the "zero point" (or room temperature operation point) with no overheating was considered to take place at the lowest current (5 mA). As seen from Figure 2, the application of current pulses as long as 500 μs (1A) results in a p-n junction overheating, Δt, of 40-100 ºC. A high Δt value is a consequence of the high series resistance amounting to 1.14 Ω, caused mainly by the high resistance of the p-layer and p-substrate. In an InAs structure the conductivity of the p- and n-type layers may differ considerably with typical values of $\sigma_p = 80(\Omega cm)^{-1}, \sigma_n = 200(\Omega cm)^{-1}$ for the p- and n-type layers respectively. Thus Joule heating and luminescence temperature quenching in the surface emitting LED is usually minimized by making the thickness of the p-layers as small as possible. Of course, this is not a compulsory argument for the photodetector whose operating current is small.

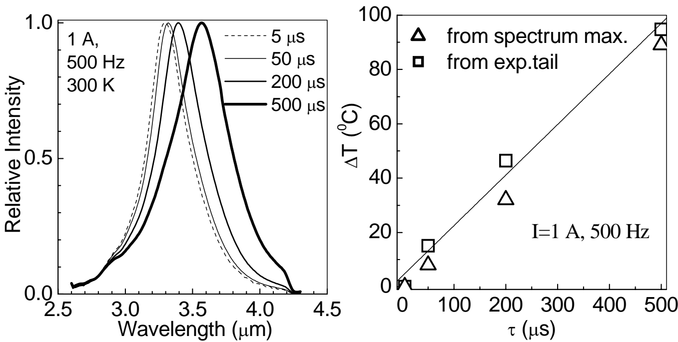


Figure 2. "Room temperature" emission spectra at fixed current of 1 A and several pulse durations in the 5-500 μs range (*left*) and overheating of the p-n junction region estimated from spectral position and short wavelength shoulder (*right*) in InGaAs homojunction LED

The need for a heatsink means that the thin p-layer has to be bonded to a header providing the shortest path for the heat from the p-n junction to the heatsink. Thus the conventional LED structure, with the output extracted perpendicular to the p-n junction, consists of a thin (5 μm or less) p-type layer adjoining active layer(s) grown onto an n-type substrate which is thinned during processing to 100-150 μm so that it is transparent for the generated radiation.

2.1 Substrate and Buffer Layers

The above construction, with radiation escaping/entering perpendicular to the device surface, is relatively easy to achieve in the ternary alloy based structures grown onto wider gap substrates, mid-IR examples being $InAs_{1-x}Sb_x$/GaSb [6], $InAs_{1-x}Sb_x$/InAs [7,8] and $InAs_{1-x}Sb_x$/InSb [8,9] systems. The transparency of the substrate (in the working wavelength range) in the latter case is not typical for compositions consisting of ternary alloys and a binary substrate and is due to the strongly non-linear composition dependence of the $InAs_{1-x}Sb_x$ energy gap.

Another version of an epi-side down structure employs graded band layers: it is grown under conditions of high substrate plasticity on relatively thin substrates. It has been shown that due to the elevated hardness of the quaternary alloy [10], the plastic deformation processes, which occur in the InAs substrate and accompany its bending ($R \to R_0 = a / grad(a)$, where a is the lattice parameter), remove most of the elastic stresses. As a result, one obtains bent structures containing graded band gap epi-layers having low dislocation density and plastically deformed substrates with high dislocation density [11]. This approach was used to make the first long wavelength (λ>4 μm) III-V room temperature LEDs [12], and high quality $InAs_{0.46}Sb_{0.54}$ layers with an energy gap as small as 90 meV (300 K) [13].

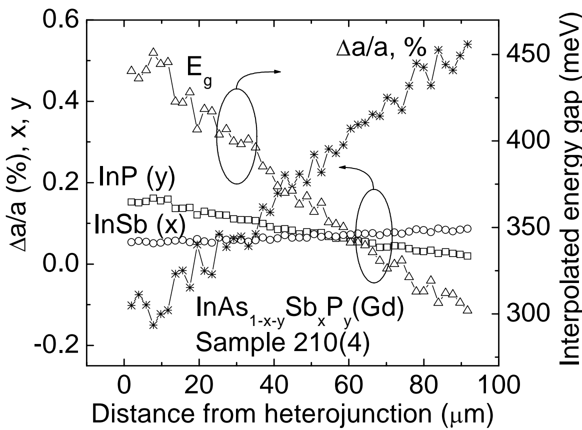


Figure 3. Distribution of composition (x, y) in graded $InAs_{1-x-y}Sb_xP_y$ layers along the growth direction and interpolated values of energy gap (E_g) and lattice mismatch (Δa/a) [14]

InAsSbP epilayers 25 - 60 μm thick have been grown onto 350 μm thick n-InAs (111) substrates at 650 – 680 ºC by the LPE method [14]. The layers were characterized by low dislocation density (10^4 cm^{-2}) due to the above-described “inverse” defect formation. The samples were curved with a radius of curvature, R, of 10 - 15 cm and were characterized by an increase of antimony and a decrease of phosphorus concentrations towards the epilayer surface as shown in Figure 3. Also shown in Figure 3 are the interpolated values of energy gap with energy gap gradient ∇=1.7 meV/μm and E_{gmin}=300 meV. The interface values of the lattice parameters of the InAs substrate and the InAsSbP are nearly identical (Δa/a<0.05%) while the narrow band gap surface is mismatched with InAs and Δa/a is as high as 0.5%.
The wide band gap part of the InAsSbP with $E_g = E_{gMAX}$ at the interface is a transparent “window” for the photons whose energy is close to the narrow band gap value of the InAsSb(P) and thus the transparency of the entire layer is defined by a relatively thin layer at the surface of the structure with $E_g = E_{gMIN}$. The effective optical thickness of the graded band gap layer with $E_g = E_{gMIN}$ is $t_{eff.} = \varepsilon / \nabla$, where

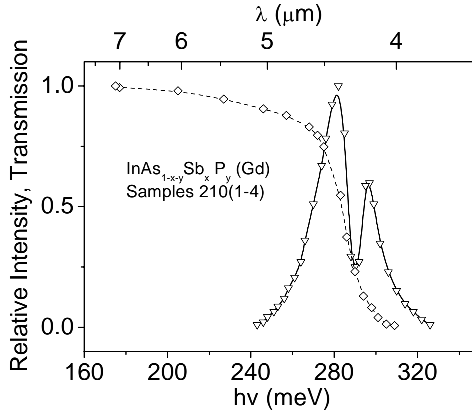


Figure 4. Emission (solid) and transmission (dashed) spectra of graded InAsSbP layer at room temperature. The *dip* at 4.3 μm is caused by the atmospheric CO_2 absorption [14].

ε -is the characteristic energy of the exponential slope of the optical density (D=-ln(T/T_o)) and ∇ is the gradient of the band gap [15]. It follows from Figure 3 that $t_{eff.} = \varepsilon / \nabla$ =7.3 meV/1.7 meV·μm^{-1}= 4.3 μm and if we assume that the absorption coefficient at the photoluminescence peak amounts to 3x10^3 cm^{-1} we are to expect transparency of around 0.3 at the above peak. The latter case is shown in Figure 4 where transparency and photoluminescence spectra measured in "transmission" geometry are presented. It is seen that transparency of the layer is quite small on the short wavelength shoulder of the photoluminescence spectrum and thus significant losses are expected at these wavelengths. It is possible to decrease the $t_{eff.}$ value by increase of ∇ however each case should be carefully considered since the gain obtained by the increase of optical transparency due to the increase of lattice parameter/energy gap gradient is often balanced by the increase of crystal defects and stresses and a related luminescence quenching.

Due to material and process issues, the creation of a graded band gap window layer is not appropriate for all growth techniques *e.g.* the growth rates in many modern procedures, such as MBE, are very low and thus such methods could hardly be implemented for the creation of the thick graded layers. On the other hand the epi-side bonded constructions can utilise narrow gap semiconductors such as InAs and InSb which are characterized by small electron effective mass and therefore strong degeneracy in the conduction band occurs at relatively low electron concentrations (n>10^{16}-10^{17} cm^{-3}) with the result that a noticeable emission [16] and transmission shift towards short wavelengths occur. Thus heavily doped InSb [17] or InAs [18] based structures can be mounted epi-side down with the radiation being extracted through a thick substrate. Figure 5 supports the above statement for n^+-InAs material by demonstrating transmission spectra of ~350 μm thick Sn- doped substrates at 20 and 60 °C [19]. As seen from Figure 5 the Moss-Burstein shift in heavily doped n^+- InAs (n=6×10^{18} cm^{-3}) makes the corresponding substrates transparent to the 3.3 μm radiation (T/T_o= 70%, 20 °C) with good prospects for the

fabrication of efficient room temperature detectors/LEDs with a thin (<100 μm) n^+-InAs "window".

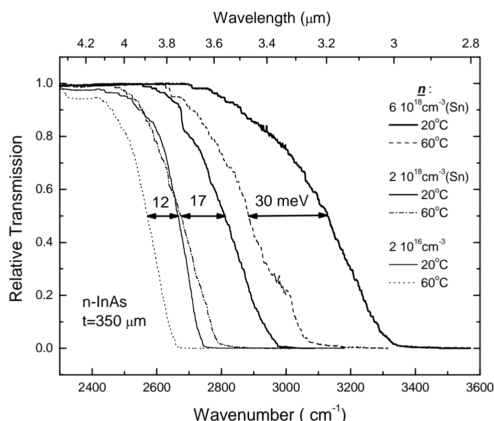


Figure 5. Transmission spectra of 350 μm thick n-InAs(Sn) at room temperature (*thick lines*) and at 60ºC (*dotted lines*) at free electron concentrations of 2×10^{16}, 2×10^{18} and 6×10^{18} cm^{-3}

2.2 Active Layer Properties

2.2.1 Layer Doping

The properties of semiconductor materials are dependent on the specific growth regimes/conditions and/or particular concentration of the incorporated dopants. Fang and co-workers [20] reported on the superior photoluminescence intensity from MOPVE grown InAs and InSb with respect to an InAsSb alloy using Ar-ion laser pumping in the range of a few to 100 mW. A decline in the PL intensity of a few hundred times, and ten times, within the 10-300 K temperature range was found for InAs and InAsSb layers respectively. The slow temperature variation in peak position of the the InAsSb was explained as a consequence of a wave-vector non-conserving transition contributing a 2kT difference between the photoluminescence maximum and the energy gap, E_g..

Fisher *et al.* [16] came to the conclusion that the photoluminescence intensity from n-InAs at 4K depended on the growth method, and had relative intensities of : 4(LPE), 2(MBE), 1.3(MOVPE) and 1 (Czochralski). It appears that the LPE growth method is still among the best techniques for producing InAsSb alloys with high quantum efficiency. Gong *et al.* [21] showed that their LPE grown layers were characterized by a narrow photoluminescence spectrum and strong temperature dependences of the PL maximum and intensity (50 meV shift and 100-fold drop over the 5 - 300 K range respectively). This was considered as an evidence for the absence of crystal defects normally present in MBE or MOVPE grown samples. Fisher *et al.* also concluded that the photoluminescence intensity at 4 K peaked for "pure" samples ($n=10^{16}$ cm^{-3}) with subsequent 20-fold decrease for the doped layers ($n=10^{18}$

cm^{-3}). The latter conclusion differs from the results of Esina [22] where the 77 K photoluminescence intensity peaked for highly doped InAs layers (n, p ~$10^{17}cm^{-3}$).

Recently several authors reported on the efficiency enhancement of InAs-, InAsSb- [23] and InGaAsSb–based devices [24] associated with the rare earth (*e.g.* gadolinium (Gd)) doping of the active layer during the LPE growth. The effect of the rare earth is attributed to a decrease in the background impurity concentration of group-VI elements due to their interaction with the rare earth and the formation of compounds with high melting temperature, and the corresponding suppression of their incorporation into the solid phase. This approach enabled a reduction in the donor concentration down to $n = 1.5\times10^{15}$ cm^{-3} which is beneficial for LED structures as well as for the photodiodes whose R_0A product is enhanced in heterostructures with lightly and heavily doped active and "window" or contact layers respectively [25].

2.2.2 Layer Thickness

The optimum thickness of the device active layer is usually a compromise between the absorption of photons with energy close to the energy gap and the rate of their generation. The simplest case is for a PD operating at low incident radiation intensity (Auger rate independent) so that an optimum value of $d_{opt}=1.26/\alpha$ is obtained for the active layer, where α is the absorption coefficient [25]. Similarly, in most investigations the influence of the LED active layer thickness on LED performance was considered without reference to non-linear effects such as Auger recombination [26] within simple models incorporating the influence of fundamental absorption on output power and emission spectrum. An example of such influence is presented in Figure 6 which shows emission spectra of 10, 20 and 50 μm thick n-InAs ($n=3\times10^{16}$ cm^{-3}) plates at room and liquid nitrogen (d=50 μm) temperatures. The emission was activated by optical pumping with a GaAs LED (λ= 0.87 μm) in a "transmission" geometry with the pump and detector located at the opposite sides of the plate.

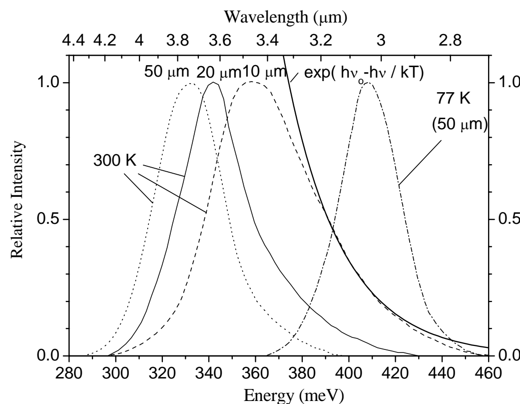


Figure 6. Emission spectra of three OP LEDs with 50, 20 and 10 μm thick n-InAs active layers at room and liquid nitrogen temperatures [14]

As seen from Figure 6 the peak wavelength shifts towards the high-energy region as the thickness decreases. The peak energy shift (from 334 to 365 meV, 300 K) is accompanied by a significant photoluminescence power enhancement and reflects the decrease of self absorption at short wavelengths. The similarity of the long wavelength shoulders, including the luminescence spectrum at 77 K, reflects the absence of absorption and is evidence for the distribution of carriers within the tail states inside the energy gap. In the 10-μm thick n-InAs slab at 300 K the high energy slope approaches the $exp(-h\nu/kT)$ value describing the carrier distribution inside the band, the latter being a key property of a properly designed LED/source.

In as much as InAs and InSb-like semiconductors are characterized by a large minority carrier diffusion length, (the value of L_p is of the order of 10-20 μm for pure ($n=10^{15}$ cm^{-3}) InAs samples [27] for example), the expected absorption in homojunction diodes is fairly strong with the result that external quantum efficiency is low. Carrier confinement of injected carriers in double heterostructures is, therefore, the most popular mid-IR LED/PD design feature.

2.2.3 Active Layer Mesa Diameter

There are many issues to be taken into account when the optimal LED active area dimensions are considered since each application calls for specific operating conditions, namely current pulse and duration. Maximum output power is achieved in a large diode with low local current density, while maximum brightness is achieved in a small LED.

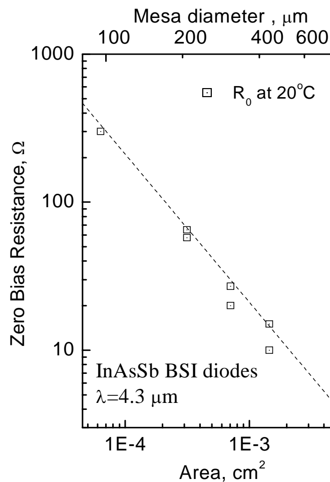


Figure 7. Zero bias resistance in InAsSbP homojunction diodes emitting at 4.3 μm *vs* mesa diameter [50]

Heatsink conditions and power consumption add uncertainties to the LED “ideal” mesa diameter definition. It seems that the photodiode has fewer such considerations since the highest signal/noise ratio is achieved with a circular mesa construction at

small current amplitudes which usually do not exceed the saturation current. It is a common feature that the forward bias (FB) current and the saturation current in reverse bias (RB) decrease linearly with mesa diameter. The latter is well illustrated in Figure 7, in terms of the dependence of zero bias resistance (R_o) on mesa diameter for InAsSb p-n junction diodes grown on to an InAsSbP buffer, as described in Section 2.1 and in Figure 3. The R_oA product amounts to ~0.02 $\Omega\cdot cm^2$ in the whole 100 - 300 μm diameter range, indicating negligible leakage current through the mesa surface. This dependence also shows that by selecting device diameter it is possible to adjust diode resistance to that of standard operational amplifiers that usually do not match the detector impedance [28].

2.3 Contacts

The room temperature internal quantum efficiency for mid-IR emitters in the 3-5 μm region is not as high as for their near-infrared counterparts, the low current values for the former being 24% [29] and 15%[30], 40 % [31] for p- and n-InAs devices respectively. The efficiency is much worse at high pumping levels due to fundamental reasons such as non-radiative Auger recombination, whose probability increases rapidly as the carrier concentration increases. To obtain high output it is tempting to operate the LED with the highest possible current amplitude, even though this will result in a reduction of quantum efficiency. At the same time high current amplitudes have a negative impact on the device performance due to the "current crowding" effect previously studied in a number of publications on near IR LEDs (see *e.g.,* [32]). The corresponding effect in the mid-IR was studied by luminescence power distribution measurements along the active surface of the structure with an infrared (IR) microscope in conjunction with an InSb thermal imaging camera sensitive in the 3-5 μm spectral range, using an appropriate pulse length for the IR camera recording time (40 ms) [33]. This permitted an IR power pattern study of the structure with a spatial resolution of 20 μm and a temperature sensitivity no poorer than 0.1°C.

Figure 8 presents direct evidence for the non-uniform current distribution in a 4.3 μm InAsSb LED composed of a 5 μm thick p-InAsSb(Zn) layer, an 80 μm thick graded band gap n-type InAsSbP "window" layer and a 350 μm thick undoped n-InAs (111) substrate. The p-InAsSb/n-InAsSbP structure was cleaved along the {110} planes into triangles with a side length of about 600 μm and soldered to a TO-18 header p-side down. Small area top contact of 50 μm in diameter was made in the centre of the triangle by indium soldering (see Figure 8 top). The reflective ring of the TO-18 header contributes to the concentration of light escaping laterally from the structure.

The middle graph in Figure 8 shows the distribution of the apparent temperature along the {111} LED surface heated up to 55°C in the a-b direction (that is, in a <011> crystallographic direction) in FB (I=100 mA, $\Delta t>0$) mode. The image and the graph reflect the drop of the signal at the contact (area denoted by the arrow C). It was verified that negligible heating is taking place in the diode so the "hot" image at C (Δt~5°C) is a "false" signal associated with the light reflected at the contact (we consider here that the radiation from the LED surface has a Lambertian distribution).

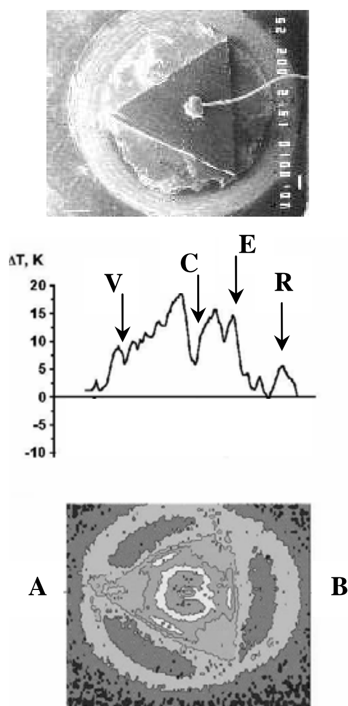


Figure 8. Photo of the InAsSb LED (*top*), IR maps of the 4.3 μm radiation emitted through the "window" layer at FB (I=600 mA) (*bottom*) and the ΔT values along the A-B direction with the following points of interest marked by *arrows*: V-vertex of the chip, C-point (dot) contact, E-edge of the chip and R-wall of the reflector ring of the TO-18 header (*middle*) [45]

As the internal quantum efficiency of InAs and related compositions is smaller than unity, it is reasonable to assume that no re-emission is taking place in the samples under investigation. Thus the radiation distribution should be very close to the spatial diode current distribution. The "dark" vertices of the triangle may be evidence for current crowding arising from bulk resistance of the 80-μm thick top n-InAsSbP layer. The decrease in current density through p-n junction areas located far from the contact is seen as a decrease in apparent temperature from 20 to 10°C in the direction C (contact)→V (triangle vertex). However, the middle areas of {011} edges close to the contact contain bright spots denoted by peak E in the "temperature" distribution. We attribute this peak to the radiation which was initially internally reflected at the surface towards the triangle cleaved edges. It is clear from the internal reflection considerations that the directions of major importance are perpendiculars from the contact to the triangle edges. Some weak radiation areas seen at the vertices (denoted by arrow V) support the above explanation, since the vertices should accumulate the radiation of the internally reflected light. Measured by a 3–5 μm infrared camera at forward current of 800 mA, the local effective "excess" emittance is 50 mW/cm^2 in the vicinity of the top contact and only 5 mW/cm^2 in the vertex at the "far end" of the structure. The above crowding results in the reduction of total (integrated over the whole surface) device efficiency

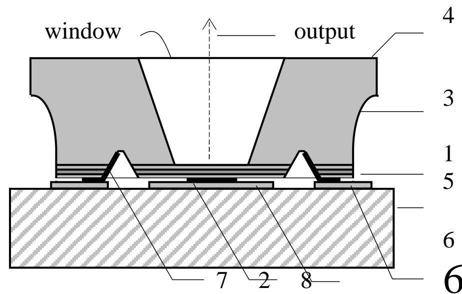


Figure 9. Schematic diagram of the flip-chip bonded Type 1 LED with a window above the anode where 1- is the double heterostructure with p-n junction (active layers), 2- is the Au round contact to p-InAsSbP confining layer, 3- are semi-cylindrical side walls, 4- is the InAs (111) surface, 5- is the Si sub-mount, 6- are Pb+Sn cathode contacts, 7- is the Au cathode contact, 8- is the Pb+Sn anode contact

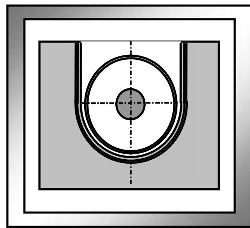


Figure 10. Bottom view of the "RC" InAs LED chip with 300 μm mesa and 100 μm circular Au anode (both are at the center) and U-shape cathode contacts. The "window" in InAs is on the opposite chip side (see Figure 9).

compared to one with uniform current spreading, since the conversion efficiency becomes surprisingly low in the areas of high current density. Thus it is evident that the contact area should be made as large as possible and/or grid- or ring-like contacts are to be used. As an example, the increase in power obtained by replacing the dot contact by a ring contact in a 3.3 μm LED device was as high as 8 [34].

Further improvements can be made with the use of a broad area contact together with a flip-chip package arrangement, since the contact can serve as an internal mirror for the generated photons [31,35,36]. Figures 9 and 10 describe such an InAsSbP/InAs LED with a mirror formed by the gold anode [37]. The InAsSbP/InAs DH structures were grown onto transparent heavily doped n^+-InAs(Sn) substrates ($n>10^{18}$ cm^{-3}) by the LPE method described previously with respect to backside illuminated photodiodes and LEDs emitting at 3.3 μm [18, 19, 38]. They comprise a 1-2 μm thick n-InAs active region grown from a Gd-doped melt and two 2-3 μm thick wide-gap $InAs_{1-x-y}Sb_xP_y$ layers ($0.05 \leq x \leq 0.09$, $0.09 \leq y \leq 0.18$) with Zn as a dopant for the cap layer ($p=2–5\times10^{17}$ cm^{-3}). The resulting heterostructures were processed by a two-side wet photolithography process into structures with a mesa diameter, D_m, of 300 μm and two non-alloyed gold circular contacts with diameters of either 240 or 100 μm. The contacts were evaporated and

then augmented by Au galvanic deposition. The cathode contact was “U” shaped as shown in Figure 10. To achieve the desired dimensions of the “windows” on the substrate side (see Figure 9), three processing steps were necessary: substrate thinning, creation of circular “wells” through the mask and finally selective etching to create the desired window width (a diameter of 100-140 μm). Structures were then cut into chips of 0.95×0.85 mm^2 dimensions and soldered onto 1.1×1.6×0.4 mm^3 semi-insulating silicon-sub-mounts at 180°C. The latter contained a "horse shoe" (or “U” shaped) Pb-Sn cathode contact matched to that described above. Finally, the Si sub-mounts with LED chips were soldered onto standard TO-18 headers.

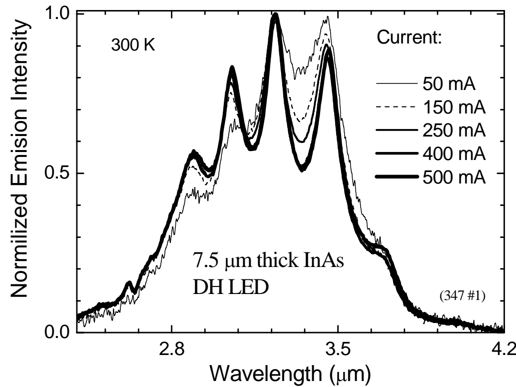


Figure 11. Room temperature electro-luminescence spectra of the “RC LED” at pumping currents in the 50-500 mA range

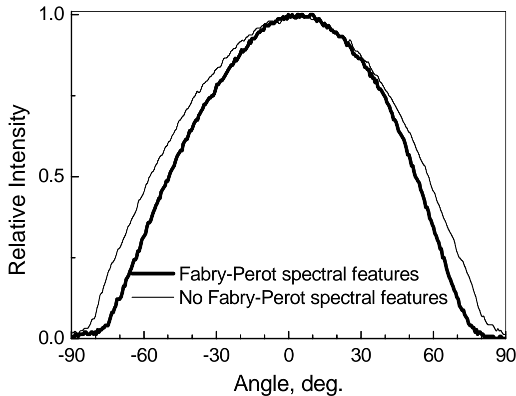


Figure 12. Far field pattern of two diodes from the same wafer with resonant (*thick curve*) and without resonant features (*thin curve*)

Electroluminescence measurements were made using 77 K CdHgTe or InSb photodiodes and a phase sensitive detection technique. The devices were pumped with pulse currents of 5 μs duration with a repetition rate of 500 Hz. Figure 11 presents typical emission spectra of a device with a 100-μm wide anode contact at pumping currents in the range 50-500 mA with clear evidence for Fabry-Perot modulation of the emission spectra by the resonator that is formed by the bottom anode contact and the top surface of the LED (see Figure 9). The spectrum envelope follows the temperature variation of the InAs energy gap (2.8×10^{-4} eV/K) and the thermalisation of carriers. The resonant mode spacing is directly proportional to the square of the wavelength ($\Delta\lambda=\lambda^2/2nL$, n=3.52) and at first glance it may seem that the device exhibits superluminescence at high pumping currents because of the narrowing of the resonant peaks. Moreover, when cooled down to 77 K the device lased perpendicular to the p-n junction at 2.95 μm at thresholds as low as 300 A/cm^2 with one dominant stable mode over the whole 1-2 A pumping current range [39]. However, the 300K light-current characteristic for this particular device follows a sub-linear dependence which is typical for all the published spontaneous LEDs in this spectral range. Without totally rejecting the possibility of superluminescence in these devices we have found a more simple explanation of the spectral features in our diodes. We believe that it is the aforementioned current crowding effect which narrows the peaks at high pumping currents. Since the window for the light escaping through the n-InAsSbP confinement layer (see Figure 9) is slightly wider than the contact, two areas of excess carrier recombination can be considered: one adjoining the circular contact and the other adjacent to the ring contact area. The latter "ring" part is expected to have weakly pronounced resonant features because of the lack of gold at the bottom LED surface in the corresponding area. The resulting total output is the sum of the spectrally modulated and unmodulated radiation; at small forward bias the brightness of the above two contributions are nearly equal. As the current is increased, current crowding occurs above the anode rendering negligible the contribution of the "ring" area. This results in an increase in brightness of the spectral peaks whose spectral separation agrees reasonably with the device parameters: n=3.52, L=7.5 μm.

It should also be noted that the device performance could be improved by decreasing the resonator length or the structure thickness, due to resonant cavity effects and the reduction of optical losses. The corresponding increase in output power is accompanied by the narrowing of the far field pattern as shown in Figure 12. The values of the resonant quality factor, *Q,* estimated from the mode linewidth were as high as 24, 32 and 80 for layers of 7.5, 13 and 45 μm thickness respectively. In addition we note that the radiation at the resonant maximum was at least 25% more intense than the corresponding value of the "non-resonant" device. Gevaux *et al.* [36] reported much stronger enhancement of the resonant peaks in their 3.3 μm MBE grown devices.

From a mechanical point of view, the proposed LED design is not ideal due to the large lateral dimensions of the active region. The contact diameter or "window part" of the device is much greater than its thickness: the ratio of the two is more than 15. However, we experienced practically no occurrence of fractures in the vicinity of the mesa edges, the thinnest device regions. This may be an indication of negligible internal stresses due to matching of the lattice and thermal expansion

coefficients in the InAsSbP/InAs system, as well as the consequence of elevated microhardness of the InAsSbP quaternary [10].

3 Choice of the operating mode

Several practical applications require mid-IR (3-6 μm) devices to operate at elevated temperatures (>80°C). Recently it was shown that InGaAs and InAsSb LEDs, emitting at the important analytical wavelengths of 3.3 and 4.3 μm respectively, exhibited "classical" p-n junction behavior up to 180°C with negligible power degradation during an operating period of 1500 hrs at elevated temperatures [40]. However, LEDs have a high value of temperature quenching coefficient and thus they exhibit poor conversion efficiency at elevated temperatures, due to both Auger recombination and high n_i magnitudes and so there is some pessimism regarding high temperature LED operation at wavelengths longer than 4 μm. For this reason, long wavelength devices often incorporate thermoelectric coolers for both LED and photodetector elements [28]. This increases the S/N ratio, although the solution is not appropriate for many practical applications.

On the other hand, a narrow gap p-n junction under reverse bias can absorb radiation; the latter is equivalent to producing negative luminescence (NL). The negative luminescence effect in semiconductors has been known since 1965, when Ivanov-Omskii and his colleagues [41] at the Ioffe Institute observed that the intensity of emission from one of the surfaces of an InSb slab subjected to crossed electric and magnetic fields was lower than the corresponding equilibrium value. Extraction of the charge carriers from the regions adjacent to the junction leads to a similar effect, that is, a reduction of their density below equilibrium values so that $(np)<(n_o p_o)=n_i^2$ (here n_o, p_o, n and p are equilibrium and non-equilibrium densities of electrons and holes, respectively; and n_i is the intrinsic charge-carrier density). Due to the extraction of charge carriers, the recombination-radiation intensity, P, decreases and becomes lower than the thermal background level, P_o, and absorption dominates over emission; this means that the luminescence appears to be "negative": $\Delta P = P - P_0 = P_0(n \cdot p / n_i^2 - 1) \leq 0$. A negative luminescence efficiency, that is, the ability to suppress the blackbody radiation at specific wavelengths, as high as 0.9 has already been reported by several authors [42, 43] and effective cooling, Δt, ranges from -25 K [42] to -54 K [44].

From the standpoint of a spectroscopic user, there is little difference between the NL and PL, save the sign of ΔP, since in most cases their emission spectra are the mirror images of each other. Moreover, in the presence of a cavity, *e.g.* Fabry-Perot resonator, the NL emission spectrum is modulated by the interference peaks in a similar way to the PL spectrum (see *e.g.* the 190°C NL spectrum by Remennyi *et al.* [35]).

In the case of transparency measurements, the existence of gas with a specific absorption band in the optical path between a forward biased LED and detector results in a decrease of the detector signal (so the gas is "cold" and thus absorbs radiation), while the same gas in the path of a reverse biased device and detector produces an increase of the detector amplitude. The latter is equivalent to gas emission measurements, with a note that the gas may be at room (ambient)

temperature and the zero level (without gas) corresponds to a point below ambient temperature. The most frequently met examples are emission spectra around 4.3 µm, which is in the region of a strong absorption of CO_2 gas. The presence of CO_2 leads to the absorption dip in PL spectra in forward bias (FB) and a "decrease" of NL emissivity under reverse bias (RB) [see *e.g.* 45]. The comparison of spectroscopic measurements using "positive" and "negative" radiation beams was considered by Stepanov [46].

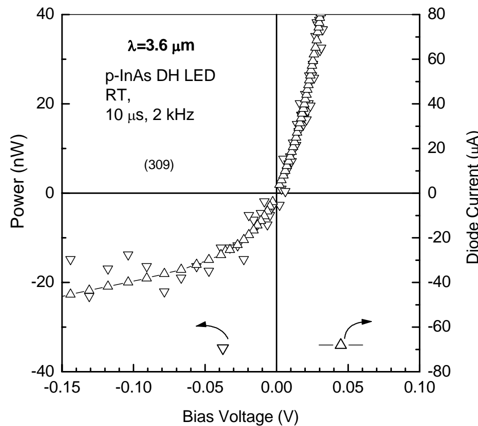


Figure 13. Output power (*triangles down*) and current (*triangles up*) *vs* bias voltage in the epi-side down bonded p-InAs DH LED. Reproduced in accordance with permission from Taylor & Francis Ltd. (http://tandf.co.uk/journals).

The suppression, or quenching, of positive luminescence power (PLP) with the temperature/wavelength increase is associated with Auger processes [47] and the ability of the system to restore the equilibrium population, which has been shown previously for the two level system from a thermodynamic balance analysis [46]. A simpler phenomenological analysis of the PLP quenching is based on the following assumptions: a) the negative luminescence power (NLP) saturates at some current, $I_{saturation}$, b) at low currents the experimental L-I curves are linear for both FB and RB currents, that is, NL and PL have the same conversion efficiencies (CE) [see *e.g.* 48]. The latter follows from the observation of similar slopes of output power and current dependencies on the applied voltage as shown in [49,42] and Figure 13. Taking into account the above assumption, we can state the following "rule of thumb" for low current injection: 1) the output power is proportional to the current regardless of the direction of the latter, 2) the CE is a ratio of NLP and saturation current, that is, (NLP/I_{sat}).

Figure 14 shows the variation in saturation current (I_{sat}, triangles up) *vs* temperature over the range 20-200°C for a DH LED, with the p-InAs active layer made in a similar manner to the LEDs described in Section 2.3. Also shown are the NLP (triangles down), NLP/I_{sat} (solid line) and experimental PL values of the conversion efficiency at I=10 mA (open circles) and I=1A (solid circles) respectively. The saturation current follows the Shockley temperature variation

defined by the thick solid line. As seen from Figure 14, the expected LED power variation (or CE at I=1A) is close enough to the data on PLP at temperatures below 120ºC. Above 120ºC the PLP quenching is enhanced. The CE_{PL} values at I=1A are well below the values at I=10 mA reflecting the sublinear current dependence of the PLP. It should be noted that the PLP is partly affected by the n^+-InAs substrate because the transparency decreases with increasing temperature. This can contribute to fast PLP fading at t>120ºC.

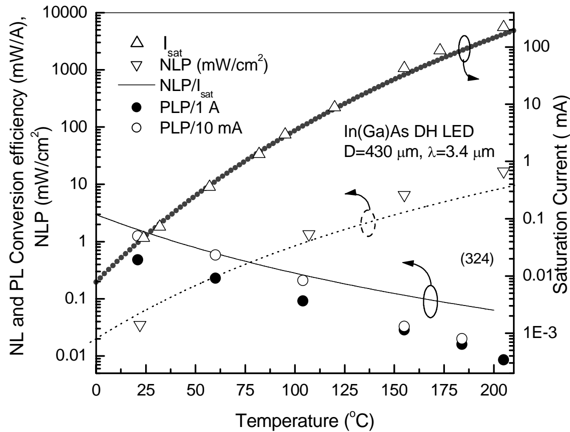


Figure 14. Temperature dependence of saturation current (*thick line*), NL power (*dotted line*) and conversion efficiency at I=10 mA (*open circles*) and I=1A (*solid circles*) correspondingly. Thin line presents the NLP/I_{sat} ratio. Reproduced in accordance with permission from Taylor & Francis Ltd. (http://tandf.co.uk/journals).

It follows from the above considerations that the diode with the lowest dark current (that is in fact a high performance photodiode) exhibits an excellent conversion efficiency (NLP/I_{sat}) in forward bias, *i.e.* it can also be used as an LED. In our case, dual use of the diode is justified by the low series resistance and broad contact geometry typical of the flip-chip construction on n^+-InAs. Indeed DH InAs diodes are characterized by fairly good detectivity ($D^*_{3.3\ \mu m} = 2\times10^{10}$ cm $Hz^{1/2}W^{-1}$) [19] and high output (0.5 mW/A). Similarly graded band gap InAsSbP LEDs emitting at 4.3 μm can be used as PDs, *e.g.* for CO_2 gas sensing [50].

It is clear that for each device there is a “crossover temperature” that indicates the starting point of the NLP superiority. Rough estimations based on experimentally observed LED performance showed [45] that the expected crossover temperature ranges from ~60°C for 5-6 μm LEDs to 180°C for 3.3 μm LEDs.

Figure 15 demonstrates the temperature dependence of PL and NL emission spectra of a 3.9 μm LED at saturation currents of 0.05, 1, 18 and 185 mA at 25, 90, 155 and 195°C respectively. (The 300 μm wide DH flip chip InGaAsSb/InAsSbP LED structure was grown onto an InAs substrate as described elsewhere [35]) . It can be seen (Figure 15b) that the NLP and PLP (integrated over the spectrum) are the same at room temperature, but at elevated temperatures the superior performance of the NLP (by a factor of almost 3) is evident as the PLP exhibits “saturation”.

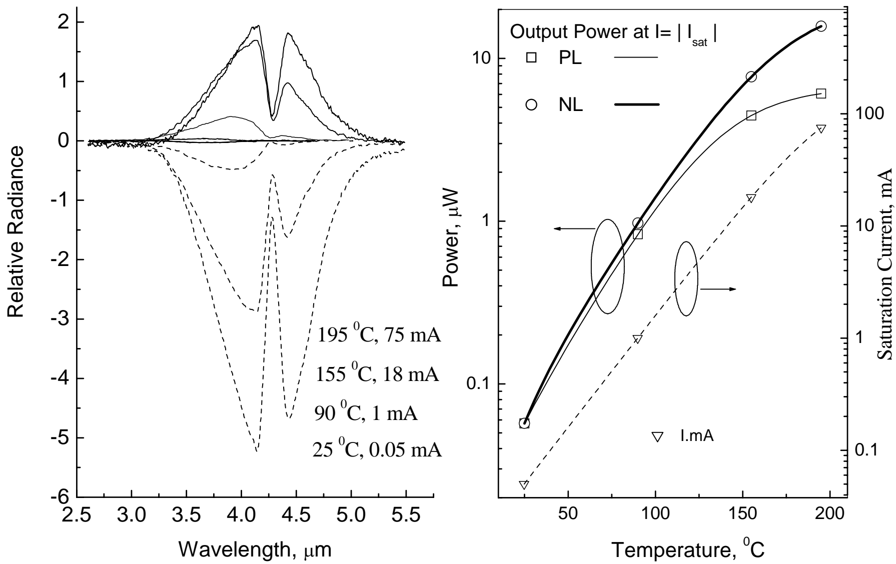


Figure 15. (**a**) PL and NL spectra of 300-μm wide InGaAsSb LED emitting at 3.9 μm and (**b**) NLP, PLP at saturation current and saturation current *vs* temperature

The superior conversion efficiency (CE) of the NLP over the PLP is typical for long wavelength mid-IR LEDs [18,39,43]. The reverse bias mode also enhances the performance of the narrow band detectors for the same reason: suppression of Auger recombination by decreasing the carrier concentration [51]. It seems that the reverse current in InAs-like narrow gap uncoated devices is not strongly effected by surface leakage due to the high conductivity of the p-n junction at elevated temperatures. Thus surface treatment/passivation is not critical for high temperature device performance in as much as saturation of the reverse current is already very well pronounced at t>60ºC while zero bias resistance varies exponentially with reciprocal temperature in most cases.

Another positive aspect of the NL mode arises from the increase in reverse-biased p-n junction effective resistance with regard to other LED layers and from the subsequent elimination of current crowding. Thus, less effort can be devoted to the contact geometry design in the NL devices. The uniform current spreading in point contact structures has been simulated for an InAsSbP diode under reverse bias and was confirmed by the 2D measurements of NLP distribution using a thermal vision infrared (3-5 μm) camera [33]. Figure 16 demonstrates another example of the 2D distribution measured at 70ºC for a p- InAsSbP/n-InAs photodiode. A schematic diagram of the 430 μm diameter mesa with a circular top contact of diameter 150 μm is shown at the top of the figure. The basic parameters are given in [45]. Esina [52] and Krier [53] have shown that in p-n junctions based on InAs and related compositions, current flow is governed by a generation-recombination mechanism (in the space charge region) at low temperatures only. The recombination in the p- and n- regions ("diffusion" mechanism) dominates at T>200 K. Indeed, the R_oA

product varies linearly as $1/n_i^2$ [45], where n_i is the intrinsic carrier concentration for an energy gap value of E_g=350 meV. This is an indication of the good quality of the p-n junction (*i.e.* negligible tunnel and leakage currents) and the high efficiency of the minority carrier extraction in a reverse biased diode. The NL, expressed as an apparent temperature decrease along the device base, was induced by reverse bias (I= -100 mA) with pulse duration of 40 ms. Note, that the figure maps out a false colour image from which the background image (t=70ºC, zero bias) was subtracted. Thus, the image represents the values of modulated apparent temperatures of about ΔT< -5ºC (or the NL power "contrast") produced by carrier extraction from the vicinity of the p-n junction.

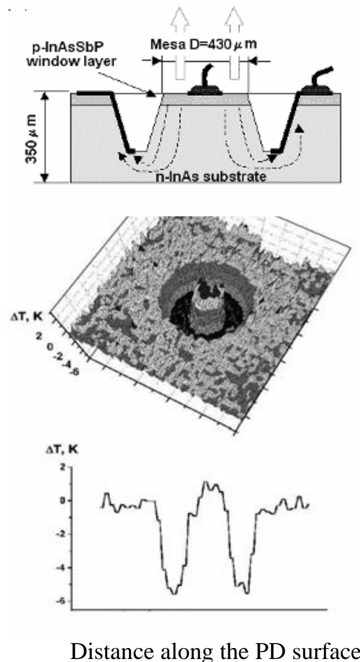


Figure 16. p-InAsSbP/n-InAs NL IR image (*top*) and apparent temperature distribution over diode surface (2D map (*middle*) and Δt values (*bottom*))

Let us point out once more the uniformity of the 2D power distribution, that is clearly seen in large area NL devices and originates from the elevated value of p-n junction resistance with respect to the FB magnitude [33] (see also Figure 1 in [35] for the 2D NL distribution in InGaAsSb LED). The uniformity of the reverse bias current spreading allows simplification of the NL device design and the use of point contact bonding in many cases, *e.g.* for fabrication of large area emitters.

Creation of the non-equilibrium state can be arranged by any kind of pumping: electrical injection/extraction in a p-n junction or device heating being only few of the possible examples. Due to experimental convenience, the most frequently used method to break thermal equilibrium is optical pumping with photons whose energy

exceeds that of the "phosphor": the method is also known as photoluminescence. Several reports have been published on room temperature photoluminescence of mid-IR III-V materials including those of n-type conductivity [20,21] and a number of optically pumped (OP) double heterostructure mid-IR lasers and near IR LEDs have been investigated (see, *e.g.*[54, 55]). However, to the best of our knowledge at the start of our work only polycrystalline IV-VI OP LEDs have been realized [56] and no attempts have been undertaken so far to create practical mid-IR III-V OP LEDs operating at ambient temperature.

In 2000 [57] we proposed the use of an intermediate layer of high refractive index (*e.g.* chalcogenide glass, n=2.4–2.8) between that of the near infrared pump and the mid-IR "phosphor", in order to increase the out-coupling of NIR radiation into the "phosphor". Several immersion lens constructions incorporating InGaAsSb/InAs, InAsSbP/InAs heterostructures and InAs or InSb plates exhibited output powers from 2 to 0.1 mW/A in the 3.3–7 μm spectral range [58], close to those of conventional (electrically pumped) LEDs. Moreover, many of the disadvantages of narrow band p-n junctions such as high tunnel and surface leakage currents are easily eliminated in a construction with separate broad band injection ("pump") and narrow band ("phosphor") recombination regions. In addition, there is a significant reduction of the device cost due to the omission of such tedious procedures as the formation of contacts and mesa architecture in a narrow gap wafer.

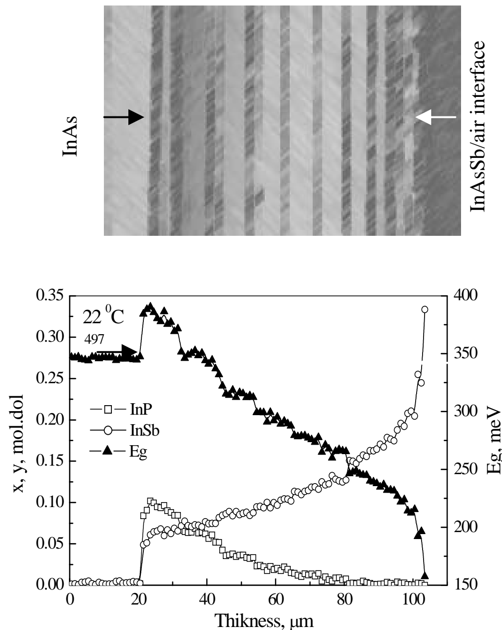


Figure 17. Photo of the (1$\underline{1}$0) cleaved edge of the InAsSb/…./InAsSbP/InAs multilayer structure (*top*) and distribution of InSb and InP content along the growth direction together with interpolated value of the energy gap (*bottom*)

Taking into account these considerations and our previous observation of reasonable InSb OP LED operation, it is obvious that the above OP LED design is both important and convenient for the narrowest energy gap devices whose growth and doping are usually not easy. Results for an OP LED emitting at 9 μm are presented in Figure 17 and 18 [59].

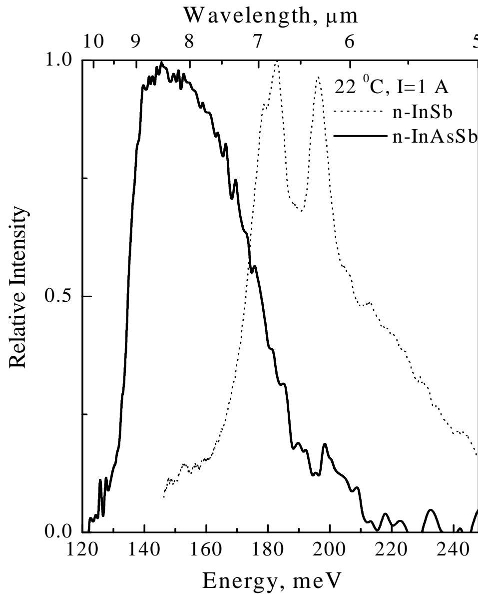


Figure 18. Room temperature Emission spectra of InSb (*dotted line*) and InAsSb (*solid line*) OP LEDs

The OP LED construction did not differ from that reported previously [14, 57, 58] and included an $InAs_{1-x}Sb_x$/…./$InAs_{1-z-y}Sb_zP_y$/InAs "phosphor" with the $InAs_{1-y}Sb_y$ side attached to a free GaAs LED surface ($\lambda \approx 0.87$ μm, $\eta_{ext} \approx 8$ %) using a chalcogenide glass with a refractive index of 2.4. The GaAs pump was activated at 2 kHz, with 15 μs pulses, and the photons generated in the narrow band part of the "phosphor" and which escaped through the InAsSbP window, were recorded by a CdHgTe cooled photodiode with maximum sensitivity at a wavelength of 9.4 μm.

Figure 17 shows a photo (top of the figure) of the cleaved and etched $(1\underline{1}0)$ plane of the 80-μm thick $InAs_{1-x}Sb_x$/…./$InAs_{1-z-y}Sb_zP_y$/InAs "phosphor" grown by the step-graded LPE method together with the chemical composition distribution along the growth direction (bottom of the figure). It is apparent from Figure 17 that the composition of the buffer (window) $InAs_{1-z-y}Sb_zP_y$ layers exhibits a step-wise character with an increase in Sb content towards the surface and a rapid decrease in the interpolated energy gap yielding an energy band offset, ∇E_g, of 2 and 11 meV/μm at the heterojunction and at the narrow band surface respectively. The latter value with a corresponding efficient optical thickness of about a fraction of a micron favors long wavelength light propagation towards the InAs substrate as considered previously in Section 2.1. It is worth mentioning that poor crystalline

quality arising from a high lattice parameter gradient is expected to decrease the conventional LED efficiency due to leakage through the dislocations "connecting" p- and n- layers of the structure. In our case the situation is not so dramatic in view of the lack of a p-n junction in the narrow gap "phosphor". Moreover, there is no direct evidence in the literature for non-radiative recombination on misfit dislocations in InAs-like materials possibly because it is generally assumed that the defect related levels in InAs are concentrated in the conduction band and not in the middle of the gap [60]. The OP LED construction is free of the above disadvantages due to the p-n junction formation in a wide gap material

Figure 18 presents the room temperature emission spectra of InSb [58] and InAsSb OP LEDs, the former being taken as a reference. The dip at 6.5 μm in both spectra is caused by the absorption of water vapor while the sharp long wavelength edge in the InAsSb OP LED relates to the decline in detector sensitivity at long wavelengths. The InAsSb emission maximum is a little smaller than the interpolated value of 150 meV, obtained through the use of known dependencies [20] which may reflect Joule heating of the "phosphor" by the GaAs pump. Nevertheless, the conversion efficiency was as high as 10 μW/A: a value considerably higher than in QC LEDs [61].

4 Out-coupling of Radiation in Mid-IR Devices

The relatively low quantum efficiency of the semiconductor mid-IR (3-5 μm) devices at room temperature makes the collimation/outcoupling of the output very critical for use in the field, *e.g.* for fiber optic applications. A^3B^5 resonant cavity (RC) mid-IR light emitting diodes (RC LEDs)/photodiodes with directed beam are currently in the research stage [62] and thus more traditional collimation techniques are still of significant importance. A narrow far field pattern for conventional LEDs can be achieved by surface patterning [63] or by other means, *e.g.* by the formation of an optical concentrator, such as a Winston cone, in the substrate material as an integrated part of a heterostructure [64], as a reflector attached to the emitter holder or as an immersion lens attached to the LED/detector outcoupling surface [65].

Strong nonradiative Auger processes that are active at high excess carrier concentration means that low density current/large active area dimensions are the common requirements for mid-IR LEDs. For example, practical InSb LEDs pumped with I=IA employ chips of lateral dimensions of 3 mm [28]. For this reason, Winston cone geometry cannot meet the requirements for high mid-IR output at a low device cost in view of the high cost of semiconductor material whose thickness is to be several times of that of the mesa diameter. However, a Winston cone structure is well matched to photodiode or NL devices, whose lateral dimensions and operating current are small [44]. Conventional reflectors, such as aluminum parabolic reflectors, are cheap, however they do not affect the internal reflection of the radiation and the external efficiency is fairly low.

The immersion lens technique has been used successfully in MIR devices to reduce internal reflection losses, to achieve a directed beam and to reach mW power levels in LEDs [18], and also to increase the D^* value in detectors [25]. This technique involves the attachment of an aplanatic hyperhemisphere [65] to a semiconductor device using a high refractive index chalcogenide glass as a "glue".

However, this technique has not, so far, been implemented in preference to more simple and cheap constructions, *e.g.* to the ones employing ball lenses.

Below, we describe InAsSbP/InAs DH and InGaAsSb/InAsSbP DH LEDs grown onto n^+-InAs (111) substrates which are coupled with spherical (or nearly spherical) lenses and emitting in the 3.3-3.9 μm range at room temperature. The heterostructures, similar to those described in the previous Sections, were processed using a two-side wet photolithography process into structures with a mesa diameter of 300 μm and non-alloyed gold circular contacts of 240 or 100 μm. To reduce the radiation losses, "windows" or thin regions were formed above the mesa as shown in Figure 9 and 19. The windows were etched in a three-stage process: substrate thinning, creation of circular "wells" through the mask and finally selective etching to obtain the desired window width (100-140 μm) and depth. Chip separation was simplified by chemically etching a rectangular network of 20-50 μm deep grooves between each chip on the epi-side of the wafer. Structures were then cut along the grooves into chips of 0.95×0.85 mm^2 dimensions which were soldered at 180ºC onto 1.1×1.6×0.4 mm^3 semi-insulating Si-sub-mounts containing the Pb-Sn cathode and anode contacts. Finally, the Si-sub-mounts and LED chips were soldered onto a standard TO-18 header.

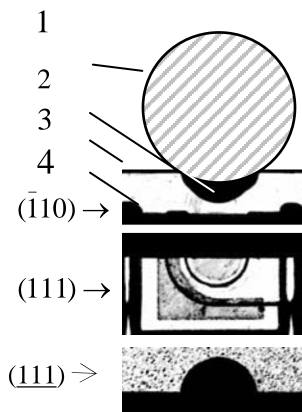


Figure 19. Photos of the three surfaces of the cleaved InGaAsSb LED sample: $(\bar{1}10)$ plane or the cross Section of the LED wafer with ball lens (1) (drawing), 2- "window" or the well filled with chalcogenide glass (drawing), 3- n-InAs surface, 4- grooves for chip disassembling /cutting, (111)- the episide surface cleaved through the centre of the mesa and anode contact, $(\bar{1}\bar{1}\bar{1})$ – the substrate side of the wafer with half of the circular well (window)

The 3.3 μm LEDs were characterized by zero substrate thickness in the region above the mesa as in Figure 9 while the InAs substrate thickness in the area adjoining the mesa of the 3.9 μm LEDs was in the 30-50 μm range. Figure 19 presents photos of three surfaces of the latter structure cleaved along the $(\bar{1}10)$ plane that passes through the mesa centre: the (111) episide surface with printed contacts (half of the circular mesa and U-contact), the cleaved $(\bar{1}10)$ plane or cross-section and finally the substrate side (the $(\bar{1}\bar{1}\bar{1})$ plane) with half of the circular "window".

The LEDs were fitted with ~1 mm wide lenses fabricated from sapphire (Edmund Scientific) or chalcogenide glass. In the latter case, the lens was moulded

onto the flat LED surface without an etched "well" (window): the lens shape being controlled by the surface tension, temperature and the amount of glass [66-68]. The former case is illustrated in Figure 19 where the patterned circle represents the sapphire ball glued to the LED surface by a chalcogenide glass. The photos of both LED types are presented in Figure 20.

Figure 20. Photos of the LEDs with chalcogenide (*right*) and sapphire (*left*) lenses mounted onto a TO-18 header

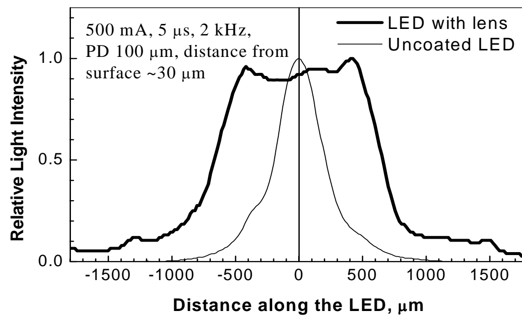


Figure 21. Distribution of radiation along the chip surface in 3.3 μm InAs LEDs; with a flat surface (*thin line*) and immersed into 1 mm chalcogenide glass (*thick line*)

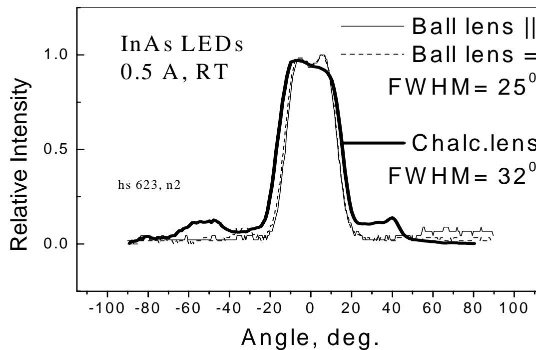


Figure 22. Far field pattern of the emitted radiation from a sapphire ball lens LED in two perpendicular planes (*thin and dashed lines* denoting "||" and "==" respectively) and from the chalcogenide glass lens LED (*thick line*)

The radiation distribution in the vicinity of the LED was measured by a movable fiber probe in a similar manner to the technique described by Remennyi [35] and Malyutenko [69]. Figure 21 shows the “near field” radiation patterns from two InAs LEDs at room temperature: one with an uncoated flat surface and one with a chalcogenide lens. As seen from Figure 21 the initially narrow near field pattern is transformed by the lens into a broader distribution. The existence of the plateau in the distribution is evidence for the collimation of radiation, which is also confirmed by the direct far field measurements shown in Figure 22. We believe that the shape of the chalcogenide lens is close to an aplanatic hemisphere since the power increase factor associated with the lens incorporation for the best devices was around 5, the latter figure being very close to the figure of merit for the immersion lens construction (n^2, n=2.4) [65]. Due to the uncertainties of the melting process the lens shape was not very reproducible and some de*via*tion in the far field pattern from one LED to another along with a non-symmetrical beam cone is inevitable.

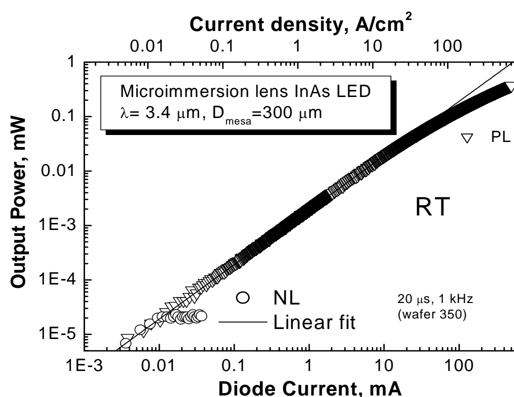


Figure 23. Pulse output power vs current in InAs LED in forward (*triangles down*), “PL” and reverse (*circles, “NL”)* bias

However, the use of sapphire lenses (with the given dimensions) together with the fixed “window” position relative to the LED emitting area enabled very similar far field patterns to be obtained in a set of LEDs fabricated from the same wafer. As an illustration, in Figure 22 we show nearly identical far field patterns of output in the ball lens LED measured in two perpendicular planes. The superior geometrical characteristics and stability of the ball lens construction with respect to temperature variations are offset in practice by the power out-coupling efficiency of the sapphire ball lens (n=1.5) LED which is several times lower than that of the chalcogenide lens construction with n=2.4. Figure 23 shows the power output *vs* current of the chalcogenide glass lens InAs LED at room temperature under forward- and reverse-bias. As seen from Figure 23 the output grows sub-linearly as the current increases: at small currents the conversion efficiency is initially equal for positive and negative luminescence with a value of 1.62 mW/A which falls to 0.72 mW/A at a forward bias current of 500 mA. The decline of the conversion efficiency at high currents in our case is not entirely related to Joule heating, since the peak position of the LED

emission is not effected by the current. This is supported by Figure 24, with two similar spectra recorded at forward currents of 10 μA and 500 mA correspondingly.

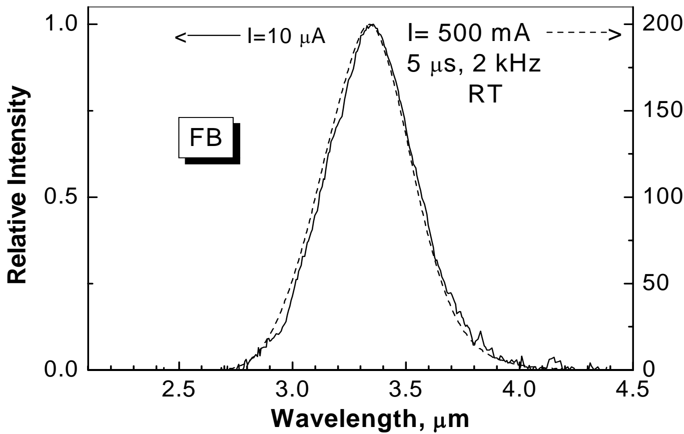


Figure 24. Emission spectra of InAs LED at low (10 μA) and high (500 mA) pulse pumping at 300 K

Thus, the implementation of the immersion lens technique, including ball and aplanatic hyperhemipshere optical elements attached to a diode by a chalcogenide glass, may both collimate the output and increase the output power by several times. The same improvements are naturally observed in the performance of the detector.

5 The Use of Diode Opto-pairs for Chemical Sensing

The selection of a particular flip-chip pair for chemical sensing should be based not only on the device parameters, which are presented in Table 1 and usually specified at room temperature, but also on other issues, particularly the operating current and temperature. The latter is a consequence of the dependence of device efficiency and emission/sensitivity peak position on the operating temperature (see, *e.g.* 70). Figure 25 presents the emission spectra of the 7.5-μm thick InAs DH LED, similar to that described in Section 2.3, in the 77-573 K temperature interval. The peak position with the kT/2 correction corresponds to the known InAs energy gap temperature dependence: Eg=415-0.276×T2/(T+83), while the short wavelength shoulder reflects the electron distribution in the conduction band: I~exp(-hν/kT). It is apparent that the same InAs diode could easily be used for H_2O, CH_4 and CO_2 detection since the emission peak overlaps with the corresponding absorption bands at 77 (3 μm), 300 (3.3 μm) and 573 K (4.3 μm) respectively. On the other hand, if necessary the proposed LED/PD construction permits the addition of a filter attached to the diode surface, which will diminish the spectral temperature variation of both the PL and NL spectra [35].

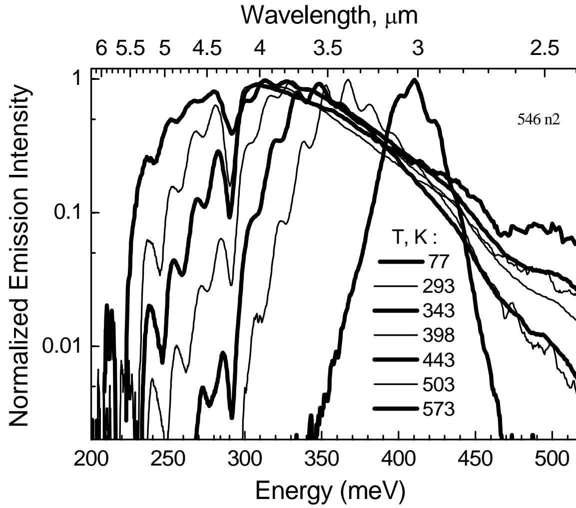


Figure 25. Emission spectra of the forward biased ("positive luminescence") 7.5 μm thick LED at 77, 293, 343, 398, 443, 503 and 573 K [39]

Table 1. (a) Typical parameters of flip-chip mid-IR LEDs grown onto n-InAs

λ, μm	Material	D, mm	t, °C	Low current CE, mW/A	High current CE, mW/A	Maximum pulse output, mW	Immersion lens	Ref.
2.8	InGaAs	0.3	18	0.35	0.5	0.5	No	[70]
2.9	InAsSbP	0.3	18	0.65	0.3	0.3	No	[70]
3.3	InAs	0.3	20	1.62	0.72	0.4	Yes	-
4.3	InAsSb	0.3	20	-	0.12	0.12	Yes	-
8	InAsSb (OP LED)	0.3		-	0.01	0.01	No	-

Table 1. (b) Typical parameters of flip-chip mid-IR photodiodes grown onto n-InAs

λ, μm	Material	D, mm mm	t, °C	$D^*_{\lambda\ max}$, cm $Hz^{1/2}W^{-1}$	$S_{U\ \lambda\ max}$, V/W	$S_{I\ \lambda\ max}$, A/W	R_o, Ω	R_oA, Ω cm^2 Ω cm	Len s	Ref.
2.8	InGaAs	0.3	18	6.23×10^{9}	1900	0.46	4100	2.90	No	[70]
2.9	InAsSbP	0.3	18	1.54×10^{10}	3620	1.49	2430	1.72	No	[70]
3.3	InAs	0.3	20	$1.63\ 10^{10}$	2864	1.79	1600	1.13	No	[19]
4.3	InAsSb	0.3	20	1×10^{9}	15	1.5	10	0.007	No	-

Conventional specifications for the spectroscopic applications of the optoelectronic components usually include the need for maximum possible signal/noise ratio accumulated within a specific time interval, say, 1 or 10s, depending on the ability of the analyte to change its composition within the measuring cell. In other words, the optically coupled diode pair should deliver to the detector the maximum possible power within the measuring time interval and also have the ability to register the emitted power with the least noise. It is hard to imagine that the measuring system could operate in continuous (CW) mode since the environment produces mid-IR background radiation whose intensity is often close to that of the LEDs and is subject to variations. Therefore, pulsing or background subtraction is normally required. In addition, pulsed current amplitudes, and hence the output power, are usually higher than those for the CW mode due to the “rest” periods when the Joule heat is dissipated at the heatsink and no additional energy is consumed from the power source.

On the other hand the L-I characteristic of most LEDs is sub-linear and thus it is clear that maximum power integrated over a certain interval could be delivered at low current and low duty cycle, say, close to the CW mode. A compromise between the two above arguments is the quasi-CW mode with a 50% duty cycle which enables the highest S/N ratio to be obtained in most cases. The pulsing frequency is usually optimized with respect to the amplifier frequency bandwidth and detector noise. Using the S/N ratio we can estimate the limit of detection (LOD), that is, the lowest gas concentration that could be detected by a diode optopair whose emission/sensitivity spectrum matches the absorption band of a gas. Taking into account Beer’s law ($U = U_0 \exp(-\alpha \cdot Lc)$) adapted for small values of exponent argument and assuming that the detector signal is directly proportional to the LED current in the quasi-CW mode, we can roughly estimate the LOD as

$$LOD = (Lc)_{test} / (\frac{\Delta U}{U} \cdot SNR \cdot t^{1/2} \cdot I) \tag{1}$$

where, $\Delta U / U$ is the signal decrement caused by the presence of a gas/liquid with concentration c_{test} in an optical path length L_{test} between the LED and photodiode and *SNR* is the signal-to-noise ratio determined for a given amplifier with bandwidth Δf, (for the unit time interval t (1 s) and unit LED current I (1 mA)). The $\Delta U / U$ value relates to the ability of the analyte to absorb within the pair sensitivity spectrum and can be estimated by calculating the overlap integral of the sample gas (or liquid) transmission spectrum and the product of the LED emission spectrum and the sensitivity of the photodiode. The typical width of the LED spectrum is around 0.1 λ_{max} and is broader than many gas bands and thus the $\Delta U / U$ value is usually not very high, the $\Delta U / U_{CH4}^{InAs}$ of 0.027 being a typical magnitude for methane gas with an InAs DH LED-PD pair.

An LED/photodiode equipped with an aplanatic hyperhemisphere [65] whose diameter is at least 10 times larger than the mesa diameter (see Figure 26) is the most appropriate option that enables the highest SNR values to be attained since the narrow far field pattern allows efficient coupling simply through the use of a spherical mirror. Estimates of the SNR for an InAs diode pair coupled in this way with a spherical mirror operating at 20 ºC give around ~4000 (Δf=1 MHz) with an LED current as small as 20 mA [50]. The corresponding LOD for methane gas is

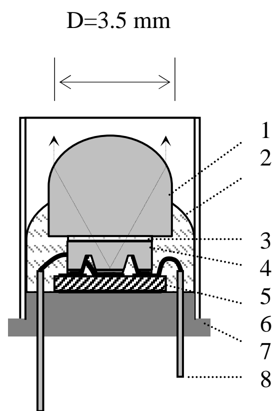


Figure 26. Construction of the immersion lens LED/photodiode. 1-CdSb or Si immersion lens D=3.5 mm, 2- epoxy, 3-chalcogenide "glue", 4 –"flip chip bonded" diode, 5-Si submount, 6-TO-39 header, 7-cathode, 8-anode [19]

calculated as 18 ppm·cm·mA·s$^{1/2}$, which means one can be optimistic about the use of these diode pairs in low voltage, low power, portable instruments.

Interaction with the sample under investigation could be realized through the absorption of the evanescent field, which exponentially decreases in the vicinity of the optical fibre surface when in contact with the analyte. The penetration depth is of the order of the wavelength and is given by, $d_p = \dfrac{\lambda}{2\pi\sqrt{n_1^2 \sin^2\theta - n_2^2}}$, where θ, n_1, n_2 are the angle of incidence, the refractive index of the fibre and ambient medium (with $n_1 > n_2$) respectively. Being an integral part of the radiation travelling along the fibre axis the evanescent field is sensitive to the presence of an absorber at the fibre surface. Provided the number of internal reflections is sufficient the output from the end of the fibre depends on absorber concentration and fibre length. Fiberoptic Evanescent Wave spectroscopy (FEWS) has numerous applications including detection of flammable gases [71] and analysis of hydrocarbons in water [72, 73] however only a few instruments employing LED sources operating in the mid-IR region have been reported. This is partly explained by a low coupling efficiency due to the non-optimized arrangements used so far.
Direct "gluing" of the fibre end to an LED surface by polymer compounds is often used for devices operating in the near IR region. However, these polymer compounds absorb radiation in the 3.3–3.4 μm range and thus there have been no reports on direct coupling of a mid-IR LED with a fibre. Here we report on the coupling of a mid-IR LED and fibre and the first experiments on liquid sensing using mid-IR A^3B^5 LEDs.

Transparent chalcogenide glass with a refractive index of 2.4 was sandwiched between the free surface of an InAs DH LED and a 300 μm thick Saphikon sapphire fibre [74] without protection layers (the construction is similar to that presented in Figure 26 except the lens is substituted by a fibre with the flat end attached to the

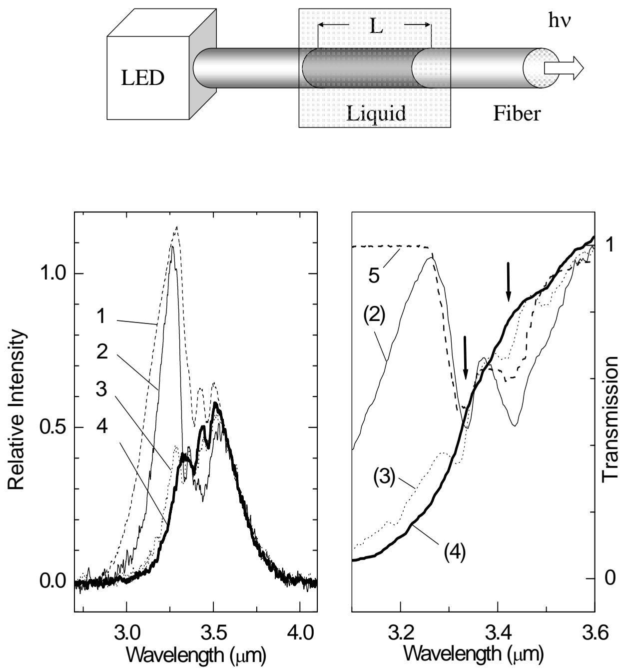


Figure 27. Emission spectra of the 3.4 μm LED coupled with 15 cm long fibre measured at an open end of the fibre (*see top*) in the presence of air (1), alcohol (2), vodka (3) and water (4) (*all at the left*) and the corresponding transmission of the fibre (*right*), (5) – direct measurement of the alcohol vapour transmission in a 10 cm long cell

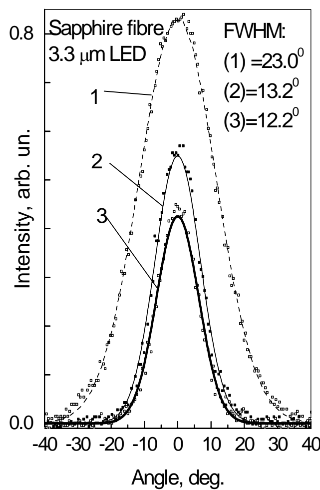


Figure 28. Far field pattern of the 3.4 μm radiation coming from the end of the fiber placed into alcohol (2) and vodka (3), (1) –initial pattern of the radiation (without liquid)

diode surface). The 15 cm long fibre was glued with epoxy into an L=12 cm long liquid cell, filled with water or alcohol, or a 40% mixture of the above, (that is vodka). The corresponding emission spectra recorded from the fibre end are shown in Figure 27 (left) together with the normalized transmission (right), defined as the ratio of the radiation intensity with liquid present in the cell to the intensity for the empty cell. As seen from Figure 27 the fibre LED spectrum is a measure of chemical composition of the liquid since the obtained transmission matches the known features of the liquids: a strong absorption at 3 μm for water and a double peak absorption band for alcohol. The absorption of an evanescent wave results in the change of the far field pattern at the free fibre end, which also contributes to knowledge about the liquid absorption features as shown in Figure 28. The narrowing of the beam coming from the fibre within the liquid is attributed to the increased absorption of the evanescent wave of the rays with multiple reflections (that is, small incident angles) and the increase of the total reflection angle.

It is clear that the use of a chalcogenide glass fibre with a refractive index close to that of the chalcogenide "glue" is beneficial in terms of the coupling efficiency: the latter amounting to 30% [75]. However, moisture and chemicals can erode the fibre and thus in practice the system is a compromise between the stability of the fibre (which is best for low index sapphire) and the coupling efficiency (which is high for chalcogenide fibres).

6 Summary

We have demonstrated that InAs-based narrow gap heterostructures exhibit a potential barrier at the p-n junction up to 300ºC and are able to operate in positive and negative luminescence modes in the 3-5 μm spectral range: the latter being preferable for elevated temperatures in terms of the output power. The optimization of mid-IR diode construction by implementing rare earth gettering, and the use of a broad mirror anode contact and graded bandgap or heavily doped "windows" has lead to "universal" flip-chip devices that are able to operate as efficient LEDs with Fabry-Perot resonant features with an output as high as ~0.5 mW/A and as photodiodes with a detectivity as high as 2×10^{10} cm $Hz^{1/2}W^{-1}$. Optical pumping using a GaAs LED appears to be an efficient way of realizing an InAsSb emitter with a conversion efficiency ~10 μW/A in the 8 μm spectral region. The coupling of the flip-chip devices with immersion lenses or fibres through the use of high index chalcogenide glass together with an appropriate choice of the bias direction at the p-n junction can yield an additional performance enhancement of a factor of 3-5. Optically coupled LED-PD pairs can be used as precise low voltage or current sensors of gases and liquids, *e.g.* with an expected limit of detection for methane gas as small as $LOD_{\Delta f=1\ MHz}$ =18 ppm·cm·mA·$s^{1/2}$.

Acknowledgements

The work was supported by Schlumberger Oilfield Services, by the SBIR/STTR program with administrative support of the U.S. Civil Research and Development Foundation for the IS of the FSU (CRDF), and by the Ministry of Science &

Industry of the Russian Federation. The author is grateful to his colleagues N. V. Zotova, N .D. Il'inskaya, S. A. Karandashev, M. A. Remennyi, N. M. Stus', N.G. Tarakanova, V.V.Shustov and V.K. Malyutenko for generous support, measurements and fruitful discussions.

References

1. P.Vujkovic-Cvijin, D.E.Cooper, J.E.van der Laan and R.E.Warren "Diode Laser Based Lidars: Next Generation" Proc.*SPIE* 1999 ,.**3758,** 142 P.Vujkovic-Cvijin *et al.* "Mobile remote sensor for leak detection in natural gas pipelines" , *Report on the Contract work #5097-260-3941 with the Gas Technology Institute, Des Plaines, IL, USA*
2. Smith, S.D., Crowder, J.G. and Hardaway, H.R., " Recent developments in the application of mid-infrared lasers, LEDs and other solid state sources to gas detection",Proc.*SPIE* 2002, **4651**, 157-172
3. N.M.Kolchanova, A.A.Popov, A.B.Bogoslavskaya, U.S.Sukach, *Tech.Phys.Lett.*, 1993, **19** 64 (in russian)
4. G.A.Sukach, A.B.Bogoslavskaya, P.F.Oleksenko, Yu.Yu.Bilinets, "Effect of Auger recombination on thermal processes in InGaAs and InAsSbP IR-emitting diodes", *Infrared Physics & Technology* 2000, **41,** 299-306
5. M. Aidaraliev, N. V. Zotova, S. A. Karandashev *et al.* "Light Emitting Diodes for the Spectral Range of λ =3.3–4.3 μm Fabricated from the InGaAs- and InAsSbP-Based Solid Solutions: Electroluminescence in the Temperature Range of 20–180°C", *Semiconductors*, 2000, **34,** 102-105
6. L.G.Bubulak, A.M.Andrews, E.R.Gertner and D.T.Longo, "Backside-illuminated InAsSb/GaSb broadband detectors", *Appl.Phys.Lett.*, 1980, **36,** 734-736
7. D.T. Cheung, A.M.Anfrews, E.R.Gertner *et al.* "Backside-illuminated $InAs_{1-x}Sb_x$ - InAs narrow-band photodetectors", *Appl.Phys.Lett.*, 1977, **30,** 587-589
8. P.K.Chiang and S.M.Bedair, "p-n junction formation in InSb and $InAs_{1-x}Sb_x$ by metalorganic chemical vapor deposition", *Appl.Phys.Lett.* 1985, **46,**. 383-385
9. J.K.Abrokwah, M.Gershenson "Liquid phase growth and characterization of $InAs_{1-x}Sb_x$ and $In_{1-x}Ga_xSb$ on (111)B InSb substrates , *Journal of Electronic materials* 1981, **10,** 379-420
10. B.A.Matveev, N.M.Stus', G.N.Talalakin *et al.* "Microhardness of InGaAs, InGaAsSb, InAsSbP semiconductor alloys enriched with InAs" *Izv.Akad.Nauk SSSR, Neorg.Mater,* 1990, **26**, 639
11. B.A. Matveev, N.M. Stus', and G.N. Talalakin, "Inverse defect formation during growth of epitaxial InAsSbP/InAs structures", *Sov.Phys.Crystallogr.*, 1988, **33**, 124-127
12. N.P. Esina, N.V. Zotova, B.A. Matveev *et al.* "Long wavelength uncooled light emitting diodes from $InAs_{1-x-y}Sb_xP_y$ solid solutions", *Sov.Tech.Phys.Lett.*, 1983, **9**, 167-168
13. N.V.Zotova, A.V.Losev, B.A.Matveev *et al.* , "Absorption edge of variable-gap $InAs_{1-x}Sb_x$ ($x < 0.54$) epitaxial layers", *Sov.Tech.Phy.Lett.*, 1990, **16,** 155 - 157
14. B.A. Matveev, N.V. Zotova, S.A. Karandashev, M.A. Remenniy, N.M. Stus', G.N. Talalakin, "III-V optically pumped mid-IR LEDs". Proc. *SPIE*, 2001, **4278**, 189-196
15. B.V. Morozov, Yu.B. Bolkhovityanov, R.S. Gabaraev *et al.* , *Sov.Phys.Semicond*, 1980, **14**, 883
16. M.Fisher, A.Krier, "Photoluminescence of epitaxial InAs produced by different growth methods", *Infrared Physics & Technolog,* 1997, **38**, 405-413

17. G.R.Nash, N.T. Gordon, M.T.Emeny, T.Ashley, “Perspectives on dynamic infrared scene projection using positive and negative luminescence” Proc. *SPIE* , 2003**, 5092,** 138-144
18. B. A. Matveev, N. V. Zotova, N .D. Il’inskaya *et al.* "Towards efficient mid-IR LED operation: optical pumping, extraction or injection of carriers?”, *J.Mod.Optics*, 2002, **49,**. 743-756
19. Matveev, B.A.; Zotova N.V., Karandashev S.A. *et al.* “Backside illuminated In(Ga)As/InAsSbP DH photodiodes for methane sensing at 3.3 μm”, Proc. *SPIE* 2002, **4650,** 173-178
20. Z.M.Fang, K.Y.Ma, D.H.Jaw *et al.* “Photoluminescence of InSb, InAs, and InAsSb grown by organometallic vapor phase epitaxy”, *J.Appl.Phys*.1990, **67,** 7034-7039
21. X.Y. Gong, T. Yamaguchi, H. Kan *et al.* “Mid-Infrared Photoluminescence from Liquid Phase Epitaxial InAsSb/InAs Multilyers”, *Jpn.J.Appl.Phys*, 1997, **36**, 738-742
22. Esina N.P., Zotova N.V. “Mechanisms of recombination of excess carriers in InAs and related solid solusions” *Sov.Phys.Sem*., 1980, **14** (*Fizika & Technika Poluprovodnikov*, v.14, No 3, pp. 316-322, in Russian).
23. A.Krier, “Physics and technology of mid-infrared light emitting diodes”, *Phil. Trans. R. Soc. Lond*. A 2001, **359** , 599-619
24. N.V..Zotova, S.A. Karandashov, B.A. Matveev *et al.* “Gadolinium -doped InGaAsSb solid solusions on an InAs substrate for light-emitting diodes operating in the spectral interval λ=3-5 μm”, *Semiconductors*, 1999, .**33** , 920-923
25. A.Rogalski “Heterostructure infrared photovoltaic detectors” *Infrared Physics and Technology* 2000, **41,** 213-238
26. A.A. Bergh & P.J. Dean, “Light-emitting diodes”, Claredon Press, Oxford, 1976, (Russian translation: ,”Mir”, Moscow 1979).
27. Http://www.ioffe.rssi.ru/SVA/NSM/Nano/index.html.
28. J.G. Crowder, T. Ashley, C.T. Eliott *et al.* “Minimally cooled InSb/InAlSb LED and photodiode devices applied to nitrogen dioxide detection at ppm levels”, *Electronics Lett*., 2000, **36**, 1867-1869
29. M.J. Kane, G. Braithwaite, M.T. Ereny *et al.* “Bulk and surface recombination in InAs/AlAs $_{0.16}$ Sb $_{0.84}$ 3.45 μm light emitting diodes”, *Appl.Phys.Lett*., 2000, **76**, 943-945
30. A.Krier, D.Chubb, S.E.Krier *et al.* “Light sources for wavelengths >2 μm grown by MBE on InP using a strain relaxed buffer , *IEE Proceedings, Optoelectronics* 1998, **145,** 292-296
31. M Aidaraliev, N V Zotova, N D Il’inskaya *et al.* “InAs and InAsSb LEDs with built-in cavities” *Semicond. Sci. Technol.* 2003, **18,** 269–272
32. Zh.I. Alferov, A.T. Gorelenok, V.G. Gruzdov *et al.* “InGaAsP/InP DH LEDs (λ=1.55 μm) with external efficiency ηe≈30%(300 K)” *Phys.Tech.Lett* 1982, **8**, (*Pis’ma v zhurnal technicheskoy fiziki*, 1992**,** **8**, 257-262, in Russian)
33. V. K. Malyutenko, O.Yu.Malyutenko, A. D. Podoltsev *et al.*, “Current crowding in InAsSb LED structures”, *Applied Physics Letters*, 2001, **79**, 4228-4230
34. D.A.Wright, V.V.Sherstnev, A.Krier *et al.* “Mid-infrared whispering gallery mode ring lasers and LEDs” *IEE Proc.-Optoelectronics* 2003*,* **150**, 314-317
35. M.A. Remennyi, B.A. Matveev, N.V. Zotova *et al.* "InGaAsSb negative luminescent devices with built-in cavities emitting at 3.9 μm ", *Physica E: Low-dimensional Systems and Nanostructures*, 2004, **20**, 548-552
36. D. Gevaux, A.Green, C.Palmer *et al.*, “Resonant-cavity light emitting diodes (RC-LEDs) and detectors for mid-IR gas-sensing applications”, *IEE Proc. Optoelectron*.,2003, **150,** 360-364
37. N.V. Zotova, N.D. Il’inskaya, S.A. Karandashev *et al.* “InAs light-emitting diodes with cavity formed by anode contact and semiconductor/air interface” *Semiconductors* 2004, **38**, 1–4

38. B. A. Matveev, M. Aidaraliev, N. V. Zotova *et al.*, In(Ga)As- and InAs(Sb)-Based Heterostructure LEDs and Detectors for the 3¸-5 μm Spectral Range", *Book of MIOMD-V* Abstracts 2002, 97-98
39. B. Matveev., N. Zotova, N. Il'inskaya *et al.*, "Spontaneous and stimulated emission in InAs LEDs with cavity formed by gold anode and semiconductor/Air interface", *phys. stat. sol. (c)* 2005, **2**, 927–930
40. M. AЇdaraliev, N. V. Zotova, S. A. Karandashev, *et al.* ,"Light Emitting Diodes for the Spectral Range λ = 3.3-4.3 μm Fabricated from InGaAs and InAsSbP Solid Solutions: Electroluminescence in the Temperature Range 20–180°C (Part 2)" *Semiconductors*, 2001, **35**, 598–604.
41. V. I. Ivanov-Omskii, B. T. Kolomiets, and V. A. Smirnov, "Radiative recombination in InSb at magneto-concentration effect", 1965 161 *Dokl. Ak. Nauk SSSR* 1956, **161**, 1308 -1309 [*Sov. Phys. Dokl*. 10, 345 (1965)].
42. J. R. Lindle, W. W. Bewley, I. Vurgaftman *et al.* "Negative luminescence from mid-wave infrared HgCdTe diode arrays" *Physica E: Low-dimensional Systems and Nanostructures* 2004, **20**, 558-562
43. B. A. Matveev, N. V. Zotova, S. A. Karandashev *et al.*, "Towards longwave (5÷ 6 μm) LED operation at 80°C : injection or extraction of carriers?", *IEE Proceedings - Optoelectronics* 2002 , **149** , 33 - 35
44. G.R. Nash, N.T. Gordon, D.J. Hall *et al.* , "Infrared negative luminescent devices and higher operating temperature detectors", *Physica E: Low-dimensional Systems and Nanostructures* 2004, **20**, 540-547
45. B.A. Matveev, M. A'daraliev, N.V. Zotova, *et al.* "Negative luminescence from InAsSbP-based diodes in the 4.0-4.3 μm range", Proc *SPIE* 2001, **4285**, 109-117
46. B.I.Stepanov ,"Basics for spectroscopy utilizing negative light beams", Minsk, 1961
47. A.M.White, "Generation-recombination processes and Auger suppression in small-bandgap detectors", *Journal of Crystal Growth* 1988, **86** , 840-848
48. L.J. Olafsen, I. Vurgaftman, W.W. Bewley *et al.* "Negative luminescence from Type-II InAs/GaSb superlattice photodiodes", *Appl.Phys.Lett*. 1999, **74**, 2681-2683
49. W. W. Bewley, M. J. Jurkovic, C. L. Felix *et al.* ,"HgCdTe Photodetectors with Negative Luminescent Efficiencies > 80%", *Appl.Phys.Lett.*, 2001, **78**, 3082-3084.
50. M A Remennyi, N V Zotova, S A Karandashev *et al.* "Low voltage episide down bonded mid-IR diode optopairs for gas sensing in the 3.3-4.3 μm spectral range" *Sensors & Actuators B: Chemical,* 2003, **91**, 256-261
51. A.M.White, "Generation-recombination processes and Auger suppression in small-bandgap detectors", *Journal of Crystal Growth* 1988, **86**, 840-848
52. N.P. Esina, N.V. Zotova, D.N. Nasledov, "Electroluninescence in InAs p-n junctions*", Fiz. Tech. Polupr.*, 1969, **3**, 1370-1373(in Russian).
53. A. Krier and Y. Mao, "High performance InAsSbP/InGaAs photodiodes for the 1.8-3.4 μm wavelength range", *Infrared Physics & Technology,* 1997*,* **38***,* 397-403
54. J.L. Malin, C.L. Felix, J.R. Meyer *et al.*, "Type II mid-IR lasers operating above room temperature", *Electron. Lett*., 1996, **32**, 1593-1594
55. M. Boroditsky, T.F. Krauss, R. Cocciol *et al.* , "Light extraction from optically pumped light-emitting diode by thin-slab photonic crystals", *Appl.Phys.Lett.,* 1999, **75**, 1036-1038.
56. Adrian John Maldan, Patent of GB No 2102942, G01N26/31, 21/35, published 09.02.83.
57. B.A. Matveev, N.V. Zotova, N.D. Il'inskaya *et al.* , "Radiation Source", GB patent #2363906 filed 21 April, 2000.
58. M. Aidaraliev, N. V. Zotova, S. A. Karandashev *et al.*, "Optically Pumped "Immersion-Lens" Infrared Light Emitting Diodes Based on Narrow-Gap III–V Semiconductors", *Semiconductors*, 2002, **36**, 828-831

59. N.V. Zotova, S.A. Karandashev, B.A. Matveev, M.A. Remennyi ,N.M. Stus', and N.G. Tarakanova. "Luminescence of Multilayer Structures Based on InAsSb at λ= 6–9 μm". *Semiconductors* 2005, **39**, 214–217
60. V.N. Brudnyi, N.G. Kolin_, A.I. Potapov V.D. Kuznetsov "Electrophysical properties of proton-irradiated InAs" *Semiconductors* 2003, **37** , 390-395
61. C.Sirtori, F.Capasso, J.Faist *et al.*, "Quantum cascade unipolar intersubband light emitting diodes in the 8-13 μm wavelength region" *Appl.Phys.Lett.*1998, **74**, 2384-2386
62. A.Green, D.Gevaux, C.Roberts, and C.Philips, "Resonant-cavity-enhanced photodetectors and LEDs in the mid-infrared", *Physica E: Low-dimensional Systems and Nanostructures*, 2004, **20**, 531-535
63. R.Windish *et al.* , "40% efficient thin-film surface textured light-emitting diodes by optimization of natural lithography" *IEEE. T.Electron.Dev*. 2000, **47**, 1492-1498
64. Ashley T., Dutton D.T., Elliott C.T. *et al.* , "Optical Concentrators for Light Emitting Diodes", Proc. *SPIE* 1998, **3289**, 43
65. R.C. Johnes. "Immersed radiation detectors", *Appl.Opt.,* 1962, **1**, 607-613
66. A.G.Fischer and C.J.Nuese, "Highly Refractive Glasses to Improve Electroluminescent Diode Efficiencies". *J.Electrochem.Soc. SOLID STATE SCIENCE* 1969, **116**, 1718-1722
67. Jaw W. Chey, Peter Sultan, Hendrik J.Gerritsen, "Resonant photoacoustic detection of methane in nitrogen using a room temperature infrared light emitting diode" *Appl.Optics*, 1987, **26**, 3192-3194
68. Esina N.P, Zotova N.V., Markov I.I. *et al.* "Gas analyzer based on semiconductor components", *J.Appl.Spectrosc*., 1985**, 42,** 465-467
69. V.K. Malyutenko, 0.Yu. Malyutenko, A. Dazzi *et al.* ," Heat transfer mapping in 3-5 μm planar light emitting structures", *J. Appl. Phys*. 2003, **93**, 9398-9400.
70. B.A. Matveev, M. Aydaraliev, N.V. Zotova *et al.*, «Flip-chip bonded InAsSbP and InGaAs LEDs and detectors for the 3 μm Spectral Region" *IEE Proc.-Optoelectronics* 2003**,** **150**, 356-359
71. S.McCabe and B.D.MacCraith, "Novel mid-infrared LED as a source for optical fibre gas sensing", *Electron. Lett*., 1993, **29,** 1719-1721
72. Messica, A., Greenstein, A., and Katzir, A., "Theory of Fiber-Optic, Evanescent-Wave Spectroscopy and Sensors," *Appl.Opt.,* 1996, **35**, 2274-2284
73. B. Mizaikoff, "Mid-Infrared Fiberoptic Evanescent Wave Sensors – A Novel Approach for Subsea Monitoring", *Meas. Sci. Technol*., 1999, **10**, 1185–1194
74. Rick K. Nubling and James A. Harrington , "Optical properties of single-crystal sapphire fibers", *Appl.Opt.,* 1997, **36**, 5934-5940
75. B. A. Matveev, N.V.Zotova, S. A. Karandashev *et al.* "3.4 μm "Flip-chip" LEDs for Fiber Optic Liquid Sensing" *Proceedings of the 1-st International Conference on Advanced Optoelectronics and Lasers (CAOL'2003), 2003 Alushta, Crimea, Ukraine* v.**2**, 138-140

QWIP Detectors for the MWIR

S. Haywood and K.T. Lai
Department of Engineering, University of Hull,
Hull, HU6 7RX, U.K.

M. Missous
School of Electrical & Electronic Engineering, University of Manchester,
Manchester, M60 1QD, U.K.

1 Introduction

Despite extensive and sustained research efforts in the last 30 years, the Mid Wavelength Infra Red (MWIR) region of the electromagnetic spectrum (defined here as 2 to 5μm) still does not enjoy the performance, cost benefits and ease of use of its visible and near-IR (NIR) counterpart (~ 0.2 to 1.6μm) [1]. The availability of high performance and low cost sources and detectors in the visible and NIR region of the spectrum has much to do with impressive development in GaAs technology for sources and in silicon CCD technologies for detectors leading to ideal source-detector combinations. This in turn has led to an explosive growth in spectroscopic measurements, which have diverse applications in scientific fields from biological and environmental systems to communications and security [2]. This success is to some extent being duplicated now in the 1.5 to ~ 2.5μm region using InP-based technology for the sources and detectors and indeed cost-effective room temperature, fast InGaAs photodiode arrays are slowly but steadily dominating this region. The fundamental materials property that is common to the (commercial) devices used in the applications above is that both emission and absorption occur *across the band gap* of the material, which then determines the operating wavelength. However as the working wavelength lengthens (> 2μm), the band gap energy required to generate these wavelengths reduces (< 0.5eV), leading to increased thermal carrier generation. This effect, which increases exponentially with reducing band gap, severely undermines the performance of mid–IR devices. Smaller band gaps, below 0.5eV, also lead to increased Auger recombination adding to the overall difficulties of making efficient devices in the infrared region. Notwithstanding these problems, the fundamental processes at play here lead to sub-nanosecond recombination rates and thus to fast devices capable of being modulated in both emission and absorption in the GHz range.

1.1 QWIPs *vs* Interband Detectors

Traditional materials used in the MWIR and long wavelength regions (LWIR) are mainly mercury cadmium telluride (MCT) and InSb. (Excellent reviews comparing the current state of QWIP and interband technologies have been published recently by Rogalksi [2,3].) MCT is soft, difficult to synthesise by any growth method and suffers from non-uniformity, making the fabrication of focal plane arrays a

challenging and costly proposition. The band gap of MCT varies as a function of its composition from HgTe, which is a semimetal with a negative band gap, to CdTe with a band gap of 1.6 eV. At a mercury composition of 80%, the band gap attains a value of 0.1 eV, sufficient for detection of IR radiation up to 12μm. In the LWIR, the high concentration of mercury is the reason for the great difficulty in achieving uniform focal plane arrays. As a result, GaAs based Quantum Well Infrared Photodetectors (QWIPs) are now dominating the field in the 8-12 μm and have achieved considerable commercial success [4]. The key reason for the great interest lies largely in the fact that these detectors are based on GaAs materials, which have mature material growth and device processing technologies. Much effort has been devoted to optimize the performance of devices operating in the 8-12μm regime, most of which are based on photoconductive multi-quantum well structures, because of their high responsivity. However in the MWIR region, where the concentration of mercury needed to synthesise devices is low, MCT has been the workhorse and 2D array technology is available based on MCT as well as InSb compounds. InSb does not suffer from the material uniformity problems which trouble MCT. However, it also requires cooling for mid–IR operation, in part because the band gap decreases with increasing temperature which exacerbates the enhanced thermal generation always associated with temperature elevation and takes the device operating cut-off wavelength above 5μm. Introduction of InAlSb (10% Al) blocking layers has successfully raised the temperature at which images can be obtained from InSb arrays (256 × 256 pixels) to 130K. Use of all InAlSb (3.5% aluminium) devices can raise this operating temperature still further to 150K while keeping the operating wavelength below 5μm [5].

One of the drawbacks of QWIP detectors is the lack of normal incidence absorption due to the polarization selection rules [6]. This necessitates additional processing in the form of gratings or polishing at 45° to achieve coupling of the light. The internal quantum efficiency is also typically ≤ 20%, compared to >70% for interband detectors due to the smaller number of carriers available to take part in transitions. Increasing the doping can enhance quantum efficiency [7] but this necessarily raises the Fermi energy and thus quantum efficiency is increased at the expense of a relatively large dark current and hence increased noise for a given operating temperature. For a detector responding in the range 3-5μm (8-9.5μm), cooling to temperatures down to 100-120K (70-77K) is typically necessary in order to reduce dark currents to sufficiently low levels. The dominant mechanism of dark current generation in QWIPs is longitudinal optical phonon excitation of carriers because of the very strong coupling in III-V materials. The most obvious way of minimizing the dark current is thus to decrease the detector temperature. It is also evident that cooling requirements become more severe for longer cut-off wavelengths. The temperature at which background limited performance (BLIP) is achieved is a useful figure of merit to compare detector operating temperature performance. A BLIP detector is defined as one in which the dark current is less than or equal to the background (300K) photocurrent (180° field of view, with a cooled sample).

It is only in the last few years that sufficient advances in epitaxial growth, synthesis and design of strained GaAs and InP quantum well structures have started to challenge the pre-eminence of MCT and InSb in the MWIR. As well as

uniformity of growth and processing, essential for focal plane arrays, the advantages that these materials potentially offer includes *high speed* switching for free space communications applications (<10ps [8]). The combination of relatively short wavelengths and the high barriers used in double barrier quantum well (DBQW) structures (see Sections 2.2. and 2.3) also enables high temperature operation for moderately doped QWIP structures [9, 10]. This chapter will thus concentrate on material issues which affect the selection rules and transition wavelength (Section2.1) as well as detector parameters such as detectivity, operating temperature (Section 2.3), operating voltage (Section 2.5) and response time (Section 2.4). We shall consider QWIPs based on GaAs and InP to access the MWIR region where Stirling cooled MCT and InSb are still used extensively. Other technologies such as antimonide-based compounds are also being intensively researched [11,12] and are showing real potential for the MWIR; however this technology is still less mature than that of either GaAs or InP and will not be considered here.

Presently there are no detectors (interband or intersubband) in the MWIR that work at room temperature with a sufficiently high detectivity ($D^* \geq 10^{10}$ cm.$Hz^{1/2}$ W^{-1}). Hence, the 'holy grail' of MWIR technology is the development of efficient room temperature detectors and we devote Section 2.3 to a study of strain-balanced QWIPS which offer the potential for near room temperature operation.

2 MWIR Transitions in QWs

2.1 Square QW Structures

Quantum wells used for QWIPs are typically designed to have two bound states in the well with their energy separation corresponding to the minimum energy of the photons to be detected. The energy of the excited state is ideally a little below the barrier edge so that excited electrons can easily escape from the QW and give rise to a photocurrent when a modest electric field is applied. This is the so-called bound-to-bound (B-B) transition (as shown in Figure 1). Alternatively, transitions from a single bound level in the well to the continuum can be used producing bound to continuum transitions (B-C in Figure 2). A detailed analysis of the respective merits and features of these types of transition can be found in the review by Levine [13]. In summary, however, B-B transitions show a high oscillator strength and narrow transition width compared to B-C transitions but require a threshold bias to be applied before any photocurrent is collected. Hence responsivity increases more slowly with bias than for a B-C QWIP which shows photoconduction at all finite values of bias. However B-C transitions are broader with lower oscillator strength than is found for B-B devices.

Figure 1. Conduction band profile of a bound-to-bound transition: (**a**) a photon having energy ħω excites an electron in the ground state to the excited state and (**b**) the photoexcited electron tunnels out of the well and thus contributes to the photocurrent

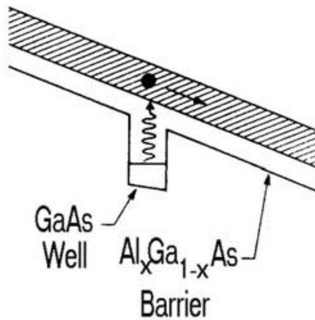


Figure 2. Conduction band structure for a bound-to-continuum (B-C) QW showing the photoexcitation and hot electron transport processes

2.1.1 (In)GaAs/AlGaAs on GaAs

At the short wavelength end of the MWIR region the separation between the bound states must be greater than 0.62eV. This then dictates large conduction band discontinuities (for n-type QWIPs). The lattice-matched system of GaAs-$Al_{0.4}Ga_{0.6}As$ (on GaAs) has a conduction band offset of ~350meV thus limiting the wavelength that can readily be obtained to greater than 5.6μm [14]. For structures in which the $Al_xGa_{(1-x)}As$ has been allowed to go indirect (x>0.45), transitions as short as 2.9μm have been reported, albeit for a normally forbidden Γ1-Γ3 transition [15]. Wavelengths as short as 1.26μm have been reached using B-C transitions in $In_{0.3}Ga_{0.7}As$/AlAs wells. However the suitability of this device for a QWIP is questionable since the X-band minimum in the AlAs is below Γ1 in the well making it difficult to populate Γ1 with electrons simply by doping. There is in general an increase in dark current for Al ratios >0.45, due to enhanced carrier escape via the AlGaAs X minimum. Therefore the preferred route to produce device structures in this system is the addition of In to the QW giving InGaAs/AlGaAs wells and retaining the direct band structure throughout. Early attempts used In compositions

below 20% to minimize the compressive strain in the QWs and the net strain throughout the structure [16,17]. Using 18% indium in the QWs, a detectivity of $D^*=4\times10^{10}cm^2\ Hz^{1/2}/W$ and a responsivity of R=12mA/W were measured for a 4.5 μm device at 95K. By using a linearly graded InGaAs buffer between the GaAs substrate and the detector active region it was shown to be possible to use higher values of indium. Chui *et al.*[18,19] grew a series of samples of $In_yGa_{(1-y)}As/Al_xGa_{(1-x)}As$ with different (y, x) values including (0.3, 0.45), (0.5, 0.45), (0.6, 0.45), (0.6, 0.67), (0.5 ,1) and (0.6, 1). These corresponded to absorption wavelengths of 3.9, 3.5, 3.0, 2.6, 2.4 and 2.2 μm respectively. Thus a peak wavelength of 3μm was achieved without using indirect barriers and 2.2 μm with indirect barriers. More recent examples of work on InGaAs/AlGaAs QWIPs include (y, x) = 0.35, 0.38 with λ_{peak} = 4.3μm [20] and a detailed theoretical analysis of this system by Choi *et al.*[21]. MWIR InGaAs/AlGaAs focal plane arrays have also been demonstrated with λ_{peak} = 5.1μm [22].

2.1.2 InGaAs/AlInAs on InP

$In_{.53}Ga_{.47}As$-$In_{.52}Al_{.48}As$ lattice-matched to InP has a band offset of 500meV allowing transitions down to ~4μm with good device characteristics. D* of 2.3×10^{10} $cm.Hz^{1/2}\ W^{-1}$ has been reported for a 4μm B-C QWIP operating at 120K [23]. The related system of InGaAs/InP quantum wells on InP has a band offset ratio in favour of the valence band making it less suitable for MWIR n-type QWIPs. InGaP compounds have been used to increase this barrier height and also allow deeper wells of high indium content ($In_{0.8}Ga_{0.2}As$) [24]. This strain-balanced structure also gives the possibility of a large critical thickness for the device. Nevertheless B-C transition energies were only just below 5μm for a 4.2nm well although BLIP performance persisted up to 135K with $D^* = 3.2\times10^{10}cm^2\ Hz^{1/2}/W$ up to 110K.

InGaAs/InP has been used to investigate p-type QWIPs and a peak response at 4.55μm was obtained using extremely narrow (1nm) wells [25]. However, the development of p-type QWIPs has been impeded by their often inferior transport and confinement properties compared to n-type QWIPs. Furthermore, the smaller electron effective mass and the higher mobility typical of the conduction band are also highly desirable for the detector's sensitivity and speed. In addition, due to the different curvatures of dispersion relations for the heavy and light hole states (different effective masses), the absorption spectrum based on the intersubband transition between the heavy and light hole states is broad [26]. Hence the absorption strength [27] and detectivity [28] are significantly smaller than for intersubband transitions in the conduction band.

2.2 Double-barrier QWs

To attain the shorter MWIR wavelengths it is beneficial to enhance the confinement of carriers in the QWs described in the previous section. Enhancement of the photocurrent collection at low bias from B-B QWIPs is also advantageous, since this indirectly reduces the dark current under device operating conditions and enhances detectivity. A specially designed structure called a double barrier QW

(DBQW) has been proposed by Schneider *et al.*[29] which simultaneously achieves both of these goals. It consists of a well layer surrounded by two high inner barriers followed by two lower outer barriers. The traditional bound-to-bound transition in simple QWIPs thus becomes a bound-to-quasi-bound transition (B-QB) as shown in Figure 3. Transition energies can be tuned more readily to reach the 2-4μm region and photoexcited carriers can tunnel through the thin high (indirect) inner barrier. Providing this barrier is thin the transit time for carriers is rapid compared to the Γ-X scattering time [30] hence few carriers are lost in this way.

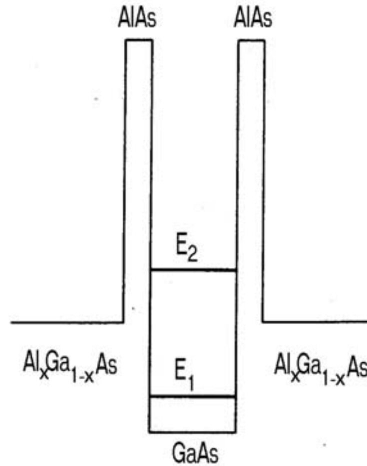


Figure 3. Schematic conduction band diagram of a bound-to-quasibound (B-QB) QW using GaAs QWs and AlAs/AlGaAs barriers

2.2.1 DBQWs on GaAs Substrates

DBQWs have been grown on GaAs substrates using both GaAs QWs [29-37] and InGaAs QWs [38-44]. Using GaAs QWs, transition energies for high performance QWIPS can be reduced to about 4μm. For example, Tsai *et al.*[36] report $D^* = 1\times 10^{11}cm^2\ Hz^{1/2}/W$ at 4.3 μm and 77K using 1.2 nm AlAs barriers and a 5nm GaAs well. Similarly Wu *et al.* report $D^* = 5\times10^{11}cm^2\ Hz^{1/2}/W$ at 80K for a 5.3μm QWIP with 3nm wells and 1.5nm barriers [37].

The introduction of strained InGaAs wells between the high AlAs barriers can significantly decrease the wavelength which can be attained and/or enhance the operating temperature Fiore *et al.* report $D^* = 4\times10^{10}cm^2\ Hz^{1/2}/W$ at 4.5μm with operation up to 95K and were able to grow a 100 period sample with 16% indium in the 2.5nm InGaAs QWs [17]. Liu *et al.* report transitions as short as 1.8μm using 2.5nm InGaAs wells with 10% indium [40]. However these transitions were weak due to loss of electrons from the high Γ1 level in the QW to the almost resonant X minimum in the AlAs barriers. In the purely Γ electron regime the shortest detection wavelength was limited to 3μm. The strain incorporation associated with indium levels greater than about 15% presents problems for the fabrication of MQWs with a large number of periods using InGaAs/AlAs and hence limits the use of this material system to produce high performance devices.

2.2.2 DBQWs on InP Substrates

Lattice-matched $In_xGa_{1-x}As/In_yAl_{1-y}As$ quantum wells on InP (x=0.52, y=0.53) produce a conduction band discontinuity of 0.51eV. This is less than the 0.62eV transition required for the lower limit of 2μm in the MWIR. The use of increased indium content (compressively strained material) in the quantum well gives a net reduction in band gap and hence increases the band discontinuity. However, this is still not sufficient for short wavelength MWIR emission. Furthermore the use of high compressive strain in a periodic structure will eventually lead to problems with dislocations. Strain compensation of the QW with tensile barrier material (by reducing the indium content) permits higher compressive, narrower band gap QWs (higher indium content) to be used. Concomitantly, there is a large increase in the band discontinuity due to the increased band gap of the barrier material (Figure 4).

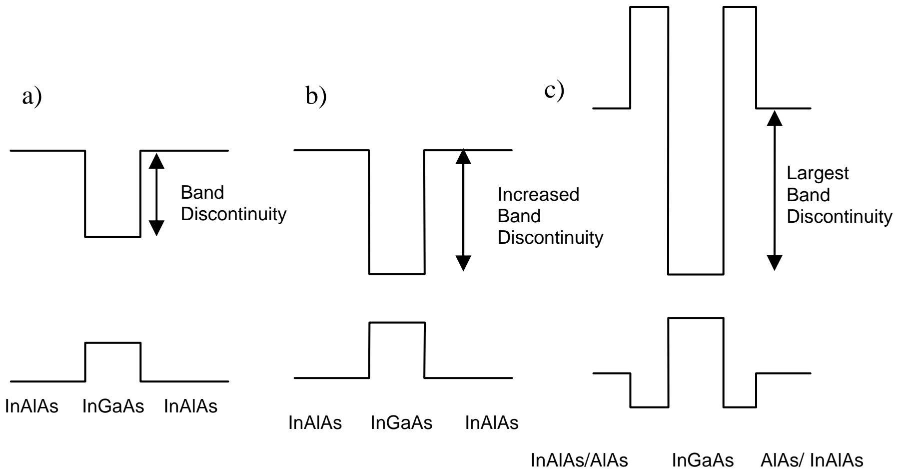


Figure 4. Schematic of: **a**) Lattice-Matched **b**) Strained and **c**) Strain-compensated QWs

Following the above picture, lattice-matched InGaAs QWs have recently been investigated with thin AlAs barriers under tensile strain with respect to the InP substrate and wide barriers made from lattice-matched AlInAs [9, 45]. This method has allowed transitions (measured in absorption) to be obtained down to 1.55μm but this can only be achieved using very thin (3ML) QWs [45]. The deeper well in InGaAs/AlInAs compared to GaAs/AlGaAs combined with the high AlAs barrier has however led to some of the highest operating temperature QWIPs yet to be produced. Lee *et al.* reported 205K operation at 3.4 μm with $D^* = 5\times10^{10}cm^2\ Hz^{1/2}/W$ at 77K [9]. The band diagram for this device which had a 4.4nm InGaAs QW is shown in Figure 5. E_F is about 273 meV below the conduction band edge of the InAlAs wide barrier. (It is this barrier height relative to E_F which primarily determines the thermal escape processes of electrons out of the well due to the indirect nature of the AlAs.)

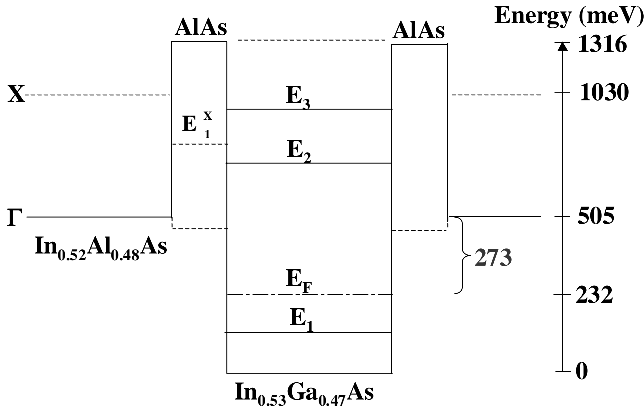


Figure 5. Band diagram for InGaAs QW lattice-matched to the InP substrate with tensile strained thin AlAs barriers and lattice-matched InAlAs wide barriers (after Lee *et al.*, [9]). The *dotted line* indicates the position of the X valley and the energy of the first confined state in the X-well (which is in the AlAs)

Intersubband transitions below this wavelength or with increased operating temperature calls for increased band offset in the conduction band. This can be achieved through the use of highly strained (but strain-balanced) $In_xGa_{(1-x)}As$-$In_yAl_{(1-y)}As$ ($x>0.8$, $y<0.3$) structures [10, 46-49]. A detailed study of work in this system is presented in the following section.

3 Strain-balanced QWIPs for High-temperature Operation

When grown on InP, it is possible to tailor the In and Al compositions in $In_xGa_{(1-x)}As$-$In_yAl_{(1-y)}As$ such that compressive and tensile strain are generated in the well and barrier regions respectively resulting in strain compensation. QWIPs rely on the growth of ultra thin layers in the build up of their active regions, where the thickness of individual layers rarely exceeds 8nm and most commonly ranges from 2 to 5 nm. This feature makes the growth of strained layers a possibility since the critical thickness of even a highly mismatched pair in the $In_xGa_{(1-x)}As$-$In_yAl_{(1-y)}As$ system ($x\sim0.8$, $y\sim0$) would still be safely above the 8nm upper limit. The $In_xGa_{(1-x)}As$-$In_yAl_{(1-y)}As$ also allows access to conduction band discontinuities >0.7eV, exactly what is needed for the short wavelength region of the MWIR.

3.1 Materials Growth and Characterisation

Molecular Beam Epitaxy is the technique of choice for the growth of QWIP structures because of the precision required in the thickness of quantum wells. The

growth of In-containing compounds dictates the use of fairly low growth temperatures (from 350 to 450°C, depending upon the indium concentration in InGaAs) because of the high indium adatom mobility which leads to clustering at high substrate growth temperatures. However as the growth temperature approaches 350°C or less, arsenic atoms start to form both anti-site and interstitial defects, compromising lattice integrity and correct compound formation. Furthermore, these temperatures are incompatible with growth of the Al-containing compounds used in the barrier where high temperatures are usually required (> 500°C).

When aluminium-containing compounds are grown at lower temperatures the (usual) excess arsenic overpressure on the surface reduces the surface diffusion of aluminium adatoms, which is already low compared to gallium and this leads to reduced crystal quality through the incorporation of point defects. It has been shown [50] that at low temperature sufficient energy can be gained from a combination of the ballistic energy and the substrate temperature to induce some surface diffusion of aluminium if the impinging arsenic flux is kept low. If the aluminium compound is to form a ternary compound with indium, besides the problem of clustering [51], a further consideration needs to be taken into account. This is that the indium sticking coefficient decreases with temperature (above 525°C). A common solution therefore is to perform growth at temperatures of 500-525°C but with a very high overpressure of arsenic, which reduces the formation of clusters.

Leaving aside the problem of introducing point defects, there are some advantages of low temperature growth, especially when growing highly strained material. At lower growth temperatures, larger critical thicknesses can be grown due to the reduced heat energy in the structure (because the wafer is kept at a lower growth temperature, the heat input into the structure is reduced compared with growth at higher temperatures) helping to conserve lattice integrity. This is of fundamental importance for the QWIPs described here. Another important aspect of low growth temperatures concerns interface abruptness and the reduction of dopant diffusion. Summarising all these effects, it is clear that there are many contradictory requirements. The way around this dilemma and to achieve high quality compounds containing aluminium and indium is to grow under *stoichiometric conditions* [50, 51].

3.2 Stoichiometric Growth Conditions

At low growth temperatures a stoichiometric mix of fluxes impinging upon the wafer surface forces the compound to form in the desired ratios, *i.e.* unwanted arsenic defects are reduced. This allows greater diffusion and migration of low mobility particles such as aluminium which, combined with the effect of ballistic energy (or heat of condensation), gives aluminium a greater mobility than expected at these lower temperatures. Problems of slow surface diffusion rates can also be addressed by slowing down growth rates by controlling the fluxes. As the surface mobility of group III adatoms is so important to flat morphology and crystal quality, this technique offers a unique opportunity to grow new material systems. This is because the adatoms are able to diffuse to steps on the surface where they incorporate preferentially allowing the surface to grow nearly layer-by-layer. In this manner many of the problems of low-temperature growth can be negated, allowing

an increase in critical thickness of strained layers and an increased overall strain in a device without degradation of the crystal lattice integrity. This technique, demonstrated some years ago [50] is called Stoichiometric Low Temperature (SLT) growth and was used throughout the growth of all the samples described in this section. The range of growth temperatures used ranged from 350 to 450°C.

The low growth temperatures thus used have the added advantage of leading to abrupt interfaces, effective dopant confinement within the quantum well and an increase in the critical thickness of the strained materials. This technique has been extensively used in the synthesis of the InGaAs-AlAs structures designed for MWIR operation [46-48] and was largely responsible for the excellent results obtained on the devices made from these materials.

3.3 Structural and Electrical Properties

For the QWIP structured described below, a large ΔE_C is achieved by increasing the indium composition in the InGaAs well to 84% and balancing the high compressive strain with thin tensile-strained AlAs inner barriers [10, 46-48] followed by a lattice matched $In_{0.52}Al_{0.48}As$ acting both as a strain "relief" layer for the next period and also as the barrier for tunnelling from the QW. With this rationale the structures shown in Figure 6 were grown. The ΔE_C for these structures is estimated to be ~675 meV with respect to the wide lattice matched InAlAs barriers and 1.475 eV for the AlAs tensile layers. To span the entire MWIR region, three structures were studied, which differed only in their well width and therefore the net strain. The well width varied from 3 nm (sample 1561) to 4.5 nm (sample 1563) to 8 nm (sample 1554). The effective strain varied from highly tensile to nearly strain-compensated.

	$In_{0.53}Ga_{0.47}As$	Cap	10nm
W	AlAs	Barrier	2.3nm
W	$In_{0.84}Ga_{0.16}As$ (Si)	QW	*d* nm
W	AlAs	Barrier	2.3nm
W	$In_{0.52}Al_{0.48}As$	Barrier	22nm
	$In_{0.52}Al_{0.48}As$	Buffer	50nm
	InP:Fe	Substrate	

Figure 6. Generic epitaxial layer profile of strained AlAs-$In_{0.84}Ga_{0.16}As$ short wavelength QWIP, the periodicity W was varied from 15 to 25. The doping in the well is 8×10^{18} cm^{-3}

Despite the very high lattice mismatch of the two constituent semiconductors, structural integrity is clearly maintained as shown in Figure 7 which depicts a Double Crystal X-Ray Diffraction (DCXRD) (004) spectrum for sample 1554.

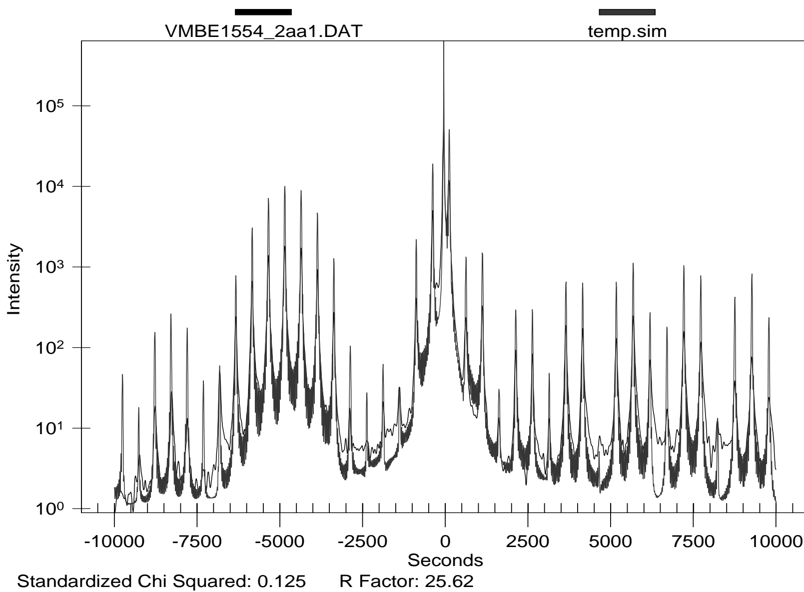


Figure 7. DCXRD (004) reflection of a strain compensated AlAs-$In_{0.84}Ga_{0.16}As$ short wavelength QWIP and the corresponding RADS simulation

This shows rich satellite features commensurate with a MQW system containing layers of both compressive and tensile strain. Also shown is a commercial RADS MERCURY dynamical simulation for the structure. The theoretical fit to the data is excellent for all three samples. However, in the case of the structure with net tensile strain, 1561, a close examination of the various satellite peaks reveals that the FWHM of the peaks are around 70arc sec for 1561 as compared to ~38 arc sec for the strain-compensated, 1554, and the less highly tensile wafer, 1563. This difference is also reflected in the transport properties since the Hall mobility is seen to decrease quite dramatically as the net tensile strain increases (see Table 1). Thus there is a limit to how much net tensile strain can be accommodated. [47]. Nevertheless, it is also clear from Table 1 that the transport properties of the QWIPs are excellent and that the high electron mobilities at room temperature attest to the abruptness and high quality of the various interfaces despite the low temperature used for growth.

The dopant confinement in these structures was studied by performing electrochemical Capacitance-Voltage (C-V) profiling. This technique gives information on dopant level as a function of depth in the sample. To study dopant activation, confinement and diffusion, two QWIP samples with well thickness of 30 Å were grown at ~350 ºC. The two structures, 1497 and 1500 are analysed below to demonstrate both effective carrier confinement in the well and the repeatability of QW doping/thickness throughout the QWIP structure. Table 2 below compares the average measured periodicity to the target (growth and DCXRD) periodicity.

Table 1. Sample growth and characterisation parameters.

	1554	1563	1561
QW periods (*W*)	15	15	25
QW thickness (L_w) nm	8	4.5	3
Average net strain per period (%)	0	-0.74	-1.3
Measured Hall QW sheet carrier Density (N_S) 10^{12} cm^{-2}	11	7.0	4.24
Hall mobility (μ) cm^2/Vsec	2000	1832	269
Fermi energy (E_F) (meV) inferred from Hall measurement	482	444	330

Table 2. Comparison of dopant periodicity

	Sample 1497	Sample 1500
Target periodicity / [Å]	265	315
Measured periodicity / [Å]	209	290

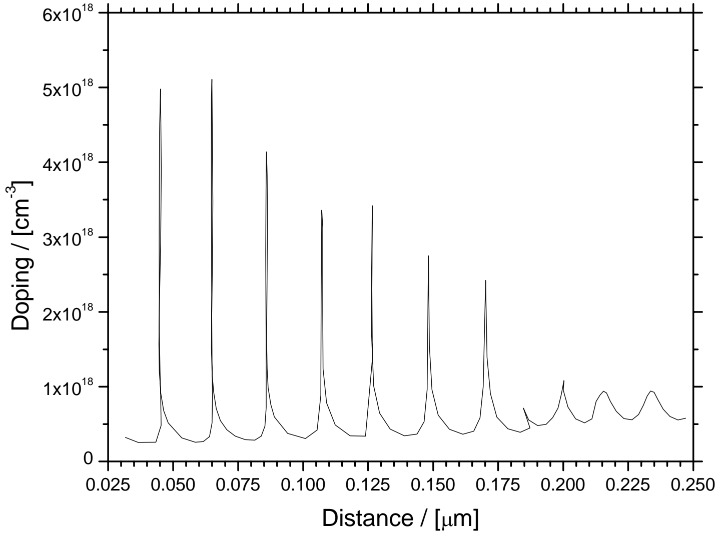


Figure 8. Sample 1497 C-V Profile

Excellent periodicity in the dopant position is noticeable in Figure 8 and the FWHM of the first 5 QW is < 30 Å (certainly limited by the resolution of the CV profiling

technique). As the depth profile proceeds deeper into the structure, the FWHM increases noticeably, the dopant level appears to decrease with some of the peaks appearing to loop. All of these effects are due to surface roughening as the etch depth increases and also to the oxidation of the AlAs confining barrier (looping). The target and measured periodicity are in very good agreement given that CV profiling is only accurate to within 10-15%. This demonstrates that the charges in the QW are activated (despite the low growth temperature used) and effective confinement is achieved.

3.4 Modelling the Intersubband Transitions

The conduction band profile and energy levels for the three structures were obtained using a three-band Kane **k.p** approximation that takes into account the effect of strain and bulk band nonparabolicity. Details of this model are given in [10]. A generic conduction band profile for the samples studied is shown in Figure 9.

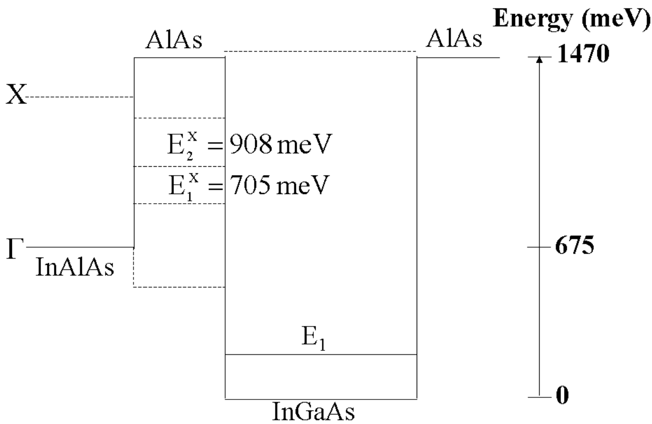


Figure 9. Generic band structure for the three samples 1561, 1563 and 1554; Γ energy levels for each of the samples are listed in Table 2

Table 3 summarizes the calculated Γ subband energy levels (at 300 K) in each of the InGaAs wells. Since the three samples 1561, 1563, and 1554 differ only in their InGaAs layer thickness, they are all characterized by the same E^X subband levels in the *X* QW in AlAs (calculated using an *X* valley effective mass of $m_{e\text{-}x} = 0.4m_0$). The energy level E_1^X in the AlAs layer is below E_2 in 1561, between E_2 and E_3 in 1563, and between E_3 and E_4 in 1554.

Table 3. Calculated Γ levels (300K) in the InGaAs QWs of samples 1561, 1563 and 1554

	Energy levels (meV)				
Sample No	$\mathbf{E_1}$	E_2	E_3	E_4	E_1+E_F
1554	82	276	521	786	564
1563	183	566	1000	1416	627
1561	301	873	1434		631

3.5 Measured Transitions

Figure 10 shows the room temperature absorbance of the three samples measured at 45° in waveguide geometry. Different polarisations of light are used to isolate allowed and forbidden geometries and the spectra are plotted as absorbance (which is defined as $-\log[T_{allowed}/T_{forbidden}]$ where T is the transmittance). No peaks were observed for parallel polarisation of the light (forbidden transition). It can be seen that the peak detection wavelength for these structures can be tuned from 2-7 μm (620-177 meV) by simply varying the InGaAs well width. These spectra were fitted using a Lorentzian lineshape and the resulting peak energies and their corresponding full-width at half-maximum (FWHM) are summarized in Table 4. For sample 1554, three peaks are observed as expected from the position of E_F given in Table 3. These peaks are assigned to $E_1 \rightarrow E_2$, $E_2 \rightarrow E_3$ and $E_3 \rightarrow E_4$ transitions in the InGaAs well A detailed analysis of this sample including the Lorentzian fit was reported in [46].

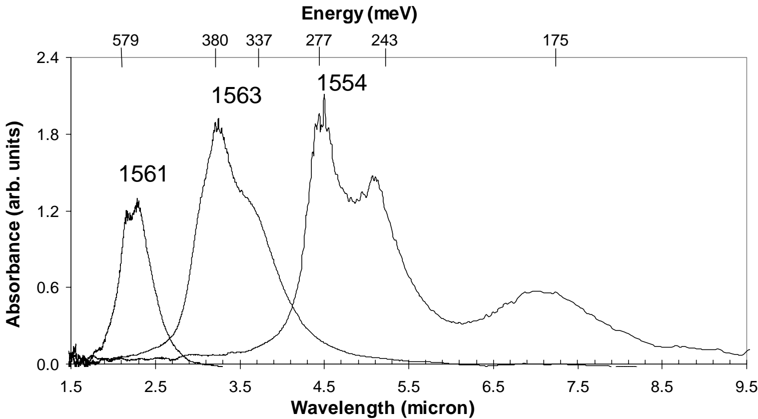


Figure 10. FTIR spectra of InGaAs/AlAs/InAlAs QWs: 1554, 1561 and 1563 (see Table 1)

Sample 1563 shows two distinct peaks: one at 337 meV and the strongest at 380 meV. In addition, there is a weak shoulder at 410 meV. From the position of E_F only two intersubband transitions, $E_1 \rightarrow E_2$ and $E_2 \rightarrow E_3$, are possible for this structure. Modelling suggests a variation of 1-2 monolayers (MLs) in thickness between the

different InGaAs wells could account for peaks at 337 and 380 meV both arising from the $E_1 \rightarrow E_2$ transition. The shoulder peak at 410 meV is assigned to the higher energy $E_2 \rightarrow E_3$ transition . Similarly the absorption spectrum for 1561 shows two peaks at 579 meV and 534 meV, one of which is in good agreement with the calculated transition energy of 572 meV for a well width of 30 Å. The separation between the observed peaks is within 50 meV and this can again be accounted for by a one ML fluctuation in thickness between different InGaAs wells.

An additional weak peak was seen for 1561 at ~207 meV in the perpendicular polarization. This peak is attributed to the $E_1^X \rightarrow E_2^X$ transition in the AlAs QW layer and is in excellent agreement with our calculated $E_1^X \rightarrow E_2^X$ transition energy of 203 meV, supporting use of $0.4m_o$ for the X-band electron effective mass.

Table 4. Peak absorption wavelengths and FWHM for transitions observed in samples 1554, 1561 and 1563. Calculated values are given in brackets

Sample No	$E_1 \rightarrow E_2$ (meV)	FWHM (meV)	$E_2 \rightarrow E_3$ (meV)	FWHM (meV)	$E_3 \rightarrow E_4$ (meV)	FWHM (meV)
1554	173 (195)	31.6	245 (246)	32	275.5 (265)	23.6
1563	337 380 (383)	54 54	410 (434)	54		
1561	534 579 (572)	60 68				

3.6 Predicted Temperature Performance

The samples described above demonstrate the usefulness of this material system for accessing the whole of the MWIR region and they have allowed investigation of the fundamental properties of this material system [46]. However, they are in themselves too highly doped to be used for QWIPs. Nevertheless, extrapolating from the observed 205 K performance of the lattice-matched InGaAs/InAlAs device operating at 3.4μm [7], it is possible to predict the temperature performance which might readily be achieved with strain-balanced structures.

It is the dark current, I_D, which limits the temperature performance and for thermionic emission and related processes this is given by

$$I_D = \mathrm{eA} v_d n_{th} \tag{1}$$

where v_d is the drift velocity and n_{th} is the thermal generation rate. The thermal generation rate can be approximated to [52]

$$n_{th} \approx \frac{\mathrm{m}^*}{\mathrm{L}\pi\hbar^2} kT.\exp[-(\Delta E_c - E_F - E_1)/kT \quad (2)$$

where ΔE_c is the band discontinuity between the wide AlInAs and the InGaAs QW and E_1 is the first energy level in the well. This assumes a band structure similar to that in Figure 11 with two confined states and the E_2 level just above the InAlAs band edge. (The significance of E_2 being just above the outer barrier is that this enables the photoexcited carrier to easily tunnel out of the well through the thin high barrier and be transported through the structure at low fields.)

Using a structure similar to 1561 but with doping of 10^{18}cm^{-3} (a QW of 42nm and about 80% indium in the InGaAs well) it is possible to achieve a device with (ΔE_c -E_F –E_1) ~ 380meV which operates at a wavelength of 3.0μm and at room temperature; 47nm and 65% In would reach 3.4μm. At 250K, n_{th} for this device is predicted to equal the 205 K value of n_{th} reported in [7]. For the 3 μm design, the same value of n_{th} is reached at 300 K. Devices have been grown with doping levels similar to this and a BLIP temperature above 250 K has been measured for the sub-3.5μm device designs [53].

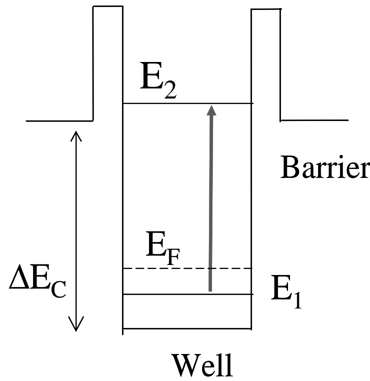


Figure 11. Conduction band edge profile showing two confined energy states and low Fermi energy for high temperature performance

4 Enhanced Detector Performance and Functionality

In this Sectionwe consider ways of enhancing QWIP performance by using specific growth and doping conditions. For example stepped wells can give an element of normal incidence absorption and allow tuning of the absorption wavelength under applied field (see Section 4.1). The position of the dopants relative to each QW can affect the responsivity under photovoltaic (PV) operating conditions and this is discussed in Section 4.2. In addition QWIPs are inherently well suited to high speed applications due to the short carrier lifetime and this is outlined in Section 4.3.

4.1 Asymmetric Wells

4.1.1 Normal Incidence Absorption

For a simple rectangular QW, we can only vary the well width and the barrier height (composition). In addition, the dipole matrix elements of transitions for odd-to-odd or even-to-even (*i.e.*, 1-3 or 2-4) quantum numbers disappear (meaning these kind of transitions are not allowed) since the envelope functions of these energy states have the same parity due to the symmetry of the well. However, the symmetry of the well can be broken by applying a strong electric field [54] or doping the QW heavily [55]. A high bias can produce a large leakage current, which will degrade the performance of the device while for the heavily doped case, the large thermionic emission associated with the high doping densities will prevent operation at higher temperatures. An alternative approach to achieving normal incidence operation is to use an asymmetric structure such as a stepped QW as shown in Figure 12.

The introduction of the step inside the well results in normally forbidden transitions becoming allowed. Furthermore, the transition energies and related oscillator strengths can now be tuned "independently" by changing the well width, the step width, and the step height. This gives an advantage over the high electric field and doping induced asymmetry in a simple square well [56-58]. A component of normal incidence absorption was recently observed in a stepped DBQW using InGaAs/GaAs QWs with AlAs/AlGaAs barriers and operating at 3μm [59,60].

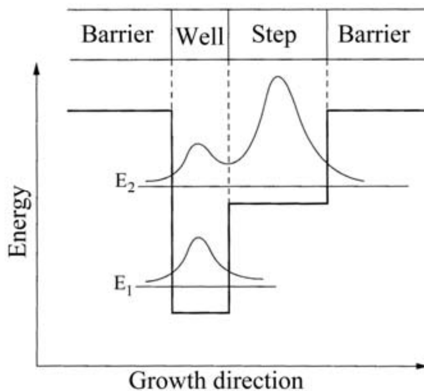


Figure 12. Conduction band profile of the asymmetric step QW

4.1.2 Voltage Tuning via the Stark Effect

When the structure in Figure 12 is illuminated with photons of energy $\hbar\omega$, electrons can either be excited from E_1 to E_2 the bound-to-bound transition or from E_1 to the delocalised levels above the $Al_xGa_{1-x}As$ barrier (bound-to-continuum transition). If a

photon excites a bound-to-bound (b) transition (from E_1 to E_2), the photoexcited carrier can then tunnel from E_2 through the tip of the triangular barrier in the presence of an applied electric field where it can then contribute to the photocurrent.

Due to the asymmetry of the structure, the subband level energy spacing of E_1 and E_2 can either be moved closer together or further apart depending on the direction of the applied field (Figure 13). Hence, we can shift to longer (red) or shorter (blue) wavelengths simply by changing the bias direction as shown below [61-65].

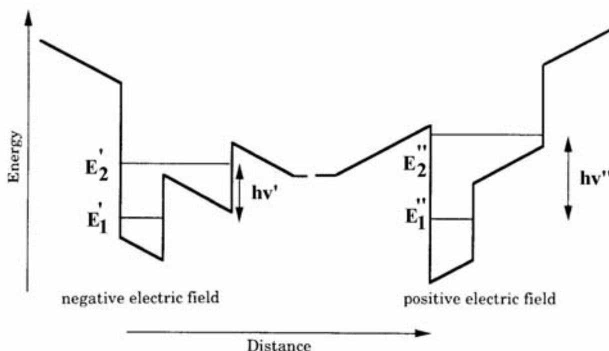


Figure 13. Modification of the conduction band profile of a stepped QW by application of an electric field (Stark effect [28]). The change in energy levels and hence in the optical transition energy depends on the polarity of the applied electric field

Large Stark shifts arise due to optical transitions from the states in the narrow well to the states in the wider well. The subband energies can either be moved closer or further apart depending on the direction of the applied field This is because the energy levels associated with different well widths have different field dependencies, hence, we can shift to longer or shorter wavelengths simply by changing the bias direction as shown in Figure 13 [61-65].

4.2 Photovoltaic Operation

Photovoltaic operation of QWIPs is preferred to photoconductive mode as it results in reduced noise and thus enhanced signal to noise ratio and detectivity. However the responsivity of QWIPs tends to be low under PV conditions unless care is taken to optimize the potential profile in the transport region between QWs. Several authors report some degree of asymmetry in the photoresponse and dark I-V characteristics of QWIPs which arises as a result of interface roughness asymmetry [7, 66]. The electron tunnelling rate is highly sensitive to the thickness of the AlAs layers and also to submonolayer differences in the interface composition. Asymmetry between the two AlAs/well interfaces and/or the AlAs/barrier interfaces can arise in the growth process. This results in asymmetric dark I-V characteristics and photoresponse (Schnieder *et al.*) Luna et al. also observed significant photoresponse at zero bias for a GaAs/AlAs/AlGaAs DBQW that was modulation

doped in the AlGaAs barrier nearest to the substrate [67]. They attribute this behaviour to enhanced asymmetry of the structure due to silicon segregation in the modulation doped sample which does not occur in well doped QWIPs [68]. Combining this doping scheme with high or wide AlGaAs barriers (35% Al or 50nm) reduced the dark current and resulted in $D^* = 7.5\times10^{11}cm^2\ Hz^{1/2}/W$ at 77K at 0V. In addition photovoltaic operation could be maintained up to 180K making this a promising route to high operating temperature QWIPs.

4.3 High-speed Operation

A key advantage of QWIPs has to be their intrinsic photocurrent response time, which is extremely short. The lifetime of optically excited carriers has been estimated to be smaller than 7 ps [8,69]. The intersubband transitions thus provide the opportunity to exceed the 1 to 2 GHz high frequency cut-off of commonly available MCT interband-detectors by more than an order of magnitude. This makes QWIPs ideally suited for heterodyne spectroscopy, since a bandwidth of at least 20 GHz is necessary to provide continuous tunability using the individual emission-lines of a CO_2 laser as a local oscillator [69]. Liu *et al.* [70] have demonstrated a heterodyne detection of two CO_2 laser beam signals with a frequency separation up to 82 GHz by mixing the down converted laser beams with a microwave signal.

Free space communications using quantum cascade lasers (QCLs) has also been shown to be limited in response speed by the available interband detectors [71]. QWIPs or QCLs in detector mode have been investigated as a potential solution to this problem [72,73]. QCLs in detection mode have a poor responsivity – about 250 times less than state-of-the-art QWIPs – due to the high capture and low escape probability of the active well [73]. However, the high frequency response combined with the ability to shift the absorption peak significantly at low voltage offers significant advantage for some applications. Perhaps most importantly this work highlights the potential for integration of QCLs and intersubband detectors to provide new directions for future telecommunications systems.

5 Conclusions

Intersubband absorption in the 3-5µm region has been demonstrated in a number of material systems based on both GaAs and InP substrates. Access to more than a small part of the MWIR involves the incorporation of strained layers. This has been successfully dealt with by the use of graded buffer layers [18,19,21] in which the strain is deliberately relaxed, by pseudomorphic growth of strained layers [7-9] and by strain-balancing [10,24,46,47]. Strain-balancing removes the restriction on total number of periods which can be grown and also obviates the need for the growth of thick buffer layers.

Achieving short wavelengths requires high barriers and many authors have used thin AlAs barriers in combination with wider low barriers to give both high transition energies and bound to quasi-bound transitions for high responsivity at low bias [29-44]. Intersubband transitions in the strained AlAs-$In_xGa_{(1-x)}As$ system have

been observed in structures grown on InP and on GaAs. With indium contents from 0.4-0.53 on InP, Smet [45] and Hirayama [8] obtained intersubband transitions as low as 1.55 μm but only in structures having 3ML (~ 9Å) well thickness and operation at ~1.8 to 2 μm could be achieved using well widths of 6 ML (~18 Å). These samples consisted of only one or two periods. With such structures there is a concern with the absorption strength as electron leakage from the first Γ-subband to the X-subband in the AlAs will cause a dramatic decrease in the electron population in the Γ-branch. This puts a lower limit on how thin the QW must be in order to have fully functional devices. Asano [41] has shown theoretically that at well widths below 7ML, the absorption should be almost completely quenched for $AlAs/In_{0.4}Ga_{0.6}As$ at a wavelength of ~ 1.9 μm.

If AlAs barriers are providing the tensile strained layer on InP very high indium content (up to 84%) InGaAs QWs can be used to produce strain-balanced structures. In this case, because of the much deeper well, intersubband absorption at 2 μm is reached with fairly wide wells (10 ML). These designs achieve both high periodicity and large Γ to X separation leading to devices with strong intersubband absorption magnitude and making the prospect of realising QWIP devices that work near room temperature a real possibility. The key to high temperature operation is maximising the barrier/well discontinuity for the chosen wavelength of operation. For any given wavelength the structure is then designed to place the excited state just above the InAlAs continuum and the ground state at an appropriate level in the well. Lee [9] reported 205K operation for a 0.5eV band discontinuity (lattice-matched InGaAs/InAlAs on InP) but with a Fermi level ~0.27eV below the barrier edge. With the much higher band offsets (~0.675eV) 250K operation should be achievable for 3.5μm devices with near room temperature operation at shorter wavelengths being a real possibility.

Acknowledgements

The materials growth and assessments have greatly benefited from the help of Dr J. Sly, Colin Mitchell and Veronica Hewer. Some of the work described in this chapter was funded by the UK EPSRC (Engineering and Physical Sciences Research Council) and this support is gratefully acknowledged.

References

1. Borman, S. Array detectors are transforming optical spectroscopy. C&E News, Mar. 18, 1996.
2. Rogalski A. Quantum well photodetectors in infrared detector technology. J. Appl. Phys. 2003; 93: 4355-4391
3. Rogalski A. Infrared detectors: An overview. Infrared Phys.Technol. 2002; 43:187-210.
4. Schneider H, Walther M, Schönbein C, Rehm R, Fleissner J, Pletschen W, Braunstein J, Koidl P, Weimann G, Ziegler J, and Cabanski W. QWIP FPAs for high-performance thermal imaging. Physica E. 2000; 7: 101-107.
5. Nash GR, Gordon NT, Hall DJ, Ashby MK, Little JC, Masterton G, Hails JE, Giess J, Haworth L, Emeny MT, Ashley T. Infrared negative luminescent devices and higher operating temperature detectors. Physica E 2004; 20:540-547.

6. Liu HC, Buchanan M, Wasilewski ZR. How good is the polarization selection rule for intersubband transitions? Appl. Phys. Lett. 1998; 72:1682-1684.
7. Liu HC, Dudek R, Shen A, Dupont E, Song, CY, Wasilewski ZR, Buchanan M. High absorption (>90%) quantum-well infrared photodetectors. Appl. Phys. Lett. 2001; 79:4237-4239
8. Hirayama Y, Smet JH, Peng L-H, Fonstad CG, Ippen EP. Feasibility of 1.55μm intersubband photonic devices using InGaAs/AlAs pseudomorphic quantum well structures. Jpn. J. Appl. Phys. Part 1 1994; 33:890-895.
9. Lee JH, Chiang JC, Li SS, and Kannam PJ. An AlAs/InGaAs/AlAs/InAlAs double-barrier quantum well infrared photodetector operating at 3.4 μm and 205K. Appl. Phys. Lett. 1999; 74:765–767.
10. Lai KT, Missous M, Gupta R and Haywood SK. Intersubband absorption in strain-compensated InAlAs/AlAs/InxGa(1-x)As quantum wells grown on InP. J. Appl. Phys. 2003; 93: 6065-6067
11. Burkle L, Fuchs F. InAs/[GaInSb] superlattices: a promising material system for IR detection. Handbook of Infra-red Detection Technologies: Chapter 5
12. Fuchs F, Rehm R, Fleißner J, Schmitz J, Pfahler Ch, Walther M, Cabanski W, Eich D, Finck M, Rode W, Wendler J, Wollrab R and Ziegler J. Focal plane arrays based on InAs/GaSb superlattices for the mid-infrared spectral range presented at MIOMD VI, St.Petersburg Sept 2003
13. Levine BF. Quantum-well infrared photodetectors. J. Appl. Phys. 1993; 74: R1-R81.
14. Levine BF, Gunapala SD, and Kopf RF. Photovoltaic GaAs quantum well infrared detectors at 4.2μm using indirect $Al_xGa_{1-x}As$ barrier. Appl. Phys. Lett. 1991; 58:1551-1553.
15. Pan JL, West LC, Walker SJ, Malik RJ, Walker JF. Inducing normally forbidden transitions within the conduction band of GaAs quantum well. Appl. Phys. Lett. 1990; 57:366-368.
16. Tsai KL, Chang KH, Lee CP, Huang KF, Huang JS, Chen HR. Two–color infrared photodetector using GaAs/AlGaAs and strained InGaAs/AlGaAs multiquantum wells. Appl. Phys. Lett. 1993; 62:3504-3506.
17. Fiore A, Rosencher E, Bois P, Nagle J, and Laurent N. Strained InGaAs/AlGaAs quantum well infrared detectors at 4.5 μm. Appl. Phys. Lett. 1994; 64:478-480.
18. Chui HC, Lord SM, Martinet E, Fejer MM, Harris JS, Jr. Intersubband transitions in high indium content InGaAs/AlGaAs quantum wells. Appl. Phys. Lett. 1993; 63:364-366
19. Chui HC, Martinet E, Fejer MM, Harris JS, Jr. Short wavelength intersubband transitions in InGaAs/AlGaAs quantum wells grown on GaAs. Appl. Phys. Lett. 1994; 64:736-738
20. Tidrow MZ, Chiang JC, Li SS, Bacher K. A high strain two stack two-color quantum well infrared photodetector. Appl. Phys. Lett. 1997: 70:859–861.
21. Choi KK, Bandara SV, Gunapala S.D, Liu W.K, Fastenau J.M. Detection Wavelength of InGaAs/AlGaAs Quantum Wells and Superlattices. J. Appl. Phys. 2002; 91:551-564.
22. Goldberg A, Wang S, Sundaram M, Uppal P, Winn M, Milne G, Stevens M. Proceedings of the 1999 Meeting of the MSS Specialty Group on Infrared Detectors, Lexington, MA (1999), edited by J. Robinson (Infrared Information Analysis Center, Ft. Belvoir, VA 2000), p.87.
23. Hasnain G, Levine BF, Silvo DL, and Cho AY. Mid-infrared detectors in the 3-5 μm band using bound to continuum state absorption in InGaAs/InAlAs multiquantum well structures. Appl. Phys. Lett. 1990; 56:770-772.
24. Maimon S, Cohen GM, Finkman E, Bashir G, Ritter D, Schacham SE. Strain Compensated InGaAs/InGaP Quantum Well Infrared Photodetector for Midwavelength Band Detection. Appl. Phys. Lett. 1999; 73: 800-802.

25. Sengupta DK, Jackson SL, Ahmari D, Kuo HC, Malin JI, Thomas S, Feng M, Stillman GE, Chang YC, Li L, Liu HC. P-type InGaAs/InP quantum well infrared photodetector with peak response at 4.55 μm. Appl. Phys. Lett. 1996; 69:3209-3211.
26. Liu HC, Oogarah T, Dupont E, Wasilewski ZR, Byloos M, Buchanan M, Szmulowicz F, Ehret J, Brown GJ. P-type Quantum Well Infrared Photodetectors Covering Wide Spectrum", Elect. Lett. 2002; 38: 909-911.
27. Pan J. L., and Fonstad C. G. Comparison of Hole and Electron Intersubband Absorption Strengths for Quantum Well Infrared Photodetectors. IEEE Trans on Electron Devices. 2000; 47:1325-1329.
28. Levine BF. Zussman A, Gunapala SD, Asom MT, Kuo JM and Hobson WS. Photoexcited Escape Probability, Optical Gain, and Noise in Quantum Well Infrared Photodetectors. J. Appl. Phys.1992; 72: 4429-4443.
29. Schneider H, Fuchs F, Dischler B, Ralston JD, and Koidl P. Intersubband absorption and infrared photodetection at 3.5 and 4.2 μm in GaAs quantum wells. Appl. Phys. Lett. 1991; 58:2234-2236.
30. Liu HC, Wilson PH, Lamm M, Steele AG, Wasilewski ZR, Li JM, Buchanan M, and Simmons JG. Low Dark Current Dual Band Infrared Photodetector Using Thin AlAs Barriers and Γ-X Mixed Intersubband Transition in GaAs Quantum Wells. Appl. Phys. Lett. 1994; 64:475-477.
31 Schneider H, Koidl P, Fuchs F, Dischier B, Schwarz K, and Ralston JD. Photovoltaic intersubband detectors for 3-5 μm using GaAs quantum wells sandwiched between AlAs tunnel barriers. Semicond. Sci. Technol. 1991; 6:C120-C123.
32. Ralston JD, Schneider H, Gallagher DFG, Kheng K, Fuchs F, Bittner P, Dischler B, and Koidl P. Novel molecular-beam epitaxially grown GaAs/AlGaAs quantum wells structures for infrared detection and integrated optics at 3-5 and 8-12 μm. J. Vac. Sci. Technol. B 1992; 10:998-1001.
33. Kiledjian MS, Schulman JN, and Wang KL. Absorption in GaAs/$Ga_{1x}Al_xAs$ quantum wells with resonant barriers for improved responsivity. Phys. Rev. B 1991; 44:5616-5621.
34. Neu G, Chen Y, Deparis C, and Massies J. Improvement of the carrier confinement by double-barrier GaAs/AlAs/(Al,Ga)As quantum well structures. Appl. Phys. Lett. 1991; 58:2111-2113.
35. Wang YH, Chiang JC, Li SS, and Ho P. A GaAs/AlAs/AlGaAs and GaAs/AlGaAs stacked quantum well infrared photodetector for 3-5 and 8-14 μm detection. J. Appl. Phys. 1994; 76:2538-2540.
36. Tsai KL, Lee CP, Chen PC, Tsang JS, Tsai CM, and Fan JC. The effect of barrier structure on the performance of double barrier quantum well infra-red photodetectors. Solid-State Electron. 1996; 39:201-204.
37. Wu WG, Jiang DS, Cui LQ, Song CY, and Zhuang Y. Structural and photoelectric studies on double barrier quantum well infrared detectors. Solid-State Electron. 1999; 43:723-727.
38. Sung B, Chui HC, Fejer MM, and Harris JS., Jr. Near-infrared wavelength intersubband transitions in high indium content InGaAs/AlAs quantum wells grown on GaAs. Electron. Lett. 1997; 33:818-820.
39. Lenchyshyn LC, Liu HC, Buchanan M, and Wasilewski ZR. Mid-wavelength infrared detection with $In_xGa_{1-x}As/Al_{0.45}Ga_{0.55}As$ multiple quantum well structures. Semicond. Sci. Technol. 1995; 10:45-48.
40. Liu HC, Buchanan M and Wasilewski ZR. Short wavelength (1-4 μm) infrared detectors using intersubband transitions in GaAs-based quantum wells. J. Appl. Phys. 1998; 83:6178-6181; erratum 2002; 91:10230.
41. Asano T, Noda S, Abe T, and Sasaki A. Jpn J.Appl. Phys. 1996; 35: 1285-1291 Near-infrared Intersubband transitions in InGaAs/AlAs Quantum Wells

42. Chin VWL, Egan RJ, Osotchan T, Vaughan MR, and Anderson SC. Noncontact thickness and composition assessment of a strained AlGaAs/AlAs/InGaAs double barrier multiple quantum well structure. J. Appl. Phys. 1996; 80:2521-2523.
43. Jancu JM, Pellegrini V, Colombelli R, Beltram F, Mueller B, Sorba L, and Franciosi A. Quantum tailoring of optical transitions in $In_xGa_{1-x}As/AlAs$ strained quantum well. Appl. Phys. Lett. 1998; 73:2621-2623.
44. Garcia CP, Nardis AD, Pellegrini V, Jancu JM, Beltram F, Müeller BH, Sorba L, and Franciosi A. 1.26 μm intersubband transitions in $In_{0.3}Ga_{0.7}As/AlAs$ quantum wells. Appl. Phys. Lett. 2000; 77:3767-3769.
45. Smet JH, Peng LH, Hirayama Y and Fonstad CG. Electron Intersubband Transitions to 0.8 eV (1.55 μm) in InGaAs/AlAs Single Quantum Wells. Appl. Phys. Lett. 1994; 64:986-987.
46. Gupta R, Lai KT, Missous M, and Haywood, SK. Subband Non-parabolicity estimated from intersubband absorption in highly multiple quantum wells. Phys Rev B 2003; 69: 033303 (1) -033303 (4)
47. Missous M, Mitchell C, Sly J, Lai KT, Gupta R. and Haywood SK. Highly strained InxGa1-xAs-InyA11-yAs (x>0.8,y<0.3) layers for short wavelength QWIP and QCL structures grown by MBE, Physica E 2004; 20: 496-502
48. Lai KT, Gupta R, Missous M, and Haywood SK. Intersubband Absorption from 2-7 μm in Strain-compensated Double-barrier $In_xGa_{1-x}As$ Multiquantum Wells. Semicon. Sci. Tech. 2004; 19:1263-1267.
49. Georgiev N, Dekorsky T, Eichorn F, Helm M, Semtsiv MP & Masselink WT. Short wavelength intersubband absorption in strain compensated InGaAs/AlAs quantum well structures grown on InP. Appl.Phys.Lett. 2003; 87:210-213
50. Missous M. Stoichiometric low-temperature GaAs and AlGaAs: a reflection high-energy electron-diffraction study. J. Appl Phys. 1995; 78: 4467-4471
51. Missous M. Stoichiometric low temperature (SLT) GaAs and AlGaAs grown by molecular beam epitaxy. Microelectronics J. 1996; 27: 393-409
52. Choi KK, Lee CY, Tidrow MZ, Chang WH, Gunapala SD. Performance assessment of quantum-well infrared photodetectors. Appl.Phys.Lett 1994; 65: 1703-1705
53. Haywood SK, Lai KT, Gupta R, Missous M. High Temperature Mid-IR Photodetectors for Gas Sensing Applications. Invited paper at MIOMD VI, St.Petersburg, Sept 2003
54. Pan JP, West LC, Walker SJ, Malik RJ and Walker JF. Inducing Normally Forbidden Transitions within the Conduction Band of GaAs Quantum Well. Appl. Phys. Lett. 1990; 57:366-368.
55. Kheng K, Ramsteiner M, Schneider H, Ralston JD, Fuchs F and Koidl P. Two-color GaAs/(AlGa)As Quantum Well Infrared Detector with Voltage-tunable Spectral Sensitivity at 3-5 and 8-12μm. Appl. Phys.Lett. 1992; 61: 666-668.
56. Yuh PF and Wang KL. Optical Transitions in a Step Quantum Well. J. Appl. Phys. 1989; 65:4377-4381.
57. Mii YJ, Wang KL, Karunasiri RPG and Yuh PF. Observation of Large Oscillator Strengths for Both 1-2 and 1-3 Intersubband Transitions of Step Quantum Well. Appl. Phys. Lett. 1990; 56: 1046-1048.
58. Li HS, Chen YW, Wang KL and Lie DYC. Intersubband Transitions in Pseudomorphic InGaAs/GaAs/AlGaAs Multiple Step Quantum Wells. J.Vac.Sci.Tech. B. 1993; 11: 1840-1843.
59. Lai KT, Haywood SK, Gupta R, and Missous M. Enhanced Intersubband Absorption in Stepped Double Barrier Quantum Wells. Elect. Lett. 2002; 38:529-530.
60. Lai KT, Haywood SK, Gupta R, and Missous M, Observation of Intersubband Absorption in the Forbidden Polarisation for a Stepped Double Barrier Quantum Well. IEE Proc.-Opto. 2003; 150: 377-380.

61. Parihar SR, Lyon SA, Santos M and Shayegan M. Voltage Tunable Quantum Well Infrared Detector. Appl. Phys. Lett. 1989; vol. 55: 2417-2419.
62. Mii YJ, Karunasiri RPG, Wang KL, Chen M and Yuh PF. Large Shark Shifts of the Local to Global State Intersubband Transitions in Step Quantum Well. Appl. Phys. Lett. 1990; 56: 1986-1988.
63. Karunasiri RPG, Mii YJ and Wang KL. Tunable Infrared Modulator and Swtich using Stark Shift in Step Quantum Wells. IEEE Elect. Device Lett. 1990; 11: 277-279.
64. Martinet E, Luc F, Rosencher E, Bois Ph and Delaitre S. Electrical Tunability of Infrared Detectors using Compositionally Asymmetric GaAs/AlGaAs Multiquantum Wells. Appl. Phys. Lett. 1992; 60: 895-897.
65. Chen WQ and Andersson TG. Intersubband Transitions for Differently Shaped Quantum Wells Under an Applied Electric Field. Appl. Phys. Lett. 1992; 60: 1591-1593.
66. Larkins EC, Schneider H, Ehret S, Fleißner J, Dischler B, Koidl P and Ralston JD. Influences of MBE growth processes on photovoltaic 3–5 _m intersubband photodetectors. IEEE Trans. Electron. Devices. 1994; 41: 511,
67. Luna E, Guzman A, Sanchez-Rochas JL, Calleja E and Munoz E. Infrared Physics and Technology 2003; 44: 383-390. Modulation Doping in 3-5µm GasAs/AlAs/AlGaAs Double Barrier quantum well infrared photodetectors: an alternative to achieve high PV performance and high temperature detection
68. Luna E, Guzman A, Trampert A, Sanchez-Rojas JL, Calleja E. On the growth conditions of 3–5 µm well-doped AlGaAs/AlAs/GaAs infrared detectors and its relation to the photovoltaic effect studied by transmission electron microscopy. Infrared Physics & Technology. 2003; 44: 391–398
69. Ehret S, Schneider H, Fleissner J, Koidl P. Untrasfast intersubband photocurrent response in quantum well infrared photodetectors. Appl. Phys.Lett. 71 (1997) 641-643.
70. Liu HC, Li J, Brown ER, McIntosh KA, Nichols KB, Manfra MJ. Quantum well intersubband heterodyne infrared detection up to 82GHz. Appl. Phys. Lett. 1995; 67: 1594-1596.
71. Blaser S., Hofstetter D, Becker M, and Faist J. Free-space optical data link using Peltier-cooled quantum cascade laser. Elect. Lett. 2001; 37:778-780.
72. Asano T, Tamura M, Yoshizawa S, and Noda S. Pump-probe measurement of ultrafast all-optical modulation based on intersubband transition in *n*-doped quantum wells. Appl. Phys. Lett. 2000; 77:19-21. InGaAs/AlAs
73. Hofstetter D, Beck M, Faist J. Quantum-cascade-laser structures as photodetectors Appl.Phys.Lett 2002; 81: 2683-2685

Negative Luminescence

T. Ashley and G.R. Nash
QinetiQ
Malvern, Worcestershire, UK

1 Introduction

An object inside a blackbody enclosure will exchange radiation with the enclosure. In equilibrium, it will absorb and emit radiation at the same rates. If the equilibrium is disturbed such that the number of excited states is increased, they will return to their initial values by increased emission of radiation, amongst other processes, whilst the absorption will be essentially unchanged, leading to a net emission which is referred to as luminescence. If the disturbance is such that the number of excited states is lower than the equilibrium level, then the emission is reduced below the absorption level, so that there is a net absorption of radiation, which is referred to as 'negative luminescence'. The concept was first formulated by Antonov-Romanovsky *et al.* [1] in the context of systems with discrete energy levels, such as atoms or molecules.

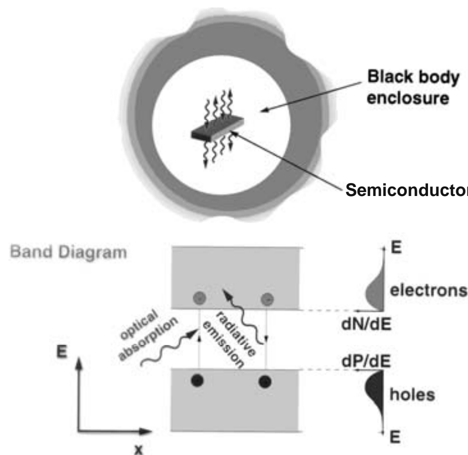


Figure 1. Exchange of radiation between a semiconductor and its environment

In the case of semiconductors, the emitted radiation is mainly due to recombination of electron-hole pairs, as illustrated in Figure 1. From the principle of detailed balance, this is offset by the absorption of radiation with photon energies greater than the band-gap leading to generation of electron-hole pairs. If such a material is held in a non-equilibrium state such that the product $np < n_o p_o$, where n, p, n_o and p_o refer to the electron and hole densities in non-equilibrium and

equilibrium situations respectively, then the radiative recombination of carriers and the associated emission of photons are reduced below the absorption rate.

Negative luminescence (NL) was first observed in semiconductors by Ivanov-Omskii *et al.* [2,3] in experiments involving the magneto-concentration effect in slabs of InSb. Observations were extended to germanium at elevated temperatures by Kessler and Mangelsdorf [4,5]. Additional experimental and theoretical work on the narrow-gap semiconductors, InSb and CdHgTe, was performed by Bolgov *et al.* [6,7], extended by Malyutenko *et al.* [8] and reviewed by Berdahl *et al.* [9].

The maximum negative luminescent power, P_{NL}, which can be developed from a device is equal to the blackbody power from an object at the temperature of the device which is lying within the absorption band of the device, *i.e.* $P_{NL} = \int S_\lambda \varepsilon_\lambda d\lambda$, where S_λ is the blackbody spectral radiant emittance and ε_λ is the semiconductor emissivity. Therefore, the narrow-gap semiconductors are particularly suited to the observation of negative luminescence because their band-gap corresponds to the peak in the blackbody emission spectrum at room temperature, as shown in Figure 2. The maximum NL powers available for specific IR bands of interest for devices at a range of temperatures are given in Table 1.

Table 1. Comparison of NL powers in various IR bands for devices at a range of temperatures

Background temperature	Power ($mWcm^{-2}$) in Wavelength range 3 – 5 µm	8 – 12 µm	12 – 24 µm
-40°C	0.03	3	8
-20°C	0.08	5	11
0°C	0.2	7	15
+20°C	0.5	11	18

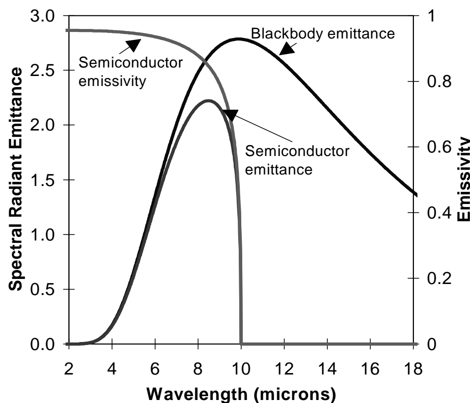


Figure 2. Relationship between the room temperature blackbody spectrum and the spectral response of a semiconductor with cut-off wavelength of 10 µm

2 Negative Luminescent Parameters

2.1 Key Parameters

Important characteristics of a NL device are its spectral properties and its negative luminescent efficiency (NLE); NL quantum efficiency (NLQE) and NL power efficiency (NLPE). The first of these are easily measured using standard spectroscopic techniques, such as Fourier Transform Infrared Spectroscopy (FTIR). In a semiconductor, the primary absorption and emission mechanisms are band-to-band, so that the negative and positive luminescence has similar spectral shapes. Often the spectral characteristics are shown in terms of the measured spectral emittance, which is the radiated power per unit wavelength (or wavenumber) per unit area. This is usually given in units of W μm^{-1} cm^{-2} (or W cm^{-1} cm^{-2}).

The negative luminescence efficiency (NLE) is the proportion of the radiation falling on the device that is absorbed, but can also be expressed as the percentage of the power radiated by the device in equilibrium that is suppressed when the device operates in NL mode (the ratio of the measured NL emittance to the equilibrium blackbody emittance). Important factors that limit the NLE efficiency of a device are the absorption within the material, the fill factor (the percentage of the device which is active) and reflection from the surface of the semiconductor. Narrow-gap semiconductors have relatively high refractive indexes n, so that the reflectivity R is typically around 30% for normal incidence radiation, where $R = \{(n-1)/(n+1)\}^2$. A broad-band anti-reflection coating can be used to decrease dramatically the reflectivity. NL efficiencies around 90% have been achieved in a range of devices, see Section s 3 and 4 of this chapter, although care has to be taken when dealing with values quoted in the literature to determine whether quoted NL efficiencies ignore reflections from the surface (*i.e.* internal NL efficiency is being quoted).

The Naval Research Laboratory, in Washington, has developed a novel self-referenced optical modulation technique [10] to measure emittance. Here the emission from the device in NL mode is referenced to that from the same device in equilibrium so the measurement does not depend on knowing the collection efficiency, sample reflectivity, or active area. This measurement yields the “internal” NL efficiency rather than the absolute efficiency, since it is insensitive to whether the device has an anti-reflective coating (with zero reflectivity, the internal and external NL efficiencies become equivalent). The “apparent” temperature of a NL device is the temperature at which an equilibrium blackbody would emit the same intensity, and is a useful Figure-of-merit for many applications. It can be determined from measured values of NL efficiency, for example a NL device (at 20 ºC) with a NLE of 80% would have an apparent temperature of 2 ºC, whereas a similar device having a NLE of 95% would have an apparent temperature of -25 ºC. Often apparent temperatures are quoted at a single wavelength, whereas for most applications it is important to determine the apparent temperature of a device over the spectral region of interest. For example, if a device is being viewed by an infrared camera with a 3-5 μm response, it is necessary to calculate the percentage of blackbody radiation suppressed over the entire wavelength range.

The above description assumes that the device is at the same physical temperature as its surroundings, so that it is in equilibrium before a bias is applied. Clearly if, for example, the device is physically hotter than its environment, it will have a net emission of radiation even without an electric field applied. We therefore define the NLE by the difference between the biased and unbiased cases, referenced to physical temperature of the device.

The final two parameters, the NL quantum efficiency (NLQE) and NL power efficiency (NLPE) are defined similarly to conventional, positive luminescence. The NLQE is the ratio between the NL photon flux and the drive current, whilst the NLPE is given by the ratio of NL power to input electrical power. The power consumption needed to achieve a certain apparent temperature is important for device applications. This depends on factors such as series resistance, the ratio between internal radiative and non-radiative processes, and optical efficiency. The effect of internal processes on device efficiency is discussed in Section 5.1.2. The optical efficiency can be improved by use of optical concentrators, as described in the following Section .

2.2 Optical Concentrators

One way in which the currents required to drive a device can be reduced is to use optical structures such as optical concentrators or lenses so that a smaller area device can be used to absorb the same amount of radiation. Such structures also increase the efficiency of the positive luminescence, as one of the most important factors

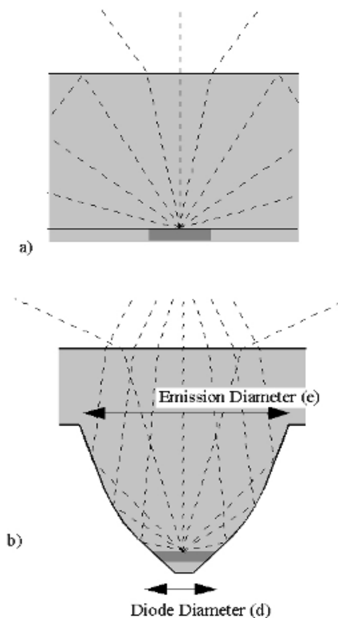


Figure 3. a) Standard LED **b**) Optical concentrator.

limiting the optical efficiency of semiconductor light emitting devices is total internal reflection of the generated light, where photons incident to the surface at angles greater than the critical angle are reflected back into the semiconductor and absorbed. This is shown schematically in Figure 3a, where light from an active layer is emitted through the substrate. Most semiconductors have a large refractive index and hence a small critical angle. Narrow-gap semiconductors have particularly large refractive indexes and correspondingly smaller critical angles. For example, InSb has a refractive index of 4, giving a critical angle of only 14°; therefore only 5% of the light generated from an active layer would be emitted from the semiconductor.

One form of optical concentration is the use of immersion lenses. These are ideally made from material of the same refractive index as the source, and also require intimate contact with the source (gap $< \lambda / 10$) to minimize attenuation. Epoxy-resin immersion lenses have been used to improve the efficiency of GaAs infrared LEDs, where epoxy-resin is deposited on the devices and surface tension causes the epoxy to form lens shapes. Unfortunately, the use of epoxy lenses for devices emitting in the mid-infrared range is inefficient due to absorption. Separate immersion lenses are ill-suited to volume production techniques, and for large area sources the immersion lens becomes impracticably large. The use of individual immersion lenses has been described for infrared detector arrays [11]; however, whether hybridized or integrated, they can be difficult to align and are not efficient for emission (or detection) over large apertures.

QinetiQ, in Malvern, has investigated the use of optical concentrators to reduce the currents needed to drive negative luminescent devices, and have developed a novel fabrication technique [12] where the optical concentrators are integrated into the substrate on which the epitaxial material is grown, as illustrated in Figure 3b, with the active area situated at the focus of the cone. Light generated from the active region at large angles, which would normally suffer total internal reflection at the opposite, radiating, surface, is instead reflected by the surface of the cone. This reflected light then impinges at the radiating surface at angles smaller than the critical angle, and is therefore emitted from the semiconductor. This allows devices

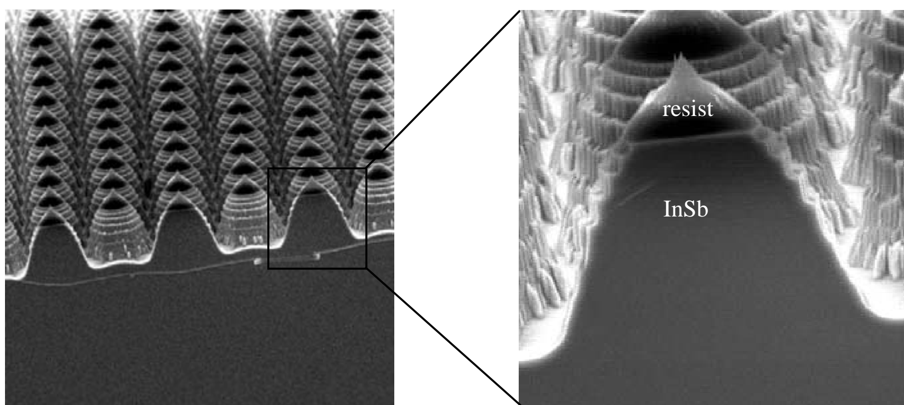


Figure 4. Integrated optical concentrators, fabricated in InSb and with near parabolic profile, on a pitch of 37 μm.

of smaller area to emit or absorb the same amount of light as a larger area device. The maximum possible area gain (in Figure 3) for a concentrator which collects all the radiation over a 2π field of view is n^2, where n is the refractive index of the material from which the concentrator is made [13]. For InSb the maximum possible area gain in this case is 16, and we have shown previously that when operated either in NL or PL mode a Winston cone [13], which has a parabolic profile, approaches the ideal limit [14]. The fabrication technique employed by Ashley *et al.* [12] consists of a dry etch process involving the repeated use of alternate chemistries that etch the semiconductor or the resist, and results in structures such as those shown in Figure 4. Nash *et al.* [15] have reported a reduction in the current needed to drive large area NL devices by a factor greater than 4 by the use of optical concentrators; see Section 4.2 of this chapter.

3 The Magnetoconcentration Effect

Of the two methods used to generate negative luminescence (NL) in semiconductors, the magnetoconcentration effect (MCE) was the first to be experimentally demonstrated, by Ivanov-Omskii *et al.* [2] in 1965. Since then, MCE has been used to generate NL in a range of material systems with groups from the Ukraine, Russia and Japan being active in the development of theory and experiment. Malyutenko's group, at the Ukrainian Academy of Sciences, has been particularly active in this field.

3.1 Principle of Operation

A charged particle moving in crossed electric and magnetic fields experiences a Lorentz force which is perpendicular both to the magnetic field and to the direction of motion of the particle. This effect can be exploited to generate negative luminescence by applying a magnetic field to a semiconductor crystal that has Ohmic contacts as illustrated in Figure 5a. Both fields are applied parallel to the emitting surface, so that electrons and holes drift towards one surface of the crystal. Regions of enhanced and depleted carrier concentrations are formed, and positive and negative luminescence can be observed from each region respectively, although in practise only one surface is usually free. Figure 5b shows schematically the variation of carrier concentration through a *2d* wide sample for both positive and negative luminescence at the left-hand-side surface. Reversing the direction of the electric and magnetic fields independently has the effect of switching the emission at one surface from positive to negative, or *vice-versa*. Reversing the direction of both fields simultaneously has no effect.

To achieve efficient NL, samples need to be near-intrinsic, have high and low surface recombination velocities at opposite faces of the sample, and a sample thickness longer than the diffusion length.

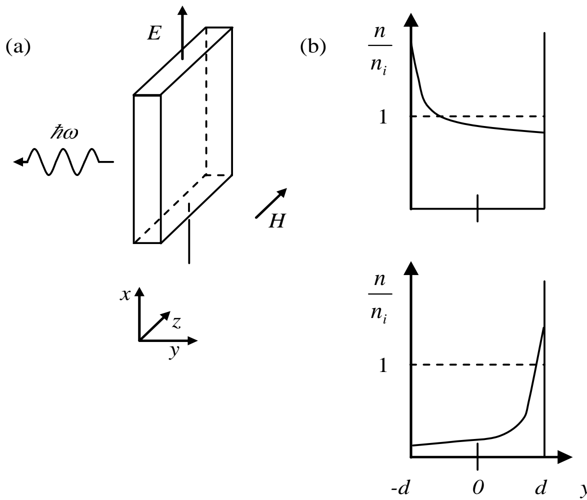


Figure 5. (**a**) MCE experimental setup (**b**) variation in carrier concentration across a semiconductor sample for positive luminescence (*top*) and negative luminescence (*bottom*)

3.2 Indium Antimonide and Cadmium Mercury Telluride Devices

The pioneering measurements of Ivanov-Omskii, Kolomiets, and Smirnov [2] were made on p-type InSb. They used magnetic fields up to 8 kOe, and microsecond electric field pulses of roughly 10 Vcm^{-1} and above. Short electrical pulses were used to ensure that Joule heating of the sample was minimised. One side of each sample was polished with fine emery powder to increase the surface recombination velocity, whereas the opposite face was subjected to electro-polishing to reduce the surface recombination velocity. Their observations showed that when Lorentz forces push carriers into the crystal, negative pulses were seen by the detector used to measure the emission from the sample. For the high current densities used in these experiments (above about 10^3 A/cm^2), the positive luminescence (PL) was more intense than the negative luminescence. In a subsequent study, they measured the spectra of the NL [3] and showed that band-to-band recombination was the primary operative mechanism, so that PL and NL have similar spectral characteristics (except that the PL can be shifted towards shorter wavelengths due to the Moss-Burstein effect under strong accumulation conditions).

Malyutenko and co-workers extended this work on InSb by concentrating on the NL aspects, measuring the field, temperature and spectral dependence of the negative luminescence power in both InSb [6-8,16] and CdHgTe [8,17,18] samples. Morimoto and Chiba also observed NL in InSb samples [19,20] in crossed electric and magnetic fields, although their aim was the fabrication of light emitting and laser diodes. In fact, much of the initial work on NL devices was motivated by the desire to produce efficient light emitters in the IR, where the small band-gap of the materials used limits the injection efficiency of standard diode structures due to the

smearing out of the barrier by thermally generated carriers. Auger processes also limit the performance of standard n-p photodiodes with small band-gaps.

A theoretical analysis of the field dependencies of both the positive and negative luminescence was performed by Bolgov *et al.* [7] and Berdahl [21], who showed that for small values of applied electric field, the radiated power emitted by a sample (the emittance) drops linearly with the electric field, and is also dependent on the surface recombination velocity, and carrier mobilities, concentrations and lifetimes. In this linear regime negative (or positive) luminescence generated using the MCE can be used to determine these parameters. For example, Berdahl and Schafer [22] extracted carrier lifetimes for InSb samples with known carrier concentration and mobility.

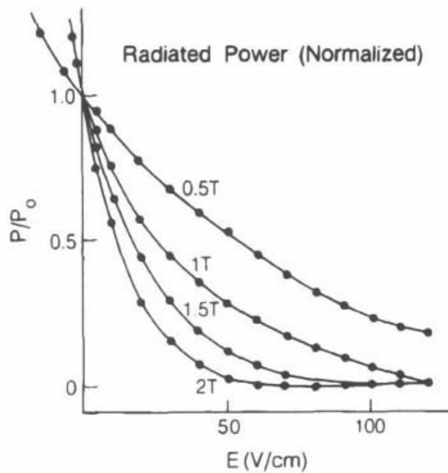


Figure 6. Positive ($P/P_0 > 1$) and negative ($P/P_0 < 1$) luminescence for a crystal of $Cd_{0.28}Hg_{0.72}Te$ at 300 K. P is the absolute radiated power per unit area, and P_0 is the value of P for $E = 0$. After [9].

For sufficiently high values of electric and magnetic fields, the NL was predicted to saturate [7,21] when its magnitude is equal to the equilibrium thermal emission from the sample, P_o, and this has been observed experimentally. Figure 6 shows the dependence of the normalised NL signal as a function of the applied electric field for a $Cd_{0.28}Hg_{0.72}Te$ sample at 300 K. Here P is the absolute radiated power per unit area (equal to the total integrated spectral emittance). Saturation of the NL signal is clearly seen at high electric field values. As P_o increases with temperature, the NL signal also increases with increasing temperature. A good review of the early MCE measurements on InSb and $Hg_xCd_{1-x}Te$ samples can be found in [9].

More recently, the MCE effect has been utilised to produce NL devices intended for application in dynamic infrared scene projection. Malyutenko and co-workers have reported promising results from an InSb based array capable of simulating a 40 K temperature drop (with the array operating at room temperature) [23,24], see Section 5.2. To achieve this temperature drop, however, a relatively large magnetic field (12 kOe) had to be applied, which increases the difficulty of making practical systems. As a common grid connection is also required for this implementation this

restricts the fill-factor that can be achieved, which also limits the apparent temperature range.

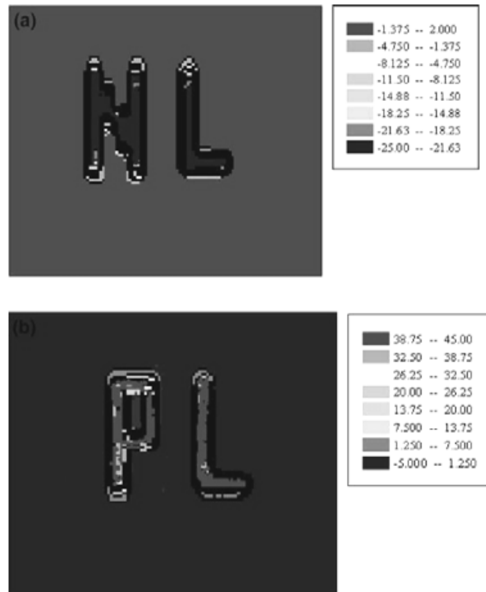


Figure 7. Measured apparent temperature maps, H = 12 kOe. (**a**) NL generation regime (E = 37 Vcm^{-1}); (**b**) PL generation regime (E = 55 Vcm^{-1}). After [25]

An alternative approach was demonstrated by Bolgov and Malyutenko [25] who employed a single, large area (15 mm × 15 mm) $Cd_xHg_{1-x}Te$ graded gap screen, which was driven in phase with the frame scanning of a single detector thermal imaging camera to generate a two-dimensional image. The free surface of the sample was polished by fine powder in order to increase the surface recombination velocity, whilst a small effective surface recombination velocity was achieved at the $Cd_xHg_{1-x}Te$ - substrate interface by the use of a graded gap transient layer [18]. This system relies on the fact that the MCE driven luminescence can be modulated at speeds, which are only limited by the lifetime of the non-equilibrium charge carrier, much faster than the camera frame rate. For example, in this case the lifetime was less than 10^{-7} s, leading to a maximum modulation rate of around 1 MHz. Figure 7 shows measured apparent temperature maps obtained with this system, for which the luminescence extended to ~7.5 μm at room temperature, for both negative and positive luminescence.

3.3 Other Material Systems

Some of the earliest MCE measurements were performed on Ge [5], due to the availability of high purity material, although little work has been reported since. As already mentioned, conventional p-n photodiodes suffer from a reduction in

efficiency as the temperature is increased, or the band-gap decreased, due to thermal smearing out of the barrier, and an increase in Auger processes. This motivated measurements of MCE induced PL and NL in InAs at elevated temperatures by Malyutenko *et al.* [26], who showed that larger changes in luminescence could be achieved using magnetic injection of carriers, compared to a conventional diode, at temperature greater than ~360 K.

3.4 Conclusions

For experimental studies, MCE is attractive as it is possible to investigate a uniform material, which therefore has a high fill-factor, with a single free surface, in contrast to reverse-biased diodes which require more complex structures. However, for many device applications the necessity to apply a relatively large magnetic field renders this approach impractical and much effort has been focussed over the last few years on the production of NL using photodiodes, which will be discussed in the next Section .

4 Photodiodes

The second commonly used method to produce negative luminescence (NL) in semiconductors exploits the reduction in carrier densities that occurs in the vicinity of a reverse biased pn junction, even in the absence of a magnetic field. The occurrence of this "carrier extraction" effect was first predicted by Berdahl in 1985 [27], and was then demonstrated shortly afterwards in CdHgTe and InSb photodiodes [28,29]. Since then there has been continuing progress towards larger, more efficient and longer wavelength devices, with groups from QinetiQ, Malvern (formerly RSRE/DRA/DERA), the Naval Research Laboratory (NRL), Washington, and the Ioffe Institute, Russia, together with their various collaborators, being particularly active in this area.

4.1 Principle of Operation

In these diode structures the negative luminescence occurs because of the reduced carrier densities which occur in the vicinity of a reverse-biased diode due to carrier extraction. In order to obtain large reductions in the carrier densities below their equilibrium values, the devices are based on a near intrinsic, narrow-gap epitaxial layer. Nearly all devices have been based on heterostuctures, designed specifically to maximise the reduction in electron and hole densities throughout an active region of thickness of the order of the optical absorption length [29]. The active layer is contained between two wide-gap layers or between one wider-gap layer and one very heavily doped layer. Examples are $\underline{p}\pi\underline{n}$, $\underline{p}\pi n^+$ or $\underline{p}\nu n^+$, where the underlined symbol means wide-gap, the plus symbol indicates high doping in excess of 10^{17} cm^{-3}, π is near-intrinsic p, and ν is near-intrinsic n. The structure contains one pn junction, which is operated in reverse bias, producing extraction of minority carriers. The other, isotype, junction is excluding to minority carriers, preventing their injection into the π or ν layer. The operation of such a device is illustrated schematically in Figure 8 for the case of a three-layer $\underline{p}\pi\underline{n}$ device. At zero bias, the

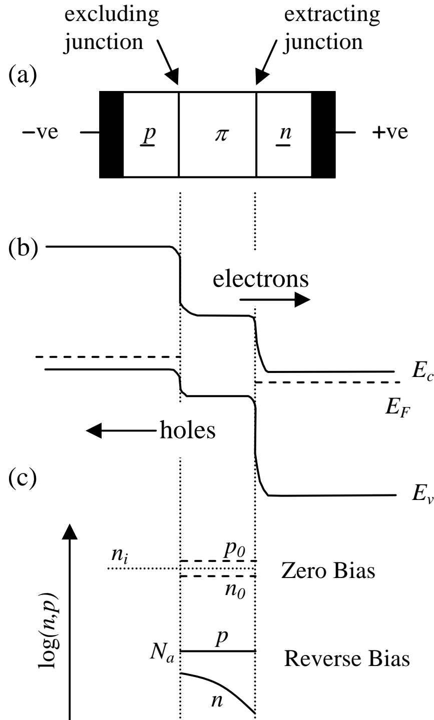


Figure 8. Schematic of an extracting diode: (**a**) the multilayer structure; (**b**) the band edges under reverse bias, where the labels E_c and E_v refer to the conduction and valence band edges respectively, and E_f is the Fermi energy; (**c**) the carrier densities in the π-region

electron and hole densities in the π region, n_0 and p_0 are close to the intrinsic value n_i, which is typically 10^{16}–10^{17} cm^{-3} (as illustrated in Figure 8c). The π region is then in equilibrium with its surroundings, emitting and absorbing photons at equal rates at wavelengths corresponding to photon energies above its band-gap. Note that the wider band-gap p and n regions are transparent to photons with energy close to or just above the band-gap of the π region, so that both can act as windows. Under reverse bias, the band edges are illustrated in Figure 8b. Electrons in the π region which diffuse to the πn junction are swept by the field into the n region. Minority carrier extraction therefore occurs at this diode junction. As the p region contains very few minority electrons it cannot inject electrons into the π region to replace those lost to the πn junction, and the pπ junction is therefore said to be excluding. For the hole current, the situation is reversed - the hole current flowing into the π region from the n region is negligible due to the very low density of minority holes in the n region. The overall result is that the electron density falls in the π region by several orders of magnitude. In order to maintain space charge neutrality the majority hole density also falls, reaching a value which is determined by the net acceptor level, N_a, and is typically 10^{15} cm^{-3}. As observed through the n region, say, the π region exhibits ‘negative luminescence’: although it continues to absorb, it does not emit radiation. It must be remembered that the p region is also transparent,

and the metal contact or substrate behind this region continues to emit radiation. For effective negative luminescence, therefore, the thickness of the π region must be comparable with or greater than an optical absorption length, typically in the range 3–5 μm.

The reduction of the carrier density in the active layer under reverse bias also leads to suppression of the Auger carrier generation processes, in particular Auger-1 and Auger-7, and their associated noise. This leads to improved performance of uncooled IR detectors based on this type of structure [30], which was pioneered by Ashley and Elliott in Malvern, and the minimisation of the current necessary to drive negative luminescent devices. This is particularly true of devices designed to operate on the 8-12 μm region. Under forward bias, the extracting and excluding contacts become injecting and accumulating, respectively, and serve to confine the electrons and holes in the active layer, with little leakage into the contact regions. This maximises the radiative recombination rate and gives efficient emission.

4.2 Indium Antimonide Devices

Initial InSb devices, which were pioneered by the group at Malvern, were relatively small, <1 mm^2 [29], whereas devices with areas of ~1 cm^2 are now being demonstrated [15,30]. Such large area devices, which are the largest negative luminescent devices so far reported, are required for use in a number of applications. The luminescent spectra of a typical InSb device, at room temperature, under forward and reverse bias are shown in Figure 9. As with negative luminescence produced by the MCE effect, the spectral properties of the NL and PL are almost the same, and in this case under both directions of bias the spectrum peaks at approximately 6.5 μm. The InSb/InAlSb heterojunction photodiodes from which these spectra were obtained [15] were grown by MBE onto degenerately doped (N_d-N_a = 2–3×10^{18} cm^{-3}) substrates in order to produce a Moss-Burstein shift in the band-gap to increase the IR transmission and allow radiation from the diodes to be emitted through the substrate [12]. The diodes had a $p^+\underline{p}\pi n^+$ structure similar to that described above, with a 3μm undoped active π region, and with the n^+ and p^+ regions doped to 3×10^{18} cm^{-3} and 1×10^{18} cm^{-3} using silicon and beryllium respectively. The $\underline{p}$ region is wider band-gap $In_{1-x}Al_xSb$, of composition $x = 0.15$. This material system is not lattice matched and the thickness of the wide-gap region is limited to values of order tens of nanometres to avoid strain relaxation.

Devices, consisting of a large array (1 cm^2) of photodiodes, were fabricated using the micromachining technique, described previously, so that each photodiode had an integrated optical concentrator. After fabrication of the optical concentrators, the devices were etched, passivated and contacted with chromium/gold. Devices were then indium bump-bonded, epitaxial side-down, to a silicon lead-out. The lead-out consisted of two gold pads only, so that the whole array of photodiodes was contacted simultaneously. Figure 10 shows infrared images of the device, taken with a MW infrared camera, operating (from left to right) under reverse, zero and forward bias. Note that under reverse and forward biases the active area of the device appears cold and hot, respectively, to the camera, even though the device is at room temperature. The large active area of the device was slightly degraded during the thinning of the substrate.

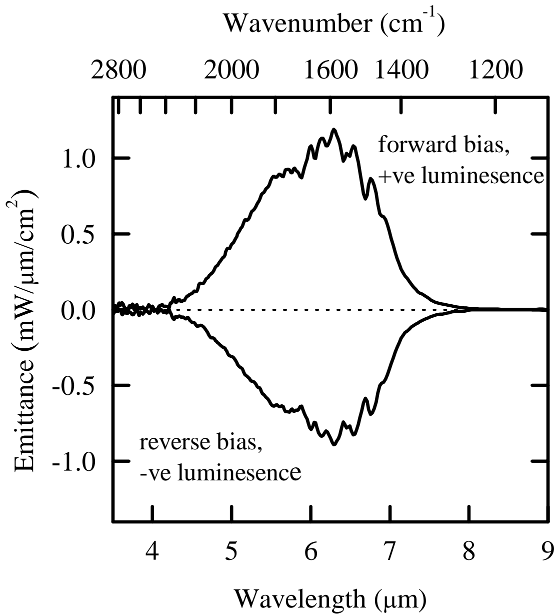


Figure 9. Measured emittance of a ~1 cm^2 InSb device under forward and reverse biases. In both cases the device was driven at 1A (~1.16 Acm^{-2}).

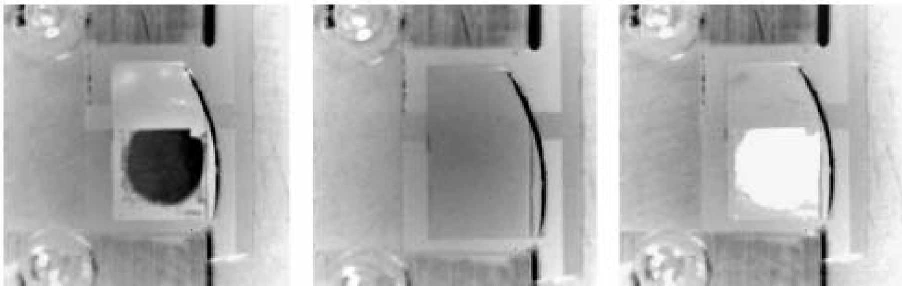


Figure 10. Infrared image of an InSb 1cm^2 negative luminescent device operating at room temperature, under (*from left to right*) reverse, zero and forward biases. The device was driven at 2A (~2.2 Acm^{-2}) in both forward and reverse biases. After [15].

The integrated spectral luminescent power from the device was also measured as a function of drive current, again at room temperature, and this is shown in Figure 11. Note that the signal of the device under reverse bias saturates at ~2.3 Acm^{-2}, when all the minority carriers have been removed from the active region and the extra current is due to breakdown effects. In contrast, no saturation of the signal is observed under forward bias. The value of the total integrated emittance measured at reverse bias saturation, 1.5 mWcm^{-2} ± 20%, is consistent with values previously

reported for similar devices fabricated without optical concentrators [29], although direct comparison is difficult as emission from the device described here was through the substrate and no anti-reflection had been applied. Negative luminescence saturation was reached, however, at much lower currents in the device with concentrators (~2 Acm^{-2} as compared to ~9 Acm^{-2}), suggesting that optical concentrators can be used to significantly reduce the currents needed to drive the device. The negative luminescent efficiency (NLE) of the device is plotted as a function of wavelength in Figure 12. The upper bound of the NLE, which is given by 1-*R* (assuming that the refractive index is independent of wavenumber), where *R* is the reflectivity, is also shown in Figure 12. The fact that the NLE approaches this maximum value suggests that the concentrator fabrication process did not adversely affect the inherent diode performance. Improvements in the electrical and optical efficiencies of the concentrators, further thinning of the substrate and the application of an anti-reflection coating should increase the NLE and hence the apparent temperature difference of this type of device.

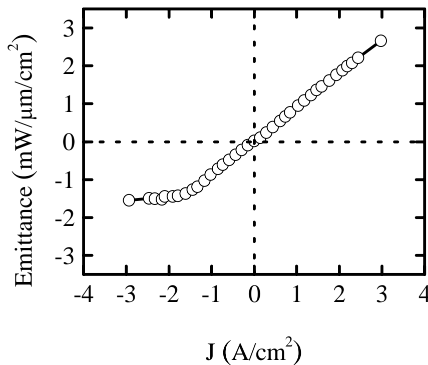


Figure 11. Measured negative luminescent emittance *vs* drive current for an InSb 1cm² device

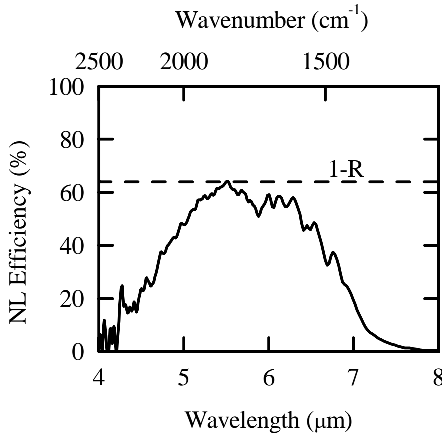


Figure 12. **N**egative luminescent efficiency as function of wavelength for an InSb 1 cm² device

4.3 Cadmium Mercury Telluride Devices

In recent years there has been considerable progress in the performance of CdHgTe negative luminescent devices. Relatively large area devices with high negative luminescent efficiencies and low reverse bias saturation currents are now being demonstrated. Coupled with increasing cut-off wavelengths this leads to devices that are suitable for applications such as the cold-shielding of uncooled IR detector arrays. The progress in device performance has been due to improvements in the quality of the material, and to the implementation of novel junction design. For example, the Naval Research Laboratory measurements of CdHgTe photodiodes (λ_{co} = 4.2 μm) from Sanders IR Imaging [10] confirmed internal NL efficiencies of up to 82%, while operating with a reverse-bias saturation current density (J_{min}) of 5 Acm^{-2}. Measurements by the same group on 800 μm × 800 μm CdHgTe on Si photodiodes (λ_{co} = 5.3 μm), grown by Raytheon Vision Systems, yielded an internal NL efficiency of 88% over the entire 3-5 μm spectral region [31]. That was accompanied by a further decrease in J_{min} to 1.3 Acm^{-2}, despite the growth on a lattice-mismatched silicon substrate.

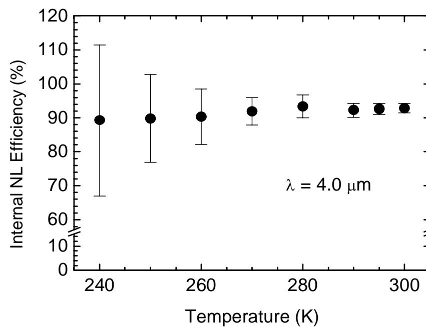
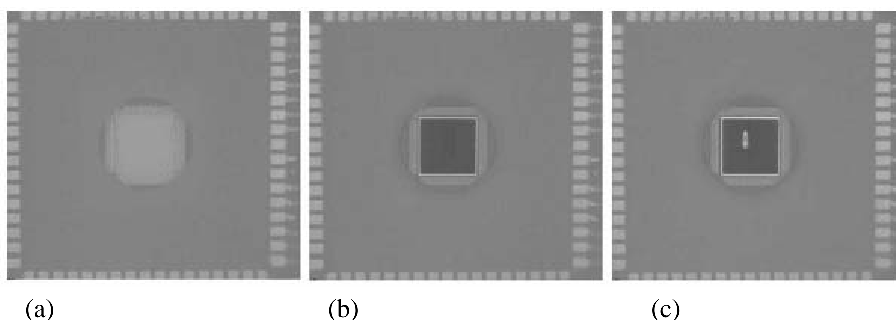


Figure 13. Internal NL efficiency *vs* heat sink temperature for a mid-wavelength CdHgTe NL device. After [32].

The NRL group have also characterized CdHgTe-on-CdTe devices with a Rockwell Scientific proprietary junction geometry [32-34]. For diodes with a cut-off wavelength of 4.6 μm the observed internal NL efficiency of 93% corresponded to suppressing the blackbody emission by a factor of 14, and the reverse-bias saturation current density was only 0.13 Acm^{-2}. The internal NL efficiency was essentially independent of temperature over the range 240-300 K, as shown in Figure 13, whereas the apparent temperature drop of the device, ΔT, reached a maximum of 54 K at room temperature.

Emission from a 5 mm × 5 mm device with a room temperature cut-off wavelength of 4.8 μm [34], was observed directly using a radiometric, InSb IR camera with a 2.6-5 μm bandpass cold filter. In this case the active area was segmented into a number of "fingers", which for future device areas scaled up to several square centimetres will ensure that a single defective diode does not render the entire device inoperable. A similar approach was also used by Nash *et al.* [35]

for devices with a longer cut-off wavelength. Figure 14 reproduces images of the device at 296 K under conditions of no bias, a saturated reverse bias (all 36 fingers), and a reverse bias of only 35 fingers. When the reverse bias is applied the device clearly appears cooler to the IR camera. Although the apparent temperature change is only ~20 K, this was not representative of the device performance as no antireflection coating had been applied and emission in the ~4.8-5 μm range was not suppressed because the device cut-off is only 4.8 μm rather than ~5 μm for an optimized mid-IR NL device. The camera and applied bias to the device were also not synchronized, so that the temperature was averaged over the frame and represents a minimum temperature change. The same team have recently extended this work by producing large area (5mm x 5mm) devices with a cut-off wavelength beyond 5μm [36].



(a) (b) (c)

Figure 14. Thermal images of (**a**) unbiased, (**b**) reverse-biased, and (**c**) reverse-biased except for one finger, 5 mm × 5 mm NL CdHgTe device at 296 K as obtained using a radiometric 2.6–5 μm camera. The bias was modulated at 3 Hz, 10% duty cycle. The delineation of the fingers is faintly visible in the biased pictures when they are magnified. After [34].

For applications such as non-uniformity correction and cold-shielding of focal plane arrays, which will be described in the next Section , negative luminescent devices with longer cut-off wavelengths are also required. For example, when used as a radiation shield, these devices could also significantly improve the performance of minimally cooled CdHgTe detectors, and of microbolometer arrays once these become background limited. The Malvern group in particular has been at the forefront of the development of long wavelength CdHgTe NL diodes, and has produced diodes with a cut-off wavelength of ~10 μm, the longest so far reported [37].

This device, which was designed at QinetiQ and grown by BAE SYSTEMS, Southampton, (on a semi-insulating GaAs substrate), comprised a 1mm diameter circular array of 52 square photodiodes on a 120 μm pitch. Electrical contact to the diodes was achieved by indium bump bonding the array, epitaxy side down, onto a gold lead-out pattern on a sapphire carrier. Figure 15 shows a typical reverse bias luminescence spectrum together with the spectral response of the whole array, $R(\lambda,T)$. The photodetector quantum efficiency $\eta_{pd}(\lambda,T)$, defined as the percentage of photons detected, has a peak of ~72% at $\lambda = 3.3$ μm, and approaches the maximum value that can be achieved without an anti-reflection coating (given by 1-

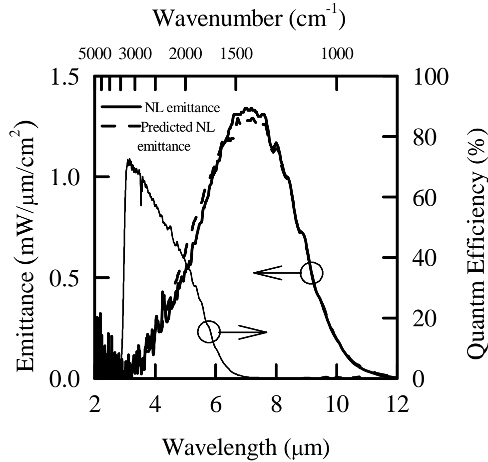


Figure 15. Measured emittance $M_{NL}(\lambda,T)$ and spectral response $R(\lambda,T)$ of a long-wavelength CdHgTe device under reverse bias

R, where *R* is the reflectivity of the uncoated surface). The addition of an anti-reflection coating would therefore lead to a maximum efficiency in excess of 90%. The mechanisms of NL and photodetection are related phenomena, so the expected NL spectral emittance $M_{NL}(\lambda,T)$ can be predicted from $R(\lambda,T)$ using

$$M_{NL}(\lambda,T) = R(\lambda,T)M_{bb}(\lambda,T) \tag{1}$$

where $M_{bb}(\lambda,T)$ is the emittance of a blackbody at the same temperature. The dotted line in Figure 15 shows $M_{NL}(\lambda,T)$ calculated this way, and the total integrated emittance in this case was 4.4 mWcm^{-2} ±15%. Note that, to enable the shape of the two spectra to be compared, the maximum in measured emittance has been scaled to that of the predicted curve. The close agreement in the shapes of the spectra confirms that the processes of NL and photodetection are closely correlated., and that the negative luminescence efficiency is, therefore, expected to be of comparable magnitude to $\eta_{pd}(\lambda,T)$. To within experimental uncertainties, this was the case.

A feature typical of LW luminescent diodes is that in reverse bias there is a dramatic increase, at a certain bias voltage, in their efficiency. This is due to suppression of the non-radiative Auger processes, which leads to pronounced Negative Differential Resistance (NDR) because of the sudden drop in the current required to drive the diodes (J_{min}). Figure 16 shows a plot of the measured emittance, together with the measured current density, as a function of applied reverse bias for an element of the same device. The reverse bias current-voltage characteristic displays negative differential conductance arising from suppression of the Auger carrier generation mechanism as a consequence of exclusion and extraction, where the current density rises to a maximum value of 86 Acm$^{-2,}$ and then falls to a minimum of ~4.5 Acm^{-2}. The measured emittance increases with reverse bias as the

maximum current is approached, and then saturates once the carrier density is reduced to near the net doping level by the effects of extraction and exclusion. In Figure 17 the internal radiative efficiency η_i is plotted as a function of reverse bias, for the same element. The extremely high value of the efficiency obtained at saturation reflects both the low trap density in this material ($N_t \sim 10^{11}cm^{-3}$, c.f. $\sim10^{13}cm^{-3}$ in vacancy doped CMT), which leads to a low Shockley-Read generation rate, and, more importantly, the suppression of Auger carrier generation as a consequence of carrier exclusion and extraction. The Auger suppression is demonstrated by the large increase in η_i (~100x) seen at ~0.7 V, which is much more pronounced than in MW devices, where Auger processes are not so strong.

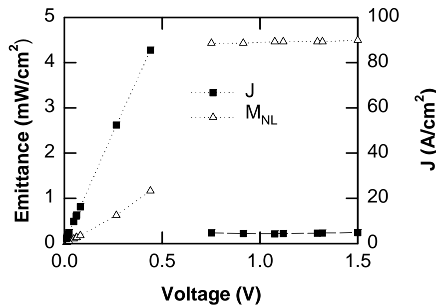


Figure 16. Measured total emittance $M_{NL}(T)$ and current density J as a function of reverse bias for a long-wavelength CdHgTe NL device [14]. The abrupt changes in both $M_{NL}(T)$ and J, indicated by the *lighter lines* (drawn as guides to the eye only), are due to negative differential conductance effects. After [37].

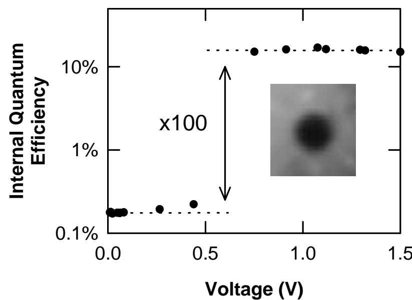


Figure 17. Internal quantum efficiency η_i as a function of reverse bias for a long-wavelength CdHgTe NL device. The *inset* shows an infrared image of emission from the device under reverse bias. After [37].

Subsequently, a CdHgTe on-GaAs photodiode employing the QinetiQ "JESS" geometry [38] for the junction, which was characterised by NRL, displayed a NL efficiency of 85% and J_{min} = 0.7 Acm^{-2}, despite a cut-off wavelength beyond 8 μm. In this device the minority carriers were fully extracted, with the consequent reduction in majority carrier density, so that the intrinsic Auger processes were

therefore fully suppressed. Very recently, QinetiQ has reported the fabrication of NL devices from CdHgTe grown on Si [39] with peak emission at 7.2 μm, but with a very low J_{min} of 0.86 Acm^{-2}. This is the first time that a long-wavelength NL device has been grown on Si. Finally, optical concentrators fabricated using the micromachining technique described previously, could be used to reduce the currents needed to drive the device even further [40].

4.4 Other Device Structures and Material Systems

Type-II superlattices have been explored as an alternative to CdHgTe as a way of producing NL devices in the technological important region between 8 μm and 12 μm. Pullin *et al.* [41] characterised an InAs/InAsSb strained-layer-superlattice (SLS) [42], for which a schematic band structure diagram is shown in Figure 18. Although designed as a detector, negative and positive luminescence were observed peaking around 4.2 μm, as shown in Figure 19. Unfortunately, the NL efficiency was relatively low (also shown in Figure 19), which was thought to be due either to Joule heating, or to reverse bias breakdown.

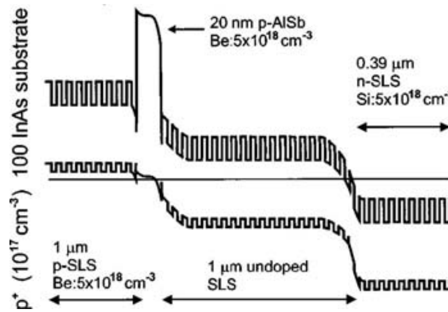


Figure 18. Schematic diagram of superlattice LED with AlSb barrier layer. After [42].

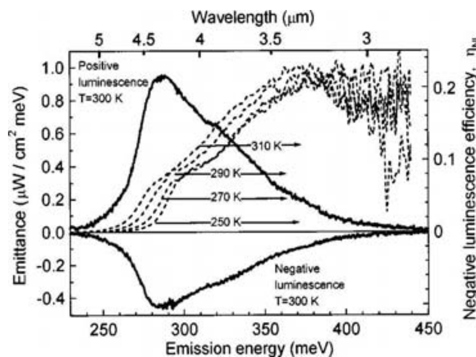


Figure 19. PL and NL for an InAs/InAsSb superlattice LED. *Dotted lines* are NLE curves in the range 250–310 K. After [41].

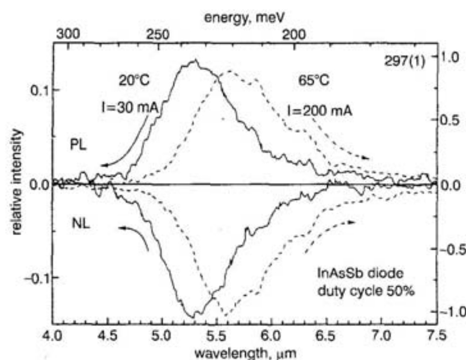


Figure 20. Negative and positive emission spectra from a InAsSb diode at 20 ºC (*solid line*) and 65 ºC (*dotted line*). After [46].

A Type-II InAs/GaSb superlattice was also investigated by Olafsen *et al.* [43], which had a similar spectral characteristic and NL efficiency. In this case simulations indicated that a thicker active p layer would improve the NL efficiency significantly.

Matveev [44-47] and co-workers at the Ioffe institute have investigated InAsSb NL devices with the aim of producing devices for gas detection in the 3-5 μm region. In particular, the motivation was to produce a device that could operate at room temperature and above, where the NL efficiency becomes higher than the PL efficiency due to suppression of Auger processes. NL devices are also free of current-crowding effects [48]. For a cut-off wavelength of ~6 μm, the NL efficiency was as high as 0.9, which is comparable to typical CdHgTe devices, at room temperature and 0.44 at 64 ºC. However, J_{min} was 20 Acm^{-2}. Figure 20 shows the NL and PL spectra measured at room temperature and 64 ºC. The same team have also integrated the devices with bandpass filters, attached with optical glue, to narrow the emission spectra and make them more suitable for gas detection [49]. Finally, InAs NL photodiodes have exhibited NL at 360 K [50], and these could be useful in situations were high temperature operation is required.

4.5 Conclusions

For comparison, Figure 21 plots calculated values of the intrinsic Auger (solid line), following the model described by Beattie and White [51], and internal radiative (dashed line) currents as a function of wavelength, at room temperature, together with the average J_{min} obtained from recent results in the literature. Note that in cases where no cut-off wavelength was given, values of these were estimated from the values of peak NL emission wavelengths by assuming that the cut-off occurs at an energy k_BT less than that of the peak. As can be seen, photodiodes with large Auger suppression are now being produced, particularly in CdHgTe. This, coupled with the high values of NL efficiency and large areas reported, suggests that NL photodiodes are now close to having the performance needed for a range of applications.

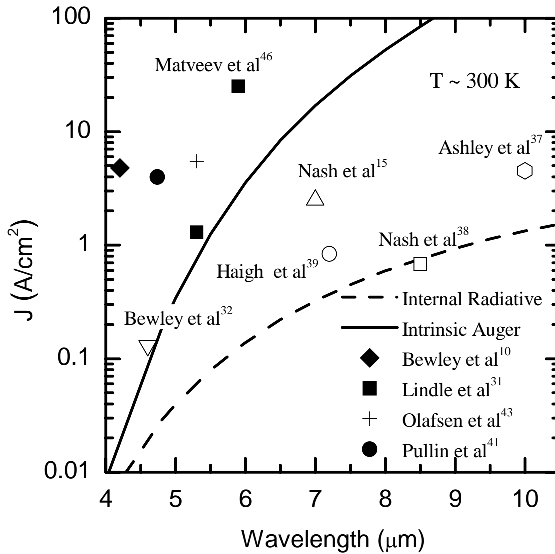


Figure 21. Calculated Auger and internal radiative currents as a function of wavelength, compared with recent published device results

5 Applications of Negative Luminescence

Several applications of negative luminescence have been proposed [52,53]. These range from the simple use as an infrared source for gas sensing, through techniques to improve the performance and testing of thermal imagers, to those which derive value from the ability to alternate the emission between positive and negative modes, so that on average the output is identical to the background level. The progress in fabricating devices with a reasonably large area and high efficiency means that the implementation of some of these is now commencing.

5.1 IR Sources for Gas Sensing

The use of IR LEDs to provide a source of infrared radiation that can be rapidly modulated for use in applications such as non-dispersive IR absorption gas sensing is described later in this book by Crowder and Smith. Many gases have their fundamental absorption in the infrared range, and so can be sensed quantifiably by the extent to which they attenuate an IR signal. Temperature stability and power consumption are key issues for some gas sensing requirements, and negative luminescent devices offer benefits for both of these in certain situations.

For use in gas sensing, all that is required is a source of modulated radiation. In conventional IR gas sensors the gas absorbs the directional radiation from the source

and re-emits it isotropically, so that less falls on the detector. When use is made of a negative luminescent source, then the isotropically emitting gas intercepts the negative luminescent radiation, so increasing the signal on the detector. Thus the negative luminescent source can be used in exactly the same way as a positive one.

5.1.1 Temperature Stabilisation

The output from an IR LED is highly temperature sensitive owing to the thermal activation of the internal, non-radiative carrier recombination processes which compete with photon generation. The emitted power, at a constant current, decreases with increasing temperature. In a gas sensor, this can be compensated for to some extent by use of a reference channel, at a wavelength not absorbed by the gas, or measurement of the temperature of the LED to enable use of algorithms in the signal processing based on knowledge of the physics of the recombination processes. This compensation, however, leads to extra complexity and cost, and has limits to its accuracy. The temperature dependence of the negative luminescent output power at constant current, conversely, increases with temperature under some conditions owing to the balance between internal recombination processes. This can be used to offset the decrease in positive output, so that the overall modulation given by the difference between the two states has less variation with temperature [54], as illustrated in Figure 22 for an InSb device.

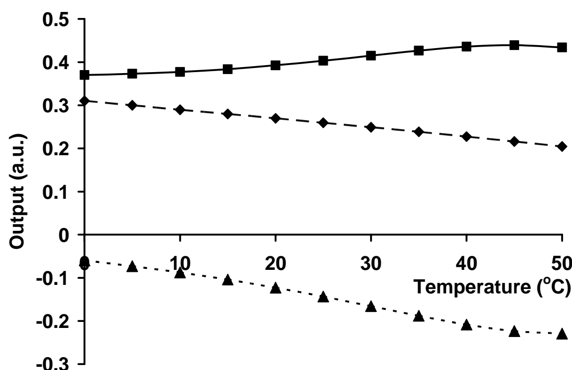


Figure 22. Output power *vs* temperature for an InSb diode operating under positive luminescent mode (*filled diamond, long dashed*); negative luminescent mode (*filled triangle, short dashed*); and positive minus negative mode (*filled square, solid*)

5.1.2 Efficient Long Wavelength IR Sources

Many gases have their fundamental absorption at around 4 μm, where the LEDs are reasonably efficient and have approximately the same efficiency under positive and negative modes of operation. Several important gases, however, have principal absorption spectra at longer wavelengths in the range up to 12 μm. As the energy-

gap of the semiconductor is reduced, in order to provide luminescence at longer wavelengths, the balance between recombination processes is altered and, in particular, the Auger process becomes increasingly dominant.

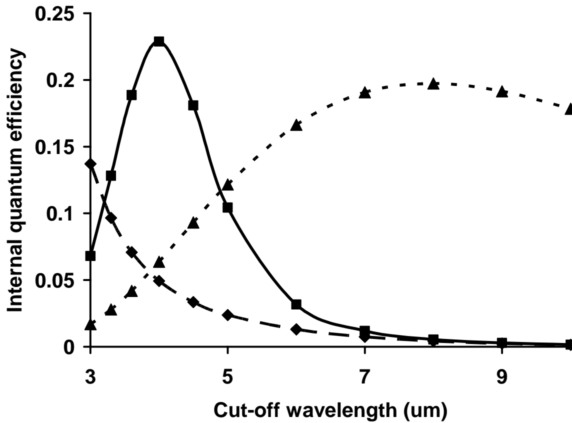


Figure 23. Internal quantum efficiency *vs* cut-off wavelength for an HgCdTe diode in equilibrium (*filled square, solid*); operating under positive luminescent mode at a fixed current density of 10 $Acm^{-2}\mu m^{-1}$ (*filled diamond, long dashed*); and negative luminescent mode at a fixed bias of -0.5 V (*filled triangle, short dashed*)

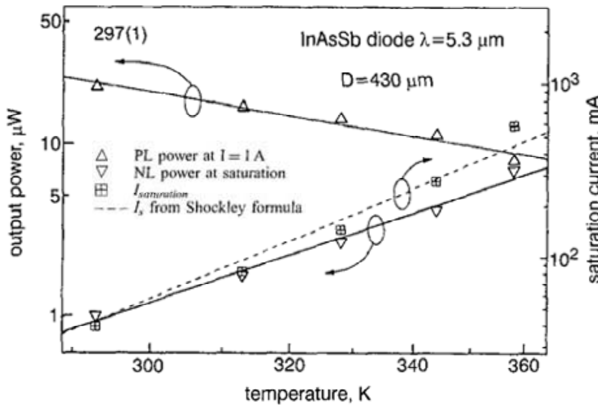


Figure 24. Temperature dependence of NL and PL power at saturation and I=1A respectively for an InAsSb diode with emission wavelength of 5.3 μm. After [46].

This is illustrated in Figure 23, where the predicted internal quantum efficiency at equilibrium for cadmium mercury telluride diodes at room temperature of varying composition is plotted *vs* cut-off wavelength, and is seen to decrease rapidly with increasing cut-off wavelength beyond 4 μm. Under forward bias, the injected electron and hole concentrations are increased, leading to a further relative increase in the Auger recombination rate and decrease in the positive luminescence

efficiency. Under reverse bias, minority carrier exclusion and extraction reduce the electron and hole concentrations, thus largely quenching the Auger processes and leading to a dramatic increase in the quantum efficiency, as shown in Figure 23. Consequently, at wavelengths in excess of about 5 μm (depending on the intensity required), the negative luminescent mode is more efficient than the positive mode. Experimental results from a LWIR CdHgTe device demonstrating this effect are presented in Figure 17 in Section 4.2 of this chapter.

Matveev *et al.* [46,47] have also considered this optimisation in the context to elevated temperature operation of In(Ga)As(Sb) diodes and shown experimentally, for example, that the maximum efficiency operation of a 5.3 μm emission wavelength device crossed over to negative luminescent mode above approximately 90°C, as illustrated in Figure 24.

5.2 Dynamic Infrared Scene Projection

To assure proper operation of IR sensor systems, and to achieve their full performance, test and calibration are required. Dynamic infrared scene projectors are a vital component of this calibration and test process, driven mainly by the need to develop hardware-in-the-loop (HWIL) simulation for the next generation of infrared imaging missiles. A dynamic infrared scene projector based on negative luminescent devices has many potential advantages compared with existing systems [52,55]. Unlike liquid crystal or micro-resistor array based systems, NL devices can appear both “hot” and “cold”. As they are solid state based, NL devices also have very fast response times, individual NL devices can be driven at frequencies greater than 1 MHz. In addition, very high array fill factors can be achieved (>90%). Finally, as these devices operate at room temperature, there is no need for cryogenic cooling.

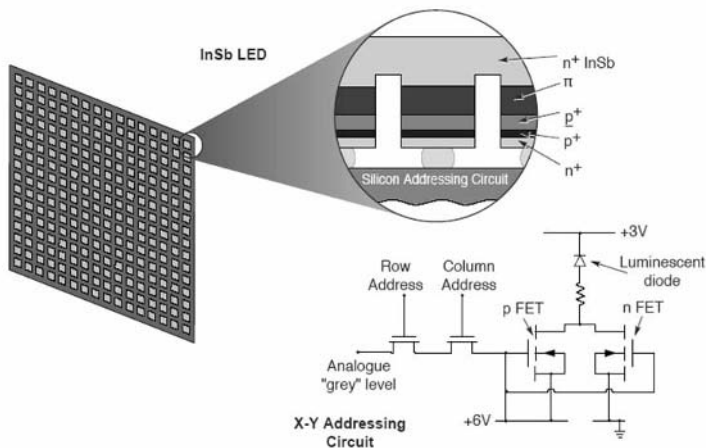


Figure 25. Schematic illustration of a dynamic IR scene projector based on positive and negative luminescent diodes. After [52].

A schematic representation of such a system is illustrated in Figure 25. To be effective in this application NL devices require a wide dynamic apparent temperature range, the appropriate spectral properties, and need to be driven in such a way as to ensure the necessary temperature resolution and response times. In addition, the simulator needs to contain a comparable number of pixels to the detector.

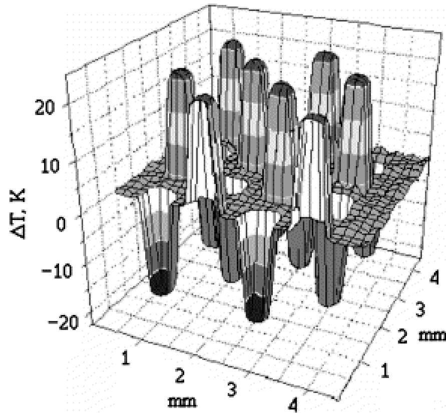


Figure 26. Pattern of temperature field, ΔT, for part of an array of InSb elements. After [24].

An example of a 2d array of sources which could be used for modulation transfer function assessment, though not for general scene projection as the pattern is fixed, has been demonstrated by Malyutenko *et al.* [24]. In this structure, the luminescence was produced using the magneto-concentration effect and alternate positive and negative contrast was achieved owing to the opposite direction of current flow in adjacent pixels. Figure 26 shows the temperature field pattern at room temperature for part of a 16 × 16 array of near intrinsic InSb elements, each 0.5 × 0.5 mm on a pitch of 1 mm, with an in-plane magnetic field of 1.2 T and electric field of 7 Vcm^{-1}. A ΔT of approximately ±20 K was achieved at this field, which, it is estimated, equates to an input power of 30 Wcm^{-2} of optically active area. Nash *et al.* [56] compare this with the power dissipation in an InSb photodiode of approximately 4 Wcm^{-2} to achieve a similar ΔT. A maximum apparent ΔT, assessed using a 3-5 μm camera, of +90 K and -40 K was achieved in an anti-reflection coated device at an electric field of 100 Vcm^{-1} for the positive and negative modes respectively.

5.3 Radiative Cooling

Berdahl proposed the use of negative luminescence to produce radiant coolers, and considered such devices as heat pumps with photons as the working fluid [27], as shown schematically in Figure 27. The reverse biased photodiode exhibits negative luminescence, *i.e.* within its absorption band, it absorbs more energy from the external emitter/absorber than it re-emits, so cooling it. Berdahl predicted that in the case of unity quantum efficiency for radiative recombination, the efficiency for cooling can approach the Carnot limit imposed by the second law of

thermodynamics. One should note, however, that electrical inefficiency in the photodiode will, under practical circumstances, lead to a rise in the temperature of the heat sink so increasing the emission of photons outside the absorption range of the photodiode which will negate the cooling effect. Narrow-gap semiconductors have a rather low ratio of radiative to non-radiative transition rates, so whilst they offer a large cooling power potential for objects around room temperature, they will tend to have a low efficiency. Ashley *et al.* [52] estimated that in practical situations cooling of the order tens of degrees would be the maximum possible, however this could be improved if materials with a high ratio of radiative to non-radiative transition rates became available. Berdahl demonstrated a cooling of 0.03 K in a blackened foil placed in proximity to an InSb MCE device [57]. Liptuga *et al.* [58] have demonstrated radiative cooling from InSb and CdHgTe and with a maximum decrease in temperature approaching 1°C.

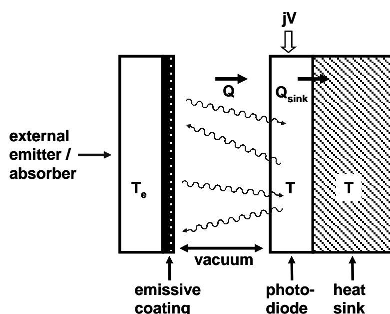


Figure 27. Physical configuration of a photodiode which exchanges radiation with an external emitter/absorber. After [27].

5.4 Uncooled IR Radiation Shields for BLIP IR Detectors

High performance infrared detectors are limited in their performance by the fluctuation in arrival rate of photons from the scene (so called BLIP), so the solid angle over which photons can fall on the detector is constrained to that of the optical system field of view by a 'cold shield'. The conventional technique is to employ a 'top-hat' metal shield which is physically cooled by the detector cryogenic encapsulation. This becomes inefficient for large arrays, because the aperture has to be opened up too large for the central elements in order to avoid vignetting the edge elements. This can be alleviated by increasing the height of the shield, but leads to problems of a very large encapsulation; excess heat capacity; insufficient cooling of the 'crown' of the hat; and excessive cool-down time. An external negative luminescent device could be used, probably in conjuction with a short conventional shield, to provide the radiation shielding [52,59], as illustrated in Figure 28. As the requirement is simply to reduce the total number of photons falling on the detector array, the uniformity of the NL radiation shield would not be critical and it could be constructed from a number of separate devices if necessary. Such structures could even, in principle, be applied to uncooled thermal detectors if their sensitivity becomes sufficiently good to benefit from radiation shielding.

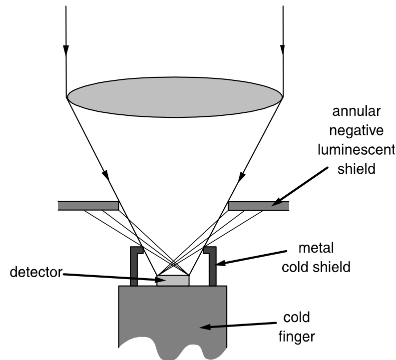


Figure 28. Schematic illustration of the use of a negative luminescent device to improve the cold-shielding efficiency of an IR detector array. After [52].

5.5 Radiometric Reference Planes for Thermal Imagers

Thermal imagers are required to resolve very small temperature differences in scenes with a mean temperature of, typically, 300 K. Current generation scanned imagers can resolve between 40 mK and 100 mK, whilst staring systems can approach 10 mK. Future generations of equipment will reduce this to a few milli-Kelvin by the use of on-focal plane background subtraction techniques. This sensitivity imposes severe demands on the uniformity of detectors within an array, and, even then, necessitates non-uniformity correction to remove fixed pattern noise. The correction is performed conventionally using Peltier devices. These have a slow response time and therefore can not follow rapid changes in mean flux level and can require several minutes to perform a correction sequence. Positive and negative luminescent sources can provide any arbitrary flux, within the range of scene temperature normally encountered, very rapidly and so overcome many of the deficiencies of the Peltier devices.

5.5.1 DC Restoration in Scanned Thermal Imagers

Scanned imagers are normally a.c. coupled, as illustrated in Figure 29, which removes offsets and enhances contrast. This introduces other image defects - droop will occur in a step function signal from the detector, because of discharge of the coupling capacitor, and undershoot will follow as the step function returns to zero; see Figure 30a. An additional problem associated with a.c. coupling is that the removal of the d.c. component removes the signal relating to the absolute temperature of the scene. This is particularly undesirable when performing radiometric monitoring or, for example, when imaging scenes with a prominent horizon. The solution adopted to restore the d.c. level and also to try to counter the effects of droop, undershoot and channel non-uniformity is illustrated in Figure 29. At the end of each line-scan the detector views a uniform reference scene and the

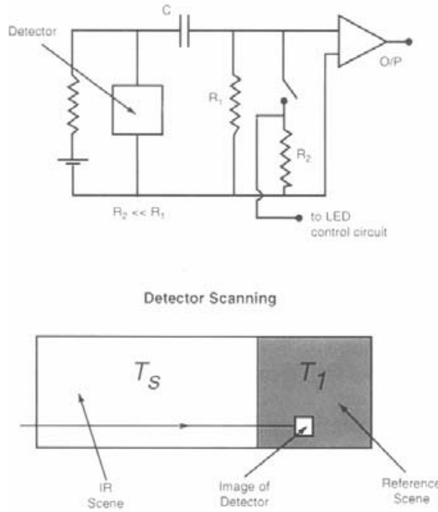


Figure 29. Schematic of the detector bias circuit and a.c. coupling of the output with a reset switch. The lower diagram shows the detector scanning across the scene then onto the reference plane. After [52]

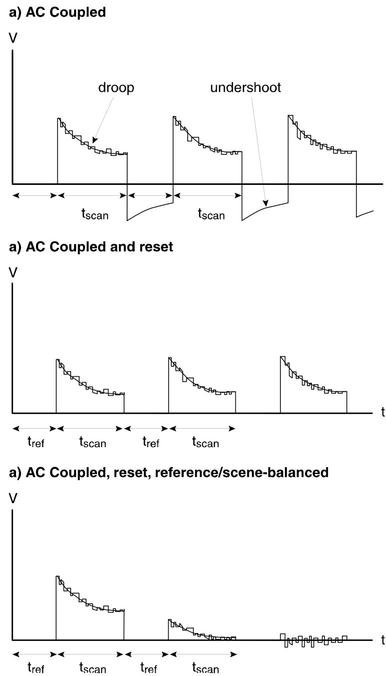


Figure 30. Detector output waveforms for: **a**) a.c. coupling, **b**) ac coupling with reset, **c**) a.c. coupling, reset and reference balancing. After [52]

channels are reset by short-circuiting the coupling capacitor to ground. Offset and droop problems still arise if the reference temperature is not close the mean scene temperature, as shown in Figure 30b, which are countered by automatically adjusting the reference temperature to correspond to that of the scene, as shown in Figure 30c. The time constant of this can be very long if the reference plane is a Peltier device, whereas the positive/negative LEDs can respond essentially instantaneously [52, 60].

5.5.2 Non-uniformity Correction in Staring Arrays

Staring array imagers are directly coupled to the scene and so are particularly sensitive to non-uniformity in the responses of individual detectors and their

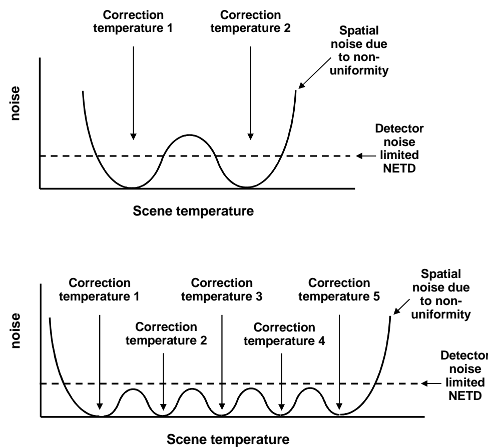


Figure 31. Schematic diagram showing the improvement in thermal imager performance that can be achieved using multi-point correction (*bottom curve*) as compared to standard two point correction (*top curve*)

associated read-out circuits. The detector outputs are related to the flux by non-linear polynomial equations, so the non-uniformity can be corrected by allowing the array to see a number of uniform reference temperatures, in order that the constants of the polynomial can be calculated and applied to the output signal. In practical systems, however, this correction has to be approximated by a linear relationship based on only two reference temperatures (‘gain and offset’ correction) owing to space and power dissipation problems of having multiple Peltier reference sources or the time for a single source to stabilise at a number of temperatures. This means that the correction is less effective for scene temperatures not equal to one of the reference temperatures, and can lead to increases in the minimum resolvable temperature difference (MRTD) of several tens of milli-Kelvin. The non-uniformity often changes with time, for example due to 1/f noise in the detector or physical temperature changes in the detector or the associated read-out electronics, so the calibration must be performed periodically - from perhaps every hour to as often as

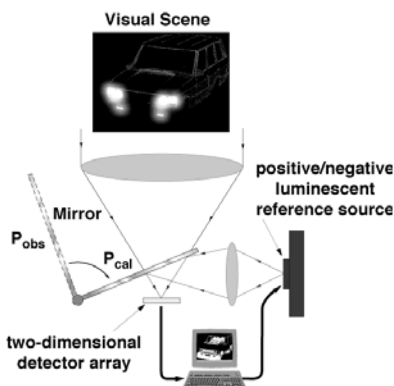


Figure 32. Illustration of the use of a positive / negative luminescent device for non-uniformity correction of staring arrays. After [52]

every frame in particularly demanding environments. In a similar manner to the scanned systems, the use of positive/negative luminescent devices will allow a larger number of effective temperatures, both above and below the physical scene temperature, to be produced virtually instantaneously and more frequently, if required, so leading to substantially improved performance [52,61].

5.6 Camouflage

As arbitrary photon fluxes can be produced, both above and below the ambient level, a wide range of apparent temperatures can be simulated when a positive/negative luminescent device is viewed with a thermal imager. Consequently, in principle, the actual temperature distribution of an object could be masked if a PL/NL device was placed in front of it or on it – hence the object's thermal signature could be modified and it would be camouflaged. In the extreme, the object could be made to appear as if it was at the same temperature as the background and hence disappear in to the backgound scene – the Infared Chameleon [53,62]. The practical difficulties of making devices which are large enough to perform this function for typical objects, and which would consume a reasonable amount of power, mean, however, that this application remains hypothetical.

6 Summary

Negative luminescent devices have progressed immensely in the 40 years since the first observation of the effect in semiconductors. Most of the progress during the first 30 years was achieved through the use of the magnetoconcentration effect. The progress has gained pace in the last ten years following the establishment of high quality epitaxial growth techniques for the narrow-gap semiconductors, which have enabled complex diode structures to be formed, so enabling negative luminescence to be produced in large area devices without the need for a magnetic field. The

technology has now progressed to the point where practical devices for a number of applications are beginning to be implemented. Consequently, the future looks cool for negative luminescence.

References

1. Antonov-Romanovsky VV, Stepanov BI, Fok MV and Khapalyuk AP. Vykhod lyuminestsentsii sistemy s tremya urovnyami energii so. Dokl. Acad. Nauk SSSR. 1955; 105:50-53
2. Ivanov-Omskii VI, Kolomiets BT and Smirnov VA. Recombination radiation in InSb due to the Magnetoconcentration Effect. Sov. Phys. Dokl. 1965; 10:345-346
3. Ivanov-Omskii VI, Kolomiets BT, and Smirnov VA. Spectrum of electroluminescence in InSb. Sov. Phys. JETP Lett. 1966; 3:185-187
4. Kessler FR and Mangelsdorf JW. Recombination radiation of germanium by free-carrier compression in crossed electric and magnetic fields. Phys. Stat. Sol. 1974; 24:557-564
5. Kessler FR and Mangelsdorf JW. Free carrier compression and galvanomagnetic luminescence of germanium at high Lorentz fields. Phys. Stat. Sol. 1981; 105:525-535
6. Bolgov SS, Malyutenko VK and Pipa VI. “Negative luminescence” in semiconductors. Sov. Tech. Phys. Lett. 1979; 5:610-611
7. Bolgov SS, Malyutenko VK and Pipa VI. Luminescence of semiconductors under carrier deficiency conditions. Sov. Phys. Semicond. 1983; 17:134-137
8. Malyutenko VK, Bolgov SS and Yablonovsky EI. A new type of IR luminescence. Infrared Phys. 1985; 25:115-119
9. Berdahl P, Malyutenko V and Morimoto T. Negative luminescence of semiconductors. Infrared Phys. 1989; 29:667-672
10. Bewley WW, Jurkovic MJ, Felix CL *et al*. HgCdTe photodectors with negative luminescent efficiencies > 80%. Appl. Phys. Lett. 2001; 78:3082-3084
11. Gordon NT, Jones CL and Purdy DJ. Application of microlenses to infrared detector arrays. Infrared Phys. 1991; 31:599-604
12. Ashley T, Dutton DT, Elliott CT, Gordon NT and Phillips TJ. Optical concentrators for light emitting diodes, Proc. SPIE 1998; 3289:43-50
13. Welford, WT and Winston R. The optics of non-imaging concentrators. Academic Press, New York, 1978.
14. Ashley T, Gordon NT and Phillips TJ. Optical modelling of cone concentrators for positive and negative IR emitters. J. Mod. Optics 1999; 46:1677-1696
15. Nash GR, Gordon NT, Ashley T, Emeny MT and Burke TM. Large-area IR negative luminescent devices. IEE Proc.: Optoelectron. 2003; 150:371-375
16. Malyutenko VK, Pipa VI, Yablonovsky EI and Kolesnikov IV. Negative-luminescence spectrum of InSb. Sov. Phys. Semicond. 1990; 24:544-547
17. Malyutenko VK, Yablonovsky EI, Bolgov SS, Beketov GV and Salyuk OY. Negative luminescence of $Cd_xHg_{1-x}Te$. Sov. Phys. Semicond. 1984; 18:211-212
18. Bolgov SS, Malyutenko VK, Pipa VI and Savchenko AP. Galvanomagnetic infrared luminescence of varying gap $Cd_xHg_{1-x}Te$/CdTe structures. Infrared Phys. 1992; 33:409-416
19. Morimoto T and Chiba M. Infrared emission from n-InSb under crossed electric and magnetic fields. Phys. Lett. 1981; 85A:395-398
20. Morimoto T and Chiba M. Characteristics of Luminescence from InSb magneto-infrared-emitting diode. Jap. J. Appl. Phys. 1984; 23:L821-L823

21. Berdahl P. Galvanomagnetic luminescence and the quantum efficiency of radiative recombination of InSb. J. Appl. Phys. 1988; 63:5846-5858
22. Berdahl P and Shaffer L. Galvanomagnetic luminescence of indium antimonide. Appl. Phys. Lett. 1985;47:1330-1332
23. Malyutenko VK, Bolgov SS and Malyutenko OYu. Multielement IR Sources with Alternating Contrast. Tech. Phys. Lett. 2001; 27:644-646
24. Malyutenko VK, Bolgov SS and Malyutenko OYu. Two-dimensional InSb array of IR emitters with alternating contrast. Infrared Phys. Technol. 2003; 44:11-15
25. Bolgov SS and Malyutenko OYu. Alternate contrast pixel-less IR image simulator. Infrared Phys. Technol. 2004; 45:249-252
26. Malyutenko VK, Yablonovsky EI, Savchenko AP, Bilinets YY and Kabatsy VN. Luminescence of InAs under magnetic injection conditions. Sov. Phys. Semicond. 1988; 22:370-372
27. Berdahl P. Radiant refrigeration by semiconductor diodes. J. Appl. Phys. 1985; 58:1369-1373
28. Bolgov SS, Malyutenko VK, Pipa VI and A.P Savchenko. Exclusion and accumulation of charge-carriers in narrow-band semiconductors at high temperatures. Pis'ma ZhTF 1989; 15:49-53
29. Ashley T, Elliott CT, Gordon, NT, Hall RS, Johnson AD and Pryce G.J. Negative Luminescence from $In_{1-x}Al_xSb$ and $Cd_xHg_{1-x}Te$ diodes. Infrared Phys. Technol. 1995; 36:1037-1044
30. Nash GR, Gordon NT, Hall DJ *et al.* Infrared negative luminescent devices and higher operating temperature detectors. Physica E 2004; 20:540-547
31. Lindle JR, Bewley WW, Vurgaftman I, Meyer JR, Varesi JB and Johnson SM. Efficient 3–5 μm negative luminescence from HgCdTe/Si photodiodes. Appl. Phys. Lett. 2003; 82:2002-2004
32. Bewley WW, Lindle JR, Vurgaftman I *et al.* Negative luminescence with 93% efficiency from midwave infrared HgCdTe diode arrays. Appl. Phys. Lett. 2003; 83:3254-3256
33. Lindle JR, Bewley WW, Vurgaftman I *et al.* Negative luminescence from mid-wave infrared HgCdTe diode arrays. Physica E 2004; 20:558-562
34. Lindle JR, Bewley WW, Vurgaftman I *et al.* A 5 mm x 5 mm mid-wavelength infrared HgCdTe photodiode array with negative luminescence efficiency >95%. J. Electron. Mat. 2004;33:600-603
35. Nash GR, Ashley T, Gordon NT, Jones CL, Maxey CD and Catchpole RA. Large Area IR Negative Luminescent Devices. Proc. SPIE 2003; 4820:132-139
36. Lindle JR, Bewley WW, Vurgaftman I *et al.* Negative Luminescence from Large-Area HgCdTe Photodiode Arrays with 4.8-6.0 μm cut-off Wavelengths. IEEE J. Quant. Electron. 2005; 41:227-233
37. Ashley T, Gordon NT, Nash GR, Jones CL, Maxey CD and Catchpole RA. Long-wavelength HgCdTe negative luminescent devices. Appl. Phys. Lett. 2001; 79:1136-1138
38. Nash GR, Ashby MK, Lindle JR *et al.* Long wavelength infrared negative luminescent devices with strong Auger suppression. J. Appl. Phys. 2003; 94:7300-7304
39. Haigh MK, Nash GR, Gordon NT *et al.* Long-wavelength HgCdTe on silicon negative luminescent devices. Appl. Phys. Lett. 2005; 86: 011910
40. Nash GR, Ashley T, Gordon NT, Jones CL, Maxey CD and Catchpole RA. Micromachined optical concentrators for IR negative luminescent devices. J. Mod. Opt. 2002; 49:811-820
41. Pullin MJ, Hardaway HR, Heber JD and Phillips CC. Type-II InAs/InAlSb strained-layer-superlattice negative luminescent devices. Appl. Phys. Lett. 1999; 75:3437-3439
42. Pullin MJ, Hardaway HR, Heber JD *et al.* Room-temperature InAsSb strained-layer superlattice light-emitting diodes at $\lambda = 4.2$ μm with AlSb barriers for improved carrier confinement. Appl. Phys. Lett. 1999; 74:2384-2386

43. Olafsen LJ, Vurgaftman I, Bewley WW *et al*. Negative luminescence from type-II InAs/GaSb superlattice photodiodes. Appl. Phys. Lett. 1999; 74:2681-2683
44. Aidaraliev M, Zotova NV, Karandashev SA *et al*. Negative Luminescence in p-InAsSbP/n-InAs diodes. Semiconductors 2001; 35:321-324
45. Matveev BA, Aydaraliev M, Zotova NV *et al*. Negative luminescence from InAsSbP based diodes in the 4.0-4.3 μm range. Proc. SPIE. 2001; 4285:109-117
46. Matveev BA, Zotova NV, Karandashev SA, Remennyi MA, Stus NM and Talalakin GN. Towards longwave (5-6 μm) LED operation at 80C: injection or extraction of carriers? IEE Proc.: Optoelectron. 2002; 149:33-35
47. Matveev BA, Zotova NV, Karandashev SA, Remennyi MA, Stus NM and Talalakin GN. Towards efficient mid-IR LED operation: extraction or injection extraction of carriers? J. Mod. Opt. 2002; 49:743-756
48. Malyutenko VK, Malyutenko OYu, Podoltsev AD *et al*. Current crowding in InAsSb light-emitting diodes. Appl. Phys. Lett. 2001; 79:4228-4230
49. Aidaraliev M, Zotova NV, Il'inskaya ND *et al*. InAs and InAsSb LEDs with built-in cavities. Semi. Sci. Technol. 2003; 18:269-272
50. Bolgov SS, Malyutenko VK and Savchenko AP. Charge-carrier exclusion in InAs. Semiconductors 1997; 31:444-445
51. Beattie AR and White AM. An analytic approximation with a wide range of applicability for electron initiated Auger transitions in narrow-gap semiconductors. J. Appl. Phys. 1996; 79:802-813
52. Ashley T, Elliott CT, Gordon NT, Phillips TJ and Hall RS. Applications of negative luminescence. Infrared Phys. Technol. 1997; 38:145-151
53. Malyutenko VK. Negative Luminescence in semiconductors: a retrospective view. Physica E. 2004; 20:553-557
54. Ashley T, Crowder JG, Mannheim VP and Smith SD. Infrared light emitting diodes. GB Patent No. 2346481, filed 1998
55. Ashley T, Elliott CT and Gordon NT. Dynamic infrared scene projector. GB Patent No. 2319664. filed 1995
56. Nash GR, Gordon NT, Emeny MT and Ashley T. Perspectives on dynamic infrared scene projection using positive and negative luminescence. Proc SPIE. 2003; 5092:138-144
57. Berdahl P. Proc. 18th Int. Conf. on Physics of Semiconductors (Ed. O. Engstrom) 1987; 2:1595
58. Liptuga AI, Malyutenko VK, Pipa VI and Levash LV. Radiative cooling under the conditions of magnetoconcentration. Semiconductors 1997; 31:423-426
59. Ashley T, Elliott CT, Gordon NT and Hall RS. Infrared optical system. GB Patent No. 0861426 filed 1995
60. Ashley T, Elliott CT, Gordon NT and Hall RS. Black level clamp for scanned thermal imaging systems. GB Patent Application No. 9520060.6 filed 1995
61. Ashley T, Elliott CT, Gordon NT and Hall RS. Thermal imaging system. GB Patent No. 2317779 filed 1995
62. Elliott CT, Ashley T and Dutton DT. How to look cool – controlling the infrared. Brochure produced for a Royal Society meeting 1997

Mid-infrared Quantum Dot Photodetectors

P. Bhattacharya[1], A. D. Stiff-Roberts[2] and S. Chakrabarti[1]

[1]Department of Electrical Engineering and Computer Science,
University of Michigan
Ann Arbor, MI USA
[2]Department of Electrical and Computer Engineering
Duke University
Durham, NC USA

1 Introduction

Infrared (IR) detectors are used in a range of imaging applications, including environmental monitoring, medical diagnosis, and industrial equipment diagnosis. The typical components of an IR camera system include: optics for light collection, a focal plane array (FPA) for detection and signal processing, a cooling system for photodetector arrays, and electronics for the digital images. A significant reduction in the cost of IR camera systems is possible if the traditional cooling systems, such as liquid-nitrogen dewars, are replaced by thermo-electric coolers. Such a design change requires the development of an IR photodetector that operates at elevated temperatures ($\geq$ 120 K). Low bandgap, compound semiconductors appropriate for mid-infrared (MIR) detection from 3-5 μm, such as HgCdTe [1-7], are inherently soft and brittle. Thus, it can be difficult to epitaxially grow and fabricate MIR devices with bulk material active regions. However, large bandgap materials that are more amenable to various fabrication techniques become an option if the advantages of extrinsic detection are realized.

An important extrinsic detection technique is to reduce the dimensionality of the photodetector active region so that *intraband* photoexcitation is possible. Through lithographical and/or epitaxial growth techniques, active region heterostructures with dimensionality on the order of the deBroglie wavelength for an electron can be designed. AlGaAs/GaAs quantum well (QW) heterostructures, which are quantum-confined in one dimension, are an important technology for IR photodetection [8-12]. Quantum dot infrared photodetectors (QDIPs), a relatively new technology discussed in this chapter, have great potential to provide the requisite detector array performance at high temperatures due to three-dimensional quantum confinement of the detector active region. Since the first extensive QDIP characterization in 1999 at the University of Michigan [13], several research groups around the world have investigated these devices [14-19], seeking to push the boundaries of state-of-the-art performance, particularly in the MIR wavelength range.

1.1 Quantum Dot Photodetectors in the Mid-infrared

Heterostructures with three-dimensional (3-D) quantum confinement are called quantum dots (QDs), and these islands are buried in a compound semiconductor of larger bandgap energy such that a confinement potential is provided. The electrical

and optical properties of QDs are extremely dependent upon their size, shape, and uniformity. Therefore, the technique used to obtain a QD ensemble is critical to the IR photodetector characteristics. The epitaxial growth technique commonly used for the formation of nanoscale, 3-D islands is discussed in more detail in Section 3. It is important to note the properties of QDs that are favorable for high performance devices. First of all, if the QDs are too small, charge carriers will not be confined; if the QDs are too large, the intraband energy transitions will lie outside the IR range of interest and have separations less than the thermal energy of electrons. Second, for the fabrication of large area devices (such as detector arrays for IR cameras), QD ensembles should have a large area density, low defect concentration, and large uniformity of QD size and shape.

For typical growth parameters used in molecular beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOCVD), an ensemble of InAs/GaAs QDs typically has lateral sizes from 15 to 25 nm and heights from 5 to 8 nm. The density of InAs/GaAs QDs ranges from 10^9 to 10^{11} cm^{-2}. Self-assembled InAs/GaAs QDs that have been grown by MBE are shown in the cross-sectional transmission electron microscopy (XTEM) images of Figure 1. It is important to note that in addition to a pyramidal shape, lens-shaped QDs are also observed in a different area of the same sample.

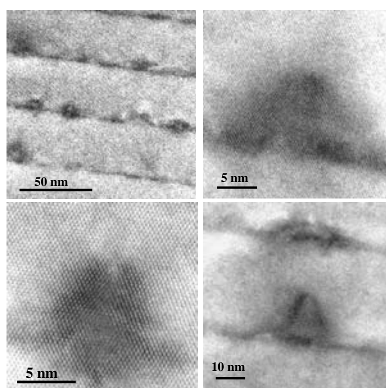


Figure 1. Cross-sectional TEM images of InAs/GaAs QDs grown by MBE, demonstrating significant size and shape variation within a single QD heterostructure. (The TEM images were obtained by the Browning group at the U. Illinois-Chicago).

InAs/GaAs quantum dot IR photodetectors (QDIPs) use these nanoscale islands for the extrinsic detection of IR light through intraband transitions in the conduction band of the InAs/GaAs heterostructure [13,14,16,20-28]. QDIPs are unipolar devices, using electrons (typically) for charge transport since the separation of confined energy levels in the conduction band corresponds to IR wavelengths of interest.

A typical device conduction band diagram under flat-band conditions is shown in Figure 2. InAs/GaAs QDIPs, like their QW counterparts, benefit from mature III-V growth and processing technologies. In addition, the quantum-mechanical nature of QDIPs leads to several additional advantages over QWIPs and the other types of IR detectors that are available. As in the case of the other photodetector technologies,

QDIPs provide multi-wavelength detection; however QDs provide many more parameters for tuning the energy spacing between bound energy levels, such as QD size and shape, strain, and material composition.

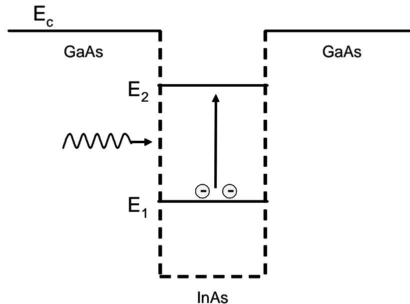


Figure 2. Schematic diagram of the conduction band in an InAs/GaAs QDIP

A second advantage of QDIPs is that QDs have a large electron relaxation time from excited states due to the phonon bottleneck. As a result, QDIPs are expected to be more efficient than QWIPs since photoexcited carriers are more likely to be collected as photocurrent before relaxing back into the ground state of the QD. Third, QDIPs are theoretically predicted to have lower dark currents than other IR photodetectors due to the 3-D quantum confinement of the electron wavefunction. This confinement leads to equivalent values for the photoionization energy and the thermionic emission activation energy for electrons captured within QDs. Thus, electrons should be photoexcited by incident IR radiation before they are thermally excited out of the QD. Both the increased electron lifetime and the reduced dark current indicate that QDIPs should be able to provide high-temperature operation (≥ 120 K). Achieving operating temperatures in this range is very important because it can reduce both the cost and weight of IR cameras by enabling the use of thermoelectric coolers. Finally, QDIPs are inherently sensitive to IR light incident in the surface-normal direction since the existence of discrete energy levels in all three dimensions eliminates absorption limitations due to polarization selection rules.

While the 3-D quantum confinement of the active region in QDIPs is expected to provide all of the advantages just related, in practice, it has been a challenge to meet all of these expectations. For some time, QDIPs have demonstrated performance that is, at best, comparable to that for QWIPs. This discrepancy between predicted and realized advantages is due in large part to the epitaxial growth mechanism of QDs. In fact, the non-uniformity of QDs, resulting from the Stranski-Krastanow growth mode, is the most significant disadvantage of QDIPs. The random fluctuations of size, material composition, and doping in QDs lead to an inherently large full-width half-maximum (FWHM) linewidth (~ 50 meV) for transitions in a QD ensemble. This large, inhomogeneously broadened linewidth has a deleterious effect on QDIP performance. Subsequently, the quantum efficiency in these devices tends to be lower than what is predicted theoretically. By improving QD uniformity, the linewidth of the QD ensemble should decrease, thus increasing the absorption coefficient and improving the performance of QDIPs. Thus the epitaxial growth and

design of unique QD heterostructures is one of the most important issues related to obtaining state-of-the-art performance in QDIPs.

1.2 Comparison of HgCdTe Photodiodes and InAs/GaAs Quantum Dot Infrared Photodetectors

Since the QDIP is an intraband detector, it differs in its principle of operation and materials requirements from interband, HgCdTe photodiodes. Several fundamental parameters that define IR detector operation are very different for InAs/GaAs QDIPs and HgCdTe photodiodes. Principle among these is the carrier lifetime, which affects the dark current very differently for these two devices. In HgCdTe photodiodes ($\tau_{carrier}$ ~ 10 μs), the dark current is primarily diffusive and depends inversely on the minority carrier lifetime. However, the dark current in a QDIP does not depend on the carrier lifetime ($\tau_{carrier}$ ~ 100 ps-1 ns) significantly; instead, it is determined primarily by thermionic emission and field-assisted tunneling rates. However, the carrier lifetime is critical in determining the responsivity, photocurrent, and gain in both types of devices (see Appendix A). The product of the detector differential resistance and optical area (R_oA product for zero bias) is an extremely important parameter for characterizing IR detectors and for determining the suitability of a particular detector for incorporation into a FPA. The knowledge of the R_oA product provides the range of background flux values over which the detector is either background-noise limited (BLIP) or Johnson-noise limited (JOLI). For large R_oA products, the $D*$ continues to improve with decreasing background flux. As a result, the transition from the BLIP condition to the JOLI condition occurs at lower background flux so that the maximum $D*$ possible increases [29].

The interpretation of the R_oA product also exhibits subtle, but important, differences for intrinsic photodiodes and extrinsic photoconductors. In a photodiode, the R_oA product is an intrinsic property of the material related to the dark current. For bulk material photodiodes, it is not possible to increase the detector area without reducing the device resistance, leading to the observation that the R_oA product is a constant for a given intrinsic material. The R_oA product in a photodiode also depends on the cutoff wavelength of the bulk material, since the wavelength is directly related to the bandgap of the material. However, the situation is different for a PC detector. In this case, the R_oA product is not an inherent property of the material. While the resistance to dark current flow at zero bias can be extrapolated from *I-V* measurements, the R_oA product is not constant, but depends on the device geometry [3, 29]. In addition, the R_oA product does not display any consistent wavelength dependence in QDIPs, since an extrinsic detection scheme is used. Therefore, a more meaningful parameter for comparison is the dark current density. In order to fully explore these differences, a comparison of the R_oA product in HgCdTe photodiodes and InAs/GaAs QDIPs has been conducted.

In Figure 3(a), experimental QDIP R_oA values as a function of the peak wavelength, λ_{peak}, determined from the dynamic resistance of *I-V* data (Figure 3 (b)), are shown. Also shown in Figure 3(a) are measured and calculated R_oA values of HgCdTe devices obtained from Rockwell Scientific. As expected, R_oA products for the QDIPs do not demonstrate the same dependence on λ_{peak} as HgCdTe detectors. It is important to note that while the dark current of junction detectors increases with

the peak wavelength, the dark current in QDIPs is far less sensitive to wavelength. Thus MIR QDIPs should have comparable J_{dark} values to QDIPs detecting longer IR wavelengths.

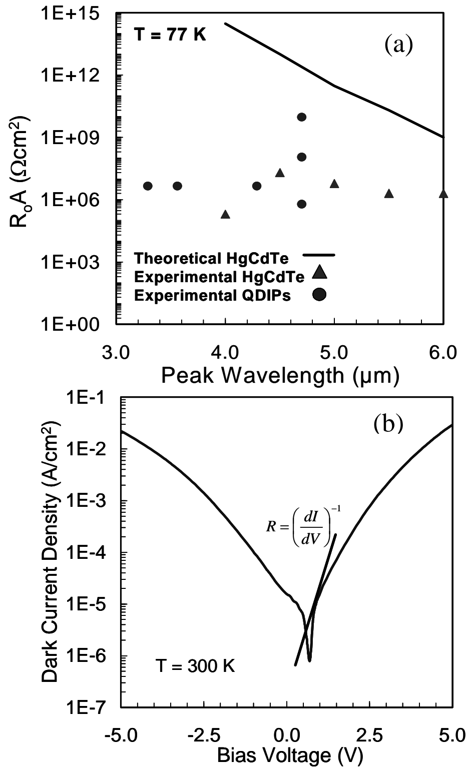


Figure 3. (**a**) R_oA values in HgCdTe photodiodes and InAs/GaAs QDIPs. HgCdTe data was obtained from Rockwell Scientific, and QDIP data was determined from dynamic resistance in *I-V* curves at operating bias, as shown in (**b**).

2 Infrared Detection with Quantum Dots

The mechanism of IR detection in QDs is shown schematically in Figure 4. When IR photons impinge on the QDIP, photo-excited carriers escape from the QD so that they can be collected as photocurrent. However, once photoexcited carriers are in the continuum, they can experience several additional processes. Possible transport mechanisms include: 1) drift in the continuum under the influence of an electric field, 2) capture into the excited state of either the same QD or a different one, and 3) collection at a device contact. These processes are shown schematically in Figure 5.

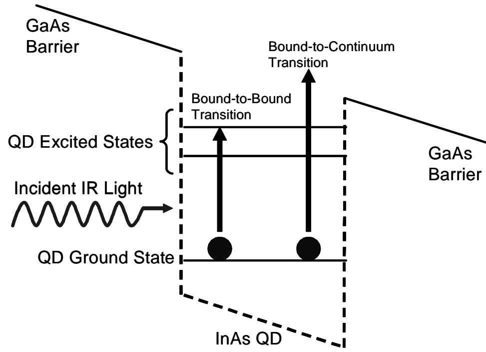


Figure 4. Schematic diagram of InAs/GaAs QDIP conduction band under bias

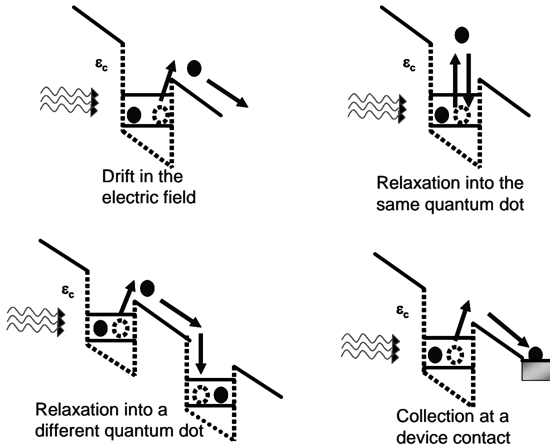


Figure 5. Schematic depiction of possible transport mechanisms for photoexcited electrons in the QDIP active region

2.1 Bound-state Energy Levels in QDs

Technological advancements in the lithography and epitaxial growth of compound semiconductors have enabled the physical realization of some fundamental concepts in quantum mechanics. In particular, for an idealized description of the dot system, the QD is the physical equivalent of the particle-in-a-box problem for three dimensions. The properties of QDs that affect the formation of discrete energy levels can be determined by solving, analytically, the time-independent Schrödinger equation for a QD with a spherical, finite barrier. In the given geometry, an electron is assumed to be captured within the potential well created by a 3-D, quantum-

confined sphere (radius, $R_0 \leq \lambda_{deBroglie}$ for an electron) of InAs buried within a GaAs matrix. Therefore, the electron has an effective mass m_1 within the potential well and an effective mass m_2 outside the well. Due to the spherical symmetry of the potential well, QDs have discrete, bound-state energy levels in three dimensions, similar to those for an atom. The potential barrier height is determined by the bandgap energy and the band line-up of the materials at the junction of the quantum-confined heterostructure, and there is a minimum confinement potential, V_o, below which bound-state energy levels do not exist. However, to determine energy levels in more realistic QDs, the 8-band **k·p** model [30] is used.

In applying the **k·p** theory to self-assembled QDs, it is necessary to take into account the unique strain distribution of these coherently-strained, quasi-zero-dimensional heterostructures. As has been demonstrated experimentally, the large built-in strain of the QDs causes a significant change in the properties of the electronic bandstructure. For example, the bandgap of InAs/GaAs QDs has been measured to be ~ 1.05 eV, even though the intrinsic bandgap of bulk InAs is 0.4 eV [30]. Thus the effect of strain is incorporated into the energy bandstructure of a semiconductor by: 1) determining the elastic strain energy of the system, and 2) including this energy in the Hamiltonian as an additional perturbation through deformation potential theory. Jiang and Singh [30] have used the valence force field (VFF) model [31,32] to determine the strain distribution in self-assembled, InAs/GaAs QDs for the eight-band **k·p** calculation of the electronic bandstructure, yielding a very detailed strain tensor for QDs and the surrounding barrier material. They then use the finite-difference method to solve the matrix equation for the perturbed, **k·p** Hamiltonian.

Using the eight-band **k·p** model developed by Jiang and Singh, the first four bound states in pyramidal, InAs/GaAs QDs with dot height of 6.8 nm and base width of 13.5 nm have been calculated. The listed energy levels are assumed to be from the InAs valence band edge. Therefore, in Figure 6, the QD intraband energy levels in the conduction band are shown schematically, assuming that the bandgap in strained InAs is 1.05 eV. According to this calculation, transitions from the QD ground state to the top excited state correspond to 9.25 μm. Thus, the electron intraband energies corresponding to MIR light occur for transitions from the QD to the continuum surrounding the QD.

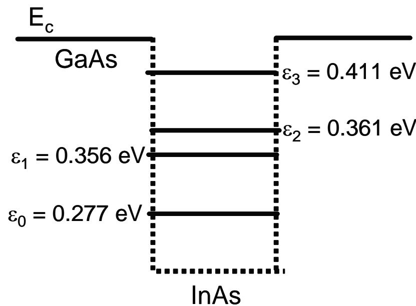


Figure 6. Calculated energy levels using the eight-band **k·p** model developed by Jiang and Singh for an InAs/GaAs QD with height = 6.8 nm and base width = 13.5 nm

2.2 Density of States and Carrier Distribution in QDs

Another unique property of QDs is the zero-dimensional density of states, which leads to favorable carrier dynamics in opto-electronic devices. The electron distribution in the active region is determined by the product of the density of states and the Fermi function for the probability that a given energy level will be occupied by an electron.

To appreciate the importance of this result, the carrier distributions in bulk material and QDs are compared. In bulk material, the density of states, shown schematically in Figure 7 (a), has a square-root dependence on the electron energy. Thus, the product of this density of states with the Fermi function yields a carrier distribution, shown schematically in Figure 7 (b) that decreases exponentially as the temperature is increased. In the case of QDs, the k-space is zero-dimensional and represented by a point. As a result, the density of states in QDs is *not* continuous for any direction. However, this infinite density of states is for an ideal QD. In reality, the size of QDs varies spatially such that the density of states is actually represented by inhomogeneously broadened δ-functions in a Gaussian distribution

$$N(\hbar\omega) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(E_{fi} - \hbar\omega)^2}{2\sigma^2} \right) \tag{1}$$

where E_{fi} is the energy separation between states f and i, and σ is the Gaussian linewidth of the transition. The QD density of states and corresponding electron distribution are shown schematically in Figure 7 (c) and (d), respectively.

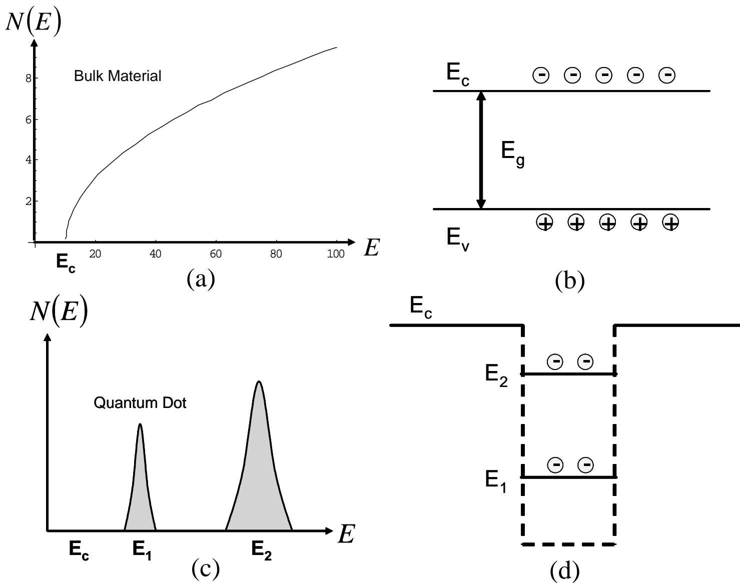


Figure 7. Schematic diagrams of the density of states in (**a**) bulk material and (**c**) QDs, and the corresponding carrier distribution in (**b**) bulk material and (d) QDs

2.3 Intraband Absorption in QDs

In general, there are two types of intraband transitions that can originate in the QD. The first type of transition is called a bound-to-bound (B-B) transition, in which ground state electrons are photoexcited to the excited state of the QD. The second type of transition is called bound-to-continuum (B-C), and in this case, an electron in the ground- or excited-state is photoexcited out of the QD into the continuum of energy levels for the barrier bulk material. These different transitions are shown in Figure 4.

For either case, the intraband absorption process requires the presence of electrons in a bound state of the QD. Electrons can be introduced in the QD by direct-doping [23], modulation doping [33], or capture of injected electrons from the device contact [17]. It has been shown that QDIP performance has a distinct tradeoff between, and dependence on, QD density and dopant concentration [34]. For this reason, it is customary to design QDIPs with doping such that there are approximately two electrons per QD ground state. For higher dopant densities, the dark current increases since electrons occupy excited-state energy levels in QDs, enabling these carriers to contribute to the noise in the device more easily by thermionic emission or field-assisted tunneling. For lower dopant densities, the absorption efficiency decreases since there are fewer electrons in the QD ground state available for photoexcitation.

The QD absorption spectrum is generally broad due to the size variation of QDs comprising the active region of a device. Therefore, QDIPs experience another tradeoff between the absorption strength and the absorption linewidth. Thus, even though each QD individually demonstrates sharp atom-like electronic states, for an array of QDs, the density of states is an inhomogeneous distribution that typically resembles a Gaussian. The linewidth of this Gaussian distribution, σ (measured in meV), significantly influences the absorption of an array of QDs since the absorption coefficient, α (measured in cm^{-1}), is inversely proportional to the linewidth [35]

$$\alpha \approx \frac{3.5x10^{5}}{\sigma} \tag{2}$$

Figure 8 shows a normalized spectral response curve measured in InAs/GaAs QDIPs and peaking in the MIR wavelength range. As demonstrated in Figure 8 the spectral response can be broad, quantified by the spectral linewidth, $\Delta\lambda/\lambda_{peak}$, where $\Delta\lambda$ is the FWHM linewidth. For spectral linewidths ≥ 0.3, the spectral response is generally believed to be due to B-C transitions as opposed to B-B transitions [24]. Another important feature of the spectral response is the long-wavelength cutoff, λ_c, which corresponds to the smallest detectable energy in the photodetector. In intrinsic photodetectors, such as bulk photodiodes, this cutoff wavelength corresponds to the bandgap energy of the material, and is therefore approximately equal to the peak wavelength. However, since QDIPs use an extrinsic detection scheme, the cutoff wavelength does not equal the peak wavelength and does not correspond to the bandgap energy. However, the cutoff wavelength does indicate the QD activation energy in the detector.

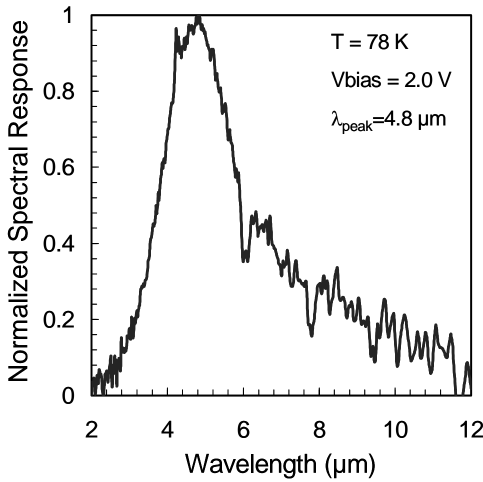


Figure 8. Normalized spectral response curve for a MIR InAs/GaAs QDIP

Another important aspect of intraband absorption in QDs is related to the polarization selection rules for incident radiation. The momentum matrix element for intraband transitions is calculated using the electron wavefunctions determined by the eight-band **k·p** model in pyramidal QDs:

$$\mathbf{p}_{fi} = \langle f|\mathbf{p}|i\rangle = \sum_{j,k} \langle \phi_{fj} | \phi_{ik} \rangle \langle u_j | u_k \rangle + \sum_j \langle \phi_{fj} |\mathbf{p}| \phi_{ij} \rangle, \tag{3}$$

where ϕ and u are the envelope and central-cell functions, respectively. In conjunction with the polarization of the electric field, the momentum matrix element determines the selection rules for permitted transitions in which photons are absorbed (or emitted). An optical transition is forbidden if the dot product between the electric field polarization and the momentum matrix element is zero, since the resulting transition rate would equal zero. Considered another way, for intraband transitions in the direction of quantum-confinement, momentum is conserved as electrons are excited from one discrete energy level to another. However, in the direction of bulk dimensions, intraband transitions in the parabolic band do not conserve momentum unless phonon-assisted scattering occurs. Therefore, QWs, which are quantum-confined in the growth direction (or z-direction) only, experience momentum conservation in the surface-normal direction. As a result, QWs cannot detect surface-normal incident light because the electric field polarization lies in the plane perpendicular to the growth direction. Therefore, lateral incidence is approximated in QWIPs by scattering surface-normal incident light. In contrast, QD devices are more flexible in that lateral- or surface-normal-incidence may be used for intraband transitions to occur. This flexibility results from the fact that QDs experience 3-D quantum confinement, and therefore conserve momentum for intraband transitions in any direction.

2.4 Phonon Bottleneck and Effective Carrier Lifetime in QDs

The carrier dynamics in 3-D, quantum-confined heterostructures differ significantly from those of bulk material and QWs. In particular, while first-order, electron-phonon interactions cause rapid relaxation in bulk and QW semiconductor materials, electron-phonon coupling is suppressed in QDs due to the magnitude of the intraband energy spacing between confined levels in QDs [36,37]. More specifically, the intraband energy spacing is > 50 meV for practical QD sizes, while the longitudinal-optical (LO) phonon energy is < 40 meV [38]. Thus the conservation of the electron energy and momentum cannot be maintained by the emission of a single, LO phonon. Similarly, single, longitudinal-acoustic (LA) phonon emission (achieved by coupling with the deformation potential) is reduced for intraband energy spacings greater than a few meV [39]. This reduction of electron-phonon scattering for QDs, which leads to longer electron lifetimes in QDs (and increased photoconductive gain in detectors), is known as the phonon bottleneck.

Several experiments have been conducted to confirm the existence of the phonon bottleneck, including electroluminescence and time-resolved photoluminescence measurements of InGaAs/GaAs QDs [40], high frequency electrical impedance measurements in $In_{0.4}Ga_{0.6}As$/GaAs QD lasers [41], and time-resolved differential transmission spectroscopy (DTS) measurements of $In_{0.4}Ga_{0.6}As$/GaAs QDs [42]. DTS measurements have resolved two time constants for relaxation from the excited state to the ground state in $In_{0.4}Ga_{0.6}As$/GaAs QDs. The short time constant (~7 ps) is governed by Auger, e-h, and multiphonon scattering in a geminate carrier configuration. The long time constant (~750 ps) is governed by phonon-mediated scattering in a non-geminate carrier configuration [42], and is evidence of the phonon bottleneck in QDs. In addition to the long electron lifetime in the QD excited states, the phonon bottleneck also promises temperature-independent performance since the electron is energetically de-coupled from the optical phonon, the properties of which depend very heavily on the temperature. Thus, the suppression of phonon scattering in QDs suggests that electron relaxation lifetimes are significantly longer than those in QWs (~1-10 ps) [43], which are dominated by scattering due to the emission of single optical phonons.

The phonon bottleneck also affects the effective carrier lifetime in QDs. The effective carrier lifetime is essential to QDIP performance in that if the lifetime is too short, a photoexcited carrier will relax to the ground state before it has a chance to be collected as photocurrent. However, if the effective carrier lifetime is long enough, an efficient detector can be realized since photoexcited carriers are more likely to contribute to the photocurrent. Boaz Kochman, *et al.*, calculated the effective electron lifetime in InAs/GaAs QDIPs using the Monte Carlo technique for simulating transport in QD heterostructures [44]. For MIR QDIPs, it is beneficial to review some of the important results from the Kochman, *et al.* model for B-C transitions. First, at low fields, the average excess electron energy above the QD heterojunction step was well below the phonon energy, and the effective carrier lifetime changed very little as a result. However, for electric fields (F) greater than 2 kV/cm, the average excess electron energy exceeded the phonon energy, and carrier capture was suppressed resulting in longer effective lifetimes. For a fixed electric field value, the dependence of the effective carrier lifetime, τ_{eff}, on the low-field

capture time, τ_c, was linear. Also, the slope of the τ_{eff} –F plot is proportional to τ_c, given by [44]

$$\frac{d\tau_{eff}}{dF} \cong 0.25 \frac{\tau_c}{F} \tag{4}$$

where F is in kV/cm and τ_c is in ps. It is important to note that for typical InAs/GaAs QDIPs, the electric field is ~ 5 kV/cm at the operating bias, which corresponds to an effective electron lifetime of ~ 0.1 ns for a capture time of 100 ps [44].

3 Self-assembly of Quantum Dots by the Stranski-Krastanow Growth Mode

The rapid improvement of epitaxial growth techniques, such as MBE and MOCVD, has enabled the routine growth of excellent compound semiconductor materials [45]. In particular, the epitaxial growth of quantum-confined heterostructures has been pursued for decades due to enhanced performance of electronic and opto-electronic devices using such materials in the active region. While MBE is well suited for the growth of QWs, in which the active region experiences quantum confinement in the growth direction only, achieving the 3-D quantum confinement required for device-quality QDs is much more challenging.

Given that quantum confinement in the z-direction (growth direction) is achieved by the epitaxial growth of a compound semiconductor heterostructure of appropriate width, the challenge is to provide quantum confinement in the lateral direction. Several fabrication and growth techniques have been investigated for obtaining lateral confinement of the device active region, however, most do not me*et al.*1 of the necessary QD requirements. For example, a straightforward approach for obtaining 3-D quantum confinement is to pattern QW structures using lithographic and etching techniques [46-48]. However, this method has not proven successful for device applications due to the inability to achieve high quality interfaces after etching and regrowth, as well as the need to use electron-beam lithography. Growth on patterned substrates or substrates with preferred crystallographic planes are additional options for achieving 3-D quantum confinement, yet, these techniques are unable to provide high-density QD ensembles. Currently, the most successful approach to achieve defect-free, multiple-layer, high-density, QD ensembles is the heteroepitaxy of coherently-strained, 3-D, self-assembled islands [49-56], a process referred to as the Stranski-Krastanow growth mode [57, 58].

Strained-layer epitaxy is controlled by the strain energy between the substrate and the epitaxially-grown overlayer, or epi-layer. Typically, the epi-layer is biaxially strained in the plane of the substrate and uniaxially strained in the perpendicular direction. The preferred growth mode for strained materials is determined by the interface energy, the surface energy, and the lattice mismatch. In lattice-mismatched systems, where the strained epi-layer has small interface energy and the lattice mismatch is >1.8 %, the Stranski-Krastanow (S-K) growth mode occurs. In the S-K growth mode, layer-by-layer growth occurs initially, and this 2-D epi-layer is called the wetting layer. However, as the epi-layer thickness increases, the strain energy in

the system also increases, and this strain energy is minimized by the formation of 3-D islands. These islands have been shown to be coherently-strained, as opposed to dislocated, which is advantageous for device applications.

Thermodynamic equilibrium is achieved when the total free energy is minimized given a fixed amount of deposited material. The driving forces for achieving equilibrium in these self-assembled QDs are elastic relaxation at the facet edges, strain-induced renormalization of the surface energy at the facets, and interaction between neighboring islands via the substrate. For a given volume of a coherently-strained island, there exists an equilibrium shape. For growth along the [100] direction in the InAs/GaAs QD system, the self-assembled QDs are small and pyramidal-shaped with a nearly square base in the (100) plane [30].

The lattice mismatch at the interface of an InAs epi-layer and a GaAs substrate is

$$f = \frac{\left(a_{InAs} - a_{GaAs}\right)}{a_{GaAs}} \approx 7\% \tag{5}$$

where the lattice constant in InAs is 6.06 Å and the lattice constant in GaAs is 5.65 Å. Therefore, the S-K growth mode for the self-assembly of 3-D islands will occur under appropriate growth conditions. The typical InAs growth rate is 0.1 ML/s, and the GaAs substrate temperature is usually cooled to 500 °C. This slow rate, coupled with the low substrate temperature, enables the self-assembly of QDs since surface kinetics are minimized and thermodynamic equilibrium by the coherent incorporation of strain can be achieved more easily. The wetting layer thickness in the InAs/GaAs QD system is 1.7 monolayers (ML), where 1 ML ~ 2.83 Å for III-V materials. The transition from the 2-D wetting layer to 3-D islands is observed using *in-situ* reflection high-energy electron diffraction (RHEED), in which a streaky pattern gives way to a spotty pattern when QDs begin to form. Once this transition is observed, InAs overgrowth occurs for 2-5 s to provide enough InAs charge (2.2 ML) for pyramidal QDs, and a 30 s growth-interrupt pause is used to allow the complete formation of QDs. After this pause, an intrinsic GaAs cap layer (~ 25-60 nm) is grown on top of the InAs QDs, thereby completing the QD potential barrier. This sequence of growth is repeated to achieve a specified number of QD layers.

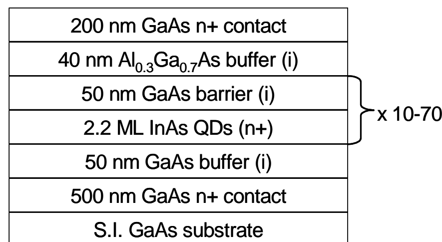


Figure 9. Schematic diagram of a typical MBE heterostructure for an InAs/GaAs QDIP featuring a single $Al_{0.3}Ga_{0.7}As$ current-blocking barrier

Typically, the QDs are directly-doped with silicon in order to provide free electrons for photoexcitation, and the carrier concentration ranges from $0.5\text{-}1\times10^{18}$ cm^{-3}. The doping density provides two electrons per dot for typical InAs/GaAs QD densities. During direct-doping, an effort is made to incorporate dopants in the QDs only and not in the wetting layer. Therefore, the silicon shutter is opened after the completion of the wetting layer, and it remains open during the subsequent dot overgrowth and growth-interrupt pause. A schematic diagram of a typical QDIP heterostructure grown by MBE is shown in Figure 9.

4 Characterization of Mid-infrared Quantum Dot Photodetectors

The standard QDIP device heterostructure comprises repeated InAs QD layers buried between GaAs barrier regions with top and bottom doped contact layers at the active region boundaries. Standard photolithography, metallization, and wet-etching techniques can be used to fabricate devices with either vertical or lateral photoconduction geometries. QDIPs using a vertical geometry typically require three lithography steps: 1) metal evaporation of the top contact, 2) device definition by mesa etch around the top contact, and 3) metal evaporation of the bottom contact around the device mesa. The device mesa height can vary from 1-4 μm depending on the device heterostructure. The optical area of the QDIP is assumed to be the detector mesa area, and surface-normal incident IR light is detected. It is important to note that the actual optical area of the QDIPs could be much larger than the mesa area when the heterostructure is grown on a semi-insulating, GaAs substrate since IR light can be reflected within the GaAs wafer. Once device fabrication is complete, the QDIP must be packaged for testing. Since QDIPs must be measured at cryogenic temperatures, the sample is typically mounted on a chip carrier designed for a given cryostat and wirebonds are used to make device contact.

In order to compare the performance of various IR detectors, there are several figures of merit that must be determined, either through direct measurement or a standard estimation. These figures of merit are used to completely characterize IR detectors, and they not only provide normalized performance parameters, but they also indicate which devices are suitable for incorporation into FPAs for IR imaging. Thus, the figures of merit discussed in this section [29,59,60] are the most significant in terms of analyzing the performance of discrete QDIPs before the fabrication of a detector array.

4.1 QDIP Heterostructure Designs

Several QDIP heterostructure designs have been investigated for use as IR photodetectors in addition to the standard InAs/GaAs QDIP. As an example, InAs QDs embedded in a strain-relieving InGaAs quantum well are known as dot-in-a-well (DWELL) heterostructures. QDIPs using a DWELL heterostructure not only permit greater control over wavelength tenability [61], but they have also demonstrated excellent device performance [62]. QDIPs using an undoped active region of InAs quantum dots grown directly on AlGaAs with a GaAs cap layer have

also demonstrated promising device performance [17]. Other III-V material systems under investigation include InAs/InAlAs QDs grown on InP [63] and InGaAs/InGaP QDs grown on GaAs [64]. Intraband transitions in the IR have also been observed for Ge/Si QDs [65-67]. These boron-doped devices use holes trapped in the QD valence band for the photogeneration of carriers, and the characterization of Ge/Si QDIPs indicates they may be an important technology for far IR detection (> 14 μm) [68-70]

Excellent device performance has been obtained using 70 layers of InAs/GaAs QDs to increase the absorption, and thereby the quantum efficiency, in QDIPs. GaAs barriers (≥ 50 nm) are required to reduce dislocation propagation in the large QD stack and to contribute to extremely low dark currents in the 70-layer QDIP. A schematic diagram of the device conduction band profile under bias is shown in Figure 10. The 70-layer heterostructure, shown in Figure 9 was grown by MBE and fabricated into devices at the University of Michigan Solid State Electronics Laboratory in Ann Arbor. Two, different 70-layer QDIP samples were fabricated (using the same growth and processing techniques) and characterized for high-temperature performance. One QDIP was measured at the U.S. Army Research Laboratory (ARL) in Adelphi, MD, and another was measured at the NASA Jet Propulsion Laboratory (JPL) in Pasadena, CA. The remainder of this chapter discusses the measured, QDIP performance, demonstrating that similar results are obtained repeatedly from these devices and that they are suitable for incorporation into FPAs operating at 150 K.

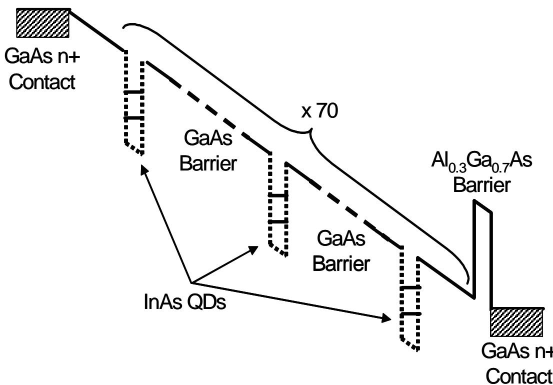
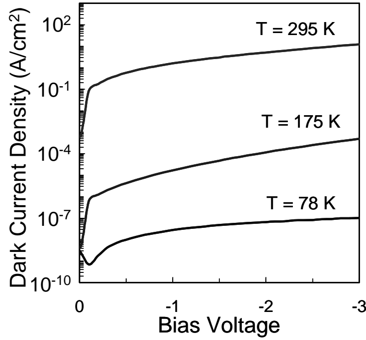


Figure 10. Schematic diagram of the conduction band profile under bias in the 70-layer, InAs/GaAs QDIP

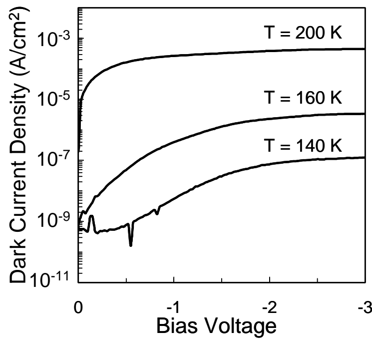
4.2 Dark Current

Bias-dependent dark current densities for the 70-layer, InAs/GaAs QDIP measured at ARL are shown in Figure 11(a) for different temperatures, exhibiting very low values (J_{dark} = 2.8×10^{-8} A/cm^2, T = 78 K, V_{bias} = -1.0 V). Even at 175 K, the dark current density was as low as 1.83×10^{-2} A/cm^2 at a bias voltage of 2.0 V. Similar

results were also obtained for the dark current density in the 70-layer QDIP sample measured at JPL. Fig. 11(b) shows the dark current densities for T = 140 K, 160 K, and 200 K, also demonstrating extremely low dark currents at elevated temperatures ($J_{dark} = 2.3\times10^{-6}$ A/cm^2, V_{bias} = -2.0 V, T = 160 K).



(a)



(b)

Figure 11. Measured dark current densities as a function of reverse bias voltage for different detector temperatures in the 70-layer, InAs/GaAs QDIP characterized at (**a**) ARL and (**b**) JPL

4.3 Mid-infrared Spectral Response in QDIPs

The MIR spectral responses of InAs/GaAs QDIPs have been investigated thoroughly for detectors featuring 10-70 layers of QDs (T = 78 K, V_{bias} = –1.0 V to 0.25 V). It is important to note that the spectral characteristics of the InAs/GaAs QDIPs featuring an $Al_{0.3}Ga_{0.7}As$ current-blocking barrier change with bias. The variation of λ_{peak} and $\Delta\lambda/\lambda$ with applied bias shown in Figure 12(a) and (b) were measured in a 10-layer MIR QDIP. However, this behavior is typical for quantum dot detectors in this wavelength range, including those using 70 QD layers in the active region. As the bias became less negative, the peak response shifted to shorter wavelengths and $\Delta\lambda/\lambda$ decreased. The blue-shift of the peak wavelength was expected due to

decreased band bending in the conduction band as the bias voltage neared 0 V. As the device approached flat-band operation, only very high-energy, photoexcited carriers in the upper excited levels of the QD could tunnel through the AlGaAs barrier in order to be collected as photocurrent. The decrease in $\Delta\lambda/\lambda$ with increasingly positive bias was not a function of device operation, but rather resulted from the strong atmospheric absorption that occurs near 3 μm. This result stresses the importance of designing a QDIP to respond in the proper atmospheric windows. It is likely that a response below 3 μm was actually the result of interband transitions between the conduction and valence bands of the InAs/GaAs QDs.

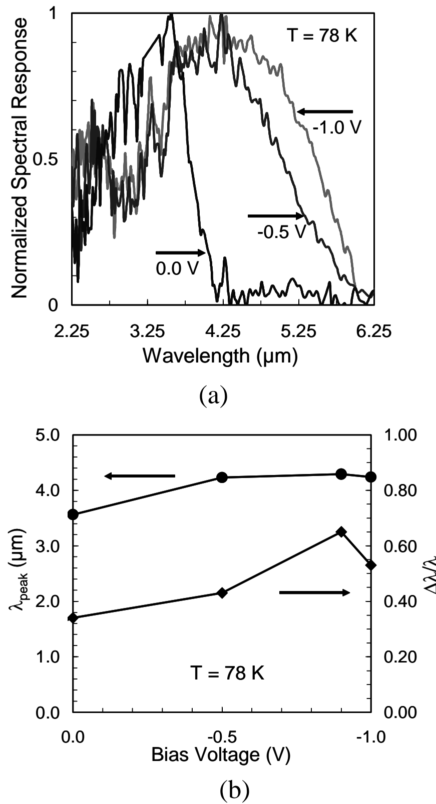


Figure 12. Bias voltage dependence of (**a**) the normalized spectral response and (**b**) the peak wavelength, λ_{peak}, and spectral linewidth, $\Delta\lambda/\lambda$, at T = 78 K in a MWIR QDIP

4.4 State-of-the-art Performance in MIR QDIPs at High Operating Temperatures

The inclusion of 70 QD layers in the InAs/GaAs QDIP heterostructure improves the conversion efficiency of the detector. For measurements conducted at ARL, at a temperature as high as 175 K, the conversion efficiency was 3.2 % for a bias voltage

of 2.0 V. In contrast, 10-layer QDIPs featuring the same AlGaAs barrier demonstrated a conversion efficiency of 0.033% for a bias voltage of 0.2 V at T = 100 K. In the 70-layer QDIP measured at ARL, for V_{bias} = 2.0 V, the responsivity was 0.12 A/W, which is almost 100 times greater than the responsivity obtained in a 10-layer QDIP (R_{peak} = 0.99 mA/W, V_{bias} = 0.2 V, T = 100 K). The maximum detectivity obtained for the 70-layer QDIP was 4.15×10^7 cmHz$^{1/2}$/W for V_{bias} = 2.0 V and T = 175 K, which is very large for such high temperature operation. This large detectivity was obtained due to the reduction of dark current by the optimization of the GaAs barrier layer thickness in the 70-layer QDIP heterostructure.

The spectral responsivity was also measured in the 70-layer QDIP characterized at JPL. As shown in Figure 13, the peak responsivity is large at temperatures as high as T = 150 K, peaking at 0.34 A/W for λ_{peak} = 3.65 μm at V_{bias} = 4.0 V. Even at a more reasonable bias of 2.0 V, this device yields a large responsivity close to that in the 70-layer QDIP characterized at ARL.

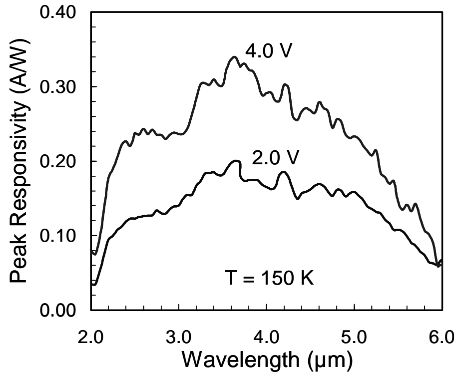


Figure 13. Measured spectral peak responsivity for different bias voltages at T = 150 K in a 70-layer, InAs/GaAs QDIP characterized at JPL

Using the measured dark current and spectral responsivity obtained at JPL, the peak specific detectivity was determined for the 70-layer, InAs/GaAs QDIP. The estimated peak detectivity for the 70-layer QDIP, shown in Table 1 as a function of temperature, is ~ 10^{11} cmHz$^{1/2}$/W at T = 100 K and ~ 6×10^9 cmHz$^{1/2}$/W at 200 K. These expected detectivities indicate that the 70-layer QDIP will provide high-temperature performance, currently unattainable in IR photodetectors, due to the simultaneous reduction of dark current by the GaAs barrier and increase of responsivity by the 70-layer absorption region.

It is important to note that the maximum, peak detectivities in the 70-layer QDIPs characterized at ARL and JPL differ by a factor of 10^3. This is due in large part to the different measurement techniques used for the two devices. While detector noise was measured directly in the ARL sample, dark current and spectral responsivity measurements were used to estimate the detectivity in the JPL sample, as noted above. Despite these differences, the most salient point is that the measured dark currents and peak responsivities are very similar for the ARL and JPL samples.

These figures of merit are fundamental in determining the overall performance of the QDIP devices.

Table.1. Peak specific detectivity in 70-layer, InAs/GaAs QDIP as a function of temperature (characterized at JPL)

Temperature (K)	Peak detectivity ($cmHz^{1/2}/W$)
100	1.25E+11
120	1.08E+11
140	5.33E+10
150	3.67E+10
160	2.92E+10
180	1.30E+10
200	6.53E+09

Table. 2. Summary of figures of merit in different QDIP designs

QDIP design	Operating temperature	Operating bias	Figure of merit	Value
70-layer QDIP (ARL)	175 K	2.0 V	J_{dark}	1.83×10^{-2} A/cm^2
			R_{peak}	0.120 A/W
			D^*_{peak}	4.15×10^{7} $cmHz^{1/2}/W$
70-layer QDIP (JPL)	150 K	2.0 V	J_{dark}	5.80×10^{-5} A/cm^2
			R_{peak}	0.201 A/W
			D^*_{peak} (estimated)	3.7×10^{10} $cmHz^{1/2}/W$

Therefore, by carefully designing the 70-layer, InAs/GaAs QDIP heterostructure to minimize dark current and maximize absorption, this device provided low dark current and relatively high responsivity at elevated operating temperatures. This performance has been verified by measurements conducted at two different laboratories. By investigating the 70-layer QDIP design, performance benchmarks necessary for FPA fabrication have been realized. The best results for MIR QDIPs measured at ARL and JPL are summarized in Table 2.

The dark current and spectral responsivity measurements conducted at JPL have confirmed that the 70-layer QDIP heterostructure is suitable for incorporation into FPAs. Table 3 shows the desired QDIP array performance for FPA fabrication at

JPL using a (350×256) Si ROIC manufactured by Indigo Systems Inc. for operation in the MIR range. Of particular importance are the ROIC bias requirements (±2.0 V), charge well capacity (18×10^6 electrons), and operating temperature (> 80 K), which are quantities that significantly affect how well the QDIP array can be hybridized in order to make an FPA.

Table. 3. QDIP array performance requirements for FPA fabrication at NASA JPL

QDIP array figure of merit at operating bias	Value
Dark current density	~ 10^{-5} A/cm^2
Peak responsivity	~ 0.100 A/W
Peak specific detectivity	~ 10^{12} cmHz$^{1/2}$/W

In comparing the measured and estimated figures of merit for the 70-layer, InAs/GaAs QDIP characterized at ARL and JPL with the required performance parameters listed in Table 3, it is important to note that the dark current and peak responsivity requirements have been met. In the case of the peak detectivity, the largest QDIP value obtained (~ 10^{11} cmHz$^{1/2}$/W in the JPL device) is still an order of magnitude below the desired value. However, this discrepancy does not preclude the fabrication and testing of a MIR QDIP FPA designed to operate at 150 K, which is an important goal to determine performance characteristics of an entire, large-scale device array.

5 Conclusions

In(Ga,Al)As/(Ga,Al)As QDIPs are positioned to become an important technology in the field of IR detection, particularly for the type of high-temperature, low-cost, high-yield detector arrays required for the military and other highly sophisticated applications. MIR detection at 150 K has been demonstrated using InAs/GaAs QDIPs, thereby advancing the state-of-the-art in QDIP performance and demonstrating that it is reasonable to incorporate these devices into an IR camera system. QDIPs have the potential to fulfil a niche market in high-temperature MIR detection. This development is significant in that it has been difficult for GaAs/AlGaAs QWIPs to provide detector performance comparable to the devices reported here at the same wavelengths. In fact, InGaAs/AlGaAs QWs are typically used to obtain MIR detection in multi-color QWIPs [71,72]. It is important to appreciate that the QDIP technology is still relatively new, and the tremendous strides that have been made in QDIP performance over the past five years by research groups around the world are only the beginning.

Appendix. Infrared Photodetector Figures of Merit

A1 Photocurrent

The photocurrent in a PC detector is determined by the number of photo-generated charge carriers that travel in the device per unit time. Thus, given the incident photon flux on the detector, (ϕ_{ph}cm^{-2}s^{-1}), the resulting photocurrent is

$$I_{ph} = qA_{opt}\phi\eta_{quant}g \tag{A1}$$

where q is the electron charge, A_{opt} is the optical area of the detector, η_{quant} is the quantum efficiency, and g is the photoconductive gain. The efficiency and gain determine the number of electrons that are photo-generated by incident photons.

A2 Dark Current

Dark current is defined as the current flowing through the detector when there is no incident photon flux, commonly referred to as the zero-field-of-view condition. The dark current is an important parameter in photodetector characterization because it determines the base current level that photocurrent must exceed in order to be detected. Possible sources of dark current in QDIPs include thermal generation of carriers in the three-dimensional QD barrier layers, thermionic emission from QDs, field-assisted tunneling emission from QDs, impact-ionization in the active region, surface leakage current, and ohmic leakage current. Of these various sources, the most significant for IR detectors is thermionic emission, which depends on the detector temperature. Therefore, by operating at cryogenic temperatures (~ 78 K), the dark current can be reduced sufficiently such that photocurrent measurements can be conducted. Thus, the dark current is often the limiting factor in IR photodetector performance since it determines the highest possible operating temperature of the device.

A3 Noise Current

The noise current is another important parameter that must be considered when characterizing QDIPs. While noise current exists under both dark and illuminated conditions, it is treated differently for dark current and photocurrent. Since the dark current is defined as the total current in the device when there is no light source present, the contribution of the noise current is included. However, the sensitivity and efficiency of a detector are determined by how well photocurrent can be distinguished from noise current. This is represented quantitatively by the signal-to-noise ratio (SNR). Therefore, noise current must be accounted for independently of photocurrent during detector characterization.

The total noise current is the summation of several different noise sources, both external and internal to the detector. External noise sources include photon noise, amplifier noise, and microphonic noise. Internal noise sources include Johnson noise, shot noise, generation-recombination noise, 1/f noise, and

temperature fluctuation noise. Of the types of noise listed, the most fundamental, or the ones which cannot be controlled through experimental procedures or material processing quality, are photon noise, Johnson noise, and generation-recombination noise.

A4 Normalized Spectral Response and Peak Responsivity

The spectral responsivity, $R(\lambda)$, is a quantity that accounts for the strength of the detector signal with respect to the monochromatic incident light power, having units of A/W or V/W depending on the type of photodetector. Thus, in a photoconductor, the spectral responsivity is the ratio of the detector photocurrent to the incident photon power as a function of wavelength:

$$R(\lambda) = \frac{I_{ph}}{P_{inc}(\lambda)} \tag{A2}$$

In practice, a normalized spectral response is defined in which the spectral shape of the detector response and the value of the responsivity at the peak wavelength are taken into account. This normalized spectral response is actually the spectral responsivity scaled by the value of the responsivity at the peak wavelength:

$$\text{Normalized spectral response} = \frac{R(\lambda)}{R(\lambda_{peak})} \tag{A3}$$

The value of the spectral responsivity at λ_{peak} is called the peak responsivity, R_{peak}, and it is an important figure of merit for characterizing the signal strength in photodetectors, usually as a function of bias voltage.

A5 Peak Specific Detectivity

As discussed earlier, the signal-to-noise ratio (SNR) determines how well the photocurrent can be distinguished from the noise current. In fact, the SNR is one of the most important factors in determining the performance and applicability of a given photodetector. Several figures of merit exist for the characterization of the SNR. The noise-equivalent power (*NEP*), with units of watts, is the incident photon power necessary to yield a SNR = 1. As such, the *NEP* depends on the spectral responsivity and the noise current:

$$NEP(\lambda) = \frac{i_{noise}}{R(\lambda)} \tag{A4}$$

where i_{noise} is the root-mean-square (rms) noise current. Given this equation, larger values of *NEP* correspond to smaller values of the SNR. Because this relationship is somewhat counter-intuitive, another figure of merit, called the detectivity (D), with units of W^{-1}, is used

$$D(\lambda) = \frac{1}{NEP(\lambda)} = \frac{R(\lambda)}{i_{noise}} = \frac{\frac{q\lambda}{hc}\eta_{quant}\, g}{i_{noise}} \tag{A5}$$

As demonstrated by this equation, the detectivity depends on several detector-specific quantities, such as the quantum efficiency, photoconductive gain, and noise current. These values, in turn, are especially sensitive to the optical area of the detector, A_{opt}, and the noise-equivalent electrical bandwidth, Δf, which is a property of the measurement technique used. Therefore, the specific detectivity, D^*, is a normalized detectivity (with units of $\mathrm{cmHz^{1/2}/W}$ - known as Jones) that enables the comparison of different photodetectors. The equation for D^* is

$$D^*(\lambda) = \frac{R(\lambda)\sqrt{A_{opt}\Delta f}}{i_{noise}} \tag{A6}$$

When $\lambda = \lambda_{peak}$, this value is called the peak specific detectivity, or the peak detectivity, D^*_{peak}.

A6 Photoconductive Gain and Quantum Efficiency

The photoconductive gain quantifies the fact that a single photo-generated carrier may generate, in turn, several additional carriers that contribute to the device photocurrent (without absorbing additional photons). The photoconductive gain can be defined as the ratio of total collected carriers to total excited carriers, whether these carriers are thermally- or photo-generated. In photodiodes, the gain equals one. However, in photoconductors, the gain may be greater than one if the carrier lifetime exceeds the carrier transit time through the device

$$g = \frac{\tau_e}{\tau_{tr}} \tag{A7}$$

where τ_e is the electron lifetime and τ_{tr} is the transit time for an electron to travel across the active region. The physical mechanism responsible for photoconductive gain in QDIPs is the maintenance of charge neutrality in the QD active region. When IR light is absorbed by the QDs, electrons are excited from the dots to the continuum states, and these electrons drift towards the contact layer. The electrons may be captured by subsequent QDs in transit, or they may reach the contact. In the latter case, an electron is injected from the opposite contact to maintain the neutral charge of QDs, and this additional carrier contributes to the photocurrent in the device. This process continues until the electron is captured by a QD. Thus, due to the long electron effective lifetime resulting from the phonon bottleneck in QDs, the gain in QDIPs can be quite large. The gain in InAs/GaAs QDIPs has typical values in the range 1-5 [27].

The photoconductive gain significantly affects the detector efficiency, which correlates the number of generated carriers to the number of incident photons. The conversion efficiency, η_{conv}, which can be determined directly from the peak

responsivity and the peak wavelength, is defined as the ratio of collected, photoexcited carriers to incident photons

$$\eta_{conv} = \frac{hcR_{peak}}{q\lambda_{peak}} \tag{A8}$$

Therefore, if the photoconductive gain is known, then the quantum efficiency, η_{quant}, or the ratio of total photoexcited carriers to incident photons, can be calculated by

$$\eta_{quant} = \frac{\eta_{conv}}{g} = \eta_{\text{int}}\left(1 - \exp\left(-\alpha l\right)\right)\left(1 - \Theta_R\right) \tag{A9}$$

where η_{int} is the internal quantum efficiency accounting for unique material properties (~ 1 in InAs/GaAs QDIPs for all practical purposes), α is the absorption coefficient, l is the absorption length, and Θ_R is the detector surface reflection coefficient. Thus, it is clear that the quantum efficiency is significantly affected by the absorption coefficient for QDs.

Acknowledgements

Collaborations with S. Gunapala, S. Bandara, and S. B. Rafol at the Jet Propulsion Laboratory in Pasadena, CA and S. W. Kennerly at the U. S. Army Research Laboratory in Adelphi, MD are gratefully acknowledged. This work is supported by the Army Research Office (MURI program) under Grant DAAD19-01-1-0462, as well as DARPA under Grant DAAD19-00-1-0394.

References

1. Piotrowski J and Gawron W. Ultimate performance of infrared photodetectors and figure of merit of detector material. Infrared Physics & Technology 1997; 38: 63-8
2. Norton PR. Status of infrared detectors. SPIE Proceedings 1998; 3379: 102-14
3. Rogalski A. Assessment of HgCdTe photodiodes and quantum well infrared photoconductors for long wavelength focal plane arrays. Infrared Physics & Technology 1999; 40: 279-94
4. Bajaj J. State-of-the-art HgCdTe infrared devices. SPIE Proceedings 2000; 3948: 42-54
5. Reine MB. Review of HgCdTe photodiodes for IR detection. SPIE Proceedings 2000; 4028: 320-30
6. Baker IM and Maxey CD. Summary of HgCdTe 2D array technology in the UK. Journal of Electronic Materials 2001; 30: 682-9
7. Varesi JB, Bornfreund RE, Childs AC *et al.*. Fabrication of high-performance large-format MWIR focal plane arrays from MBE-grown HgCdTe on 4" silicon substrates. Journal of Electronic Materials 2001; 30: 566-73
8. Levine BF. Quantum-well infrared photodetectors. Journal of Applied Physics 1993; 74: R1-R81
9. Gunapala SD and Bandara KMSV: Homojunction and Quantum-Well Infrared Detectors. In: MH Francombe, JL Vossen (ed). Academic Press, San Diego, 1995, pp.113-237

10. Gunapala SD, Bandara SV, Singh A *et al.*. 640x 486 long-wavelength two-color GaAs/AlGaAs quantum well infrared photodetector (QWIP) focal plane array camera. IEEE Transactions on Electron Devices 2000; 47: 963-71
11. Pan JL and Jr. CGF. Theory, fabrication, and characterization of quantum well infrared photodetectors. Material Science and Engineering. R, Reports: a review journal 2000; 28: 65-147
12. Tidrow MZ. Device physics and state-of-the-art of quantum well infrared photodetectors and arrays. Material Science and Engineering B 2000; 74: 45-51
13. Phillips J, Bhattacharya P, Kennerly SW, Beekman DW and Dutta M. Self-assembled InAs-GaAs quantum-dot intersubband detectors. IEEE Journal of Quantum Electronics 1999; 35: 936-43
14. Liu HC, Gao M, McCafferey J, Wasilewski ZR and Fafard S. Quantum dot infrared photodetectors. Applied Physics Letters 2001; 78: 79-81
15. Tang S-F, Lin S-Y and Lee S-C. Near-room-temperature operation of an InAs/GaAs quantum-dot infrared photodetector. Applied Physics Letters 2001; 78: 2428-30
16. Wang SY, Lin SD, Wu HW and Lee CP. Low dark current quantum-dot infrared photodetectors with an AlGaAs current blocking layer. Applied Physics Letters 2001; 78: 1023-5
17. Chen Z, Baklenov O, Kim ET *et al.*. Normal incidence InAs/AlxGa1-xAs quantum dot infrared photodetectors with undoped active region. Journal of Applied Physics 2001; 89: 4558-63
18. Kang YH, Park J, Lee UH and Hong S. Effect of the dot size distribution on quantum dot infrared photoresponse and temperature-dependent dark current. Applied Physics Letters 2003; 82: 1099-101
19. Krishna S, Raghavan S, Winckel Gv *et al.*. Two color InAs/InGaAs dots-in-a-well detector with background-limited performance at 91 K. Applied Physics Letters 2003; 82: 2574-6
20. Ryzhii V. The theory of quantum-dot infrared phototransistors. Semiconductor Science and Technology 1996; 11: 759-65
21. Berryman KW, Lyon SA and Segev M. Mid-infrared photoconductivity in InAs quantum dots. Applied Physics Letters 1997; 70: 1861-3
22. Maimon S, Finkman E, Bahir G *et al.*. Intersublevel transitions in InAs/GaAs quantum dots infrared photodetectors. Applied Physics Letters 1998; 73: 2003-5
23. Phillips J, Kamath K and Bhattacharya P. Far-infrared photoconductivity in self-organized InAs quantum dots. Applied Physics Letters 1998; 72: 2020-2
24. Pan D, Towe E and Kennerly S. Normal-incidence intersubband (In, Ga)As/GaAs quantum dot infrared photodetectors. Applied Physics Letters 1998; 73: 1937-9
25. Xu SJ, Chua SJ, Mei T *et al.*. Characteristics of InGaAs quantum dot infrared photodetectors. Applied Physics Letters 1998; 73: 3153-5
26. Pan D, Towe E and Kennerly S. Photovoltaic quantum-dot infrared detectors. Applied Physics Letters 2000; 76: 3301-3
27. Stiff AD, Krishna S, Bhattacharya P and Kennerly S. Normal-incidence, high-temperature, mid-infrared InAs-GaAs vertical quantum-dot infrared photodetector. IEEE Journal of Quantum Electronics 2001; 37: 1412-9
28. Krishna S, Stiff-Roberts AD, Phillips JD, Bhattacharya P and Kennerly SW. Hot dot detectors: infrared quantum dot intersubband photodetectors are a promising technology for multiwavelength IR detection. IEEE Circuits & Devices 2002; 18: 14-24
29. Dereniak EL and Boreman GD. Infrared Detectors and Systems. Wiley, New York, 1996
30. Jiang H and Singh J. Strain distribution and electronic spectra of InAs/GaAs self-assembled dots: an eight-band study. Physical Review B 1997; 56: 4696-701
31. Keating PN. Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure. Physical Review 1966; 145: 637-45

32. Martin RM. Elastic properties of ZnS structure semiconductors. Physical Review B 1970; 1: 4005-11
33. Horiguchi N, Futatsugi T, Nakata Y *et al.*. Quantum Dot Infrared Photodetector Using Modulation Doped InAs Self-Assembled Quantum Dots. Japanese Journal of Applied Physics 1999; 38: 2559-61
34. Phillips J. Evaluation of the fundamental properties of quantum dot infrared detectors. Journal of Applied Physics 2002; 91: 4590-4
35. Singh J. Electronic and Optoelectronic Properties of Semiconductor Structures. Cambridge University Press, New York, 2003
36. Bockelmann U and Bastard G. Phonon scattering and energy relaxation in two-, one-, and zero-dimensional electron gases. Physical Review B 1990; 42: 8947-51
37. Benisty H, Sotomayor-Torres CM and Weisbuch C. Intrinsic mechanism for the poor luminescence properties of quantum-box systems. Physical Review B 1991; 44: 10945-8
38. Bhattacharya P, Kamath KK, Singh J *et al.*. In(Ga)As/GaAs self-organized quantum dot lasers: DC and small-signal modulation properties. IEEE Transactions on Electron Devices 1999; 46: 871-83
39. Ohnesorge B, Albrecht M, Oshinowo J and Forchel A. Rapid carrier relaxation in self-assembled $In_xGa_{1-x}As$/GaAs quantum dots. Physical Review B 1996; 54: 11532-8
40. Mukai K, Ohtsuka N, Shoji H and Sugawara M. Phonon bottleneck in self-formed $In_xGa_{1-x}As$/GaAs quantum dots by electroluminescence and time-resolved photoluminescence. Physical Review B 1996; 54: R5243-R6
41. Klotzkin D, Kamath K and Bhattacharya P. Quantum capture times at room temperature in high-speed $In_{0.4}Ga_{0.6}As$-GaAs self-organised quantum-dot lasers. IEEE Photonics Technology Letters 1997; 9: 1301-3
42. Urayama J, Norris TB, Singh J and Bhattacharya P. Observation of phonon bottleneck in quantum dot electronic relaxation. Physical review Letters 2001; 86: 4930-3
43. Klotzkin D and Bhattacharya P. Temperature dependence of dynamic and DC characteristics of quantum-well and quantum-dot lasers: a comparative study. Journal of Lightwave Technology 1999; 17: 1634-42
44. Kochman B, Stiff-Roberts AD, Chakrabarti S *et al.*. Absorption, carrier lifetime, and gain in InAs-GaAs quantum-dot infrared photodetectors. IEEE Journal of Quantum Electronics 2003; 39: 459-67
45. Parker EHC, ed. 1985. New York: Plenum Press
46. Wang PD, Torres CMS, Benisty H, Weisbuch C and Beaumont SP. Radiative recombination in GaAs-$Al_xGa_{1-x}As$ quantum dots. Applied Physics Letters 1992; 61: 946-8
47. Steffen R, Koch T, Oshinowo J, Faller F and Forchel A. Photoluminescence study of deep etched InGaAs/GaAs quantum wires and dots defined by low-voltage electron beam lithography. Applied Physics Letters 1996; 68: 223-5
48. Ueno H, Moriyasu K, Wada Y *et al.*. Conductance through laterally coupled quantum dots. Japanese Journal of Applied Physics 1999; 38: 332-5
49. Goldstein L, Glas F, Marzin JY, Charasse MN and LeRoux G. Growth by molecular beam epitaxy and characterization of InAs/GaAs strained-layer superlattices. Applied Physics Letters 1985; 47: 1099-101
50. Fujita S, Matsuda Y and Sasaki A. Blue luminescence of a ZnSe-$ZnS_{0.1}Se_{0.9}$ strained-layer superlattice on a GaAs substrate grown by low=pressure organometallic vapor phase epitaxy. Applied Physics Letters 1985; 47: 955-7
51. Berger PR, Chang K, Bhattacharya PK and Singh J. A study of strain-related effects in the molecular-beam epitaxy growth of $In_xGa_{1-x}As$ on GaAs using reflection high-energy electron diffraction. Journal of Vacuum Science and Technology B 1987; 5: 1162-6
52. Tabuchi M, Noda S and Sasaki A. Strain energy and critical thickness of heteroepitaxial InGaAs layers on GaAs substrate. Journal of Crystal Growth 1991; 115: 169-73

53. Leonard D, Krishnamurthy M, Reaves CM, Denbaars SP and Petroff PM. Direct formation of quantum-sized dots from uniform coherent islands of InGaAs on GaAs surfaces. Applied Physics Letters 1993; 63: 3203-5
54. Xie Q, Chen P, Kalburge A *et al.*. Realization of optically active strained InAs island quantum boxes on GaAs (100) via molecular beam epitaxy and the role of island induced strain fields. Journal of Crystal Growth 1995; 150: 357-63
55. Bimberg D, Grundmann M and Ledentsov NN. Growth, spectroscopy, and laser application of self-ordered III-V quantum dots. Materials Research Society Bulletin 1998; 23: 31-4
56. Bhattacharya P, Kamath K, Phillips J and Klotzkin D. Self-organized growth of In(Ga)As/GaAs quantum dots and their opto-electronic device applications. Bulletin of Material Science 1999; 22: 519-29
57. Stranski IN and Krastanow L. Theory of orientation separation of ionic crystals. Sitzungsberichte 1938; Abteilung IIb. 146: 797-810
58. Bimberg D, Grundmann M and Ledentsov NN. Quantum Dot Heterostructures. John Wiley & Sons, Chichester, 1999
59. Caniou J. Passive Infrared Detection: Theory and Applications. Kluwer Academic Publishers, Boston, 1999
60. Rogalski A. Infrared Detectors. Gordon and Breach Science Publishers, Singapore, 2000
61. Kim E-T, Chen Z and Madhukar A. Tailoring detection bands of InAs quantum-dot infrared photodetectors using InxGa1-xAs strain-relieving quantum wells. Applied Physics Letters 2001; 79: 3341-3
62. Raghavan S, Rotella P, Stintz A *et al.*. High-responsivity, normal-incidence long-wave infrared (l=7.2mm) $InAs/In_{0.15}Ga_{0.85}As$ dots-in-a-well detector. Applied Physics Letters 2002; 81: 1369-71
63. Weber A, Gauthier-Lafaye O, Julien FH *et al.*. Strong normal-incidence infrared absorption in self-organized InAs/InAlAs quantum dots grown on InP(001). Applied Physics Letters 1999; 74: 413-5
64. Kim S, Mohseni H, Erdtmann M *et al.*. Growth and characterization of InGaAs/InGaP quantum dots for mid-infrared photoconductive detector. Applied Physics Letters 1998; 73: 963-5
65. Liu JL, Wu WG, Balandin A, Jin GL and Wang KL. Intersubband absorption in boron-doped multiple Ge quantum dots. Applied Physics Letters 1999; 74: 185-7
66. Boucaud P, Thanh VL, Sauvage S, DéBarre D and Bouchier D. Intraband absorption in Ge/Si self-assembled quantum dots. Applied Physics Letters 1999; 74: 401-3
67. Liu JL, Wu WG, Balandin A *et al.*. Observation of inter-sub-level transitions in modulation-doped Ge quantum dots. Applied Physics Letters 1999; 75: 1745-7
68. Yakimov AI, Dvurechenskii AV, Proskuryakov YY *et al.*. Normal-incidence infrared photoconductivity in Si p-i-n diode with embedded Ge self-assembled quantum dot. Applied Physics Letters 1999; 75: 1413-5
69. Miesner C, Röthig O, Brunner K and Abstreiter G. Mid-infrared photocurrent measurements on self-assembled Ge dots in Si. Physica E 2000; 7: 146-50
70. Yakimov AI, Dvurechenskii AV, Nikiforov AI and Proskuryakov YY. Interlevel Ge/Si quantum dot infrared photodetector. Journal of Applied Physics 2001; 89: 5676-81
71. Jiang X, Li SS and Tidrow MZ. Investigation of a multistack voltage-tunable four-color quantum-well infrared photodetector for mid- and long-wavelength infrared detection. IEEE Journal of Quantum Electronics 1999; 35: 1685-92
72. Li SS, Kim S-H, Moon J-H and Lee JH. A two-stack, multi-color quantum well infrared photodetector for mid- and long-wavelength infrared detection. Infrared Physics & Technology 2003; 44: 235-41

Quantum Photovoltaic Devices Based on Antimony Compound Semiconductors

Y. Wei, A. Gin and M. Razeghi*
Center for Quantum Devices, Electrical and Computer Engineering Department, Northwestern University, Evanston, IL 60208, USA

1 Introduction

Since the discovery of infrared light in 1800 [1,2] detection techniques have steadily improved with the development of science and technology. The idea of Type II InAs/GaSb superlattices was first introduced by Sai-Halasz and Esaki in the 1970s [3,4,5]. The concept of type II band alignment between InAs and GaSb has introduced new opportunities for bandstructure engineering. Calculations using the empirical tight-binding method (ETBM) show that the bandgap of these superlattices can be tuned between zero and 0.6 eV. This enables one to envisage multicolor infrared detectors. The bandstructure can also be tuned using different specific layer thicknesses, which was not possible with previous infrared technologies. This capability allowed designers to choose a bandstructure that would lead to lower Auger recombination and thus higher operating temperatures. High quality uncooled infrared detectors in the long wavelength infrared range (LWIR, 7-14 μm) might therefore become possible. The use of III-V materials leads to a much more stable system than the HgCdTe technology, especially in the very long wavelength infrared range (VLWIR, >14 μm). The uniformity of the Type II superlattices in this wavelength range is also unmatched compared with HgCdTe. In addition, the superlattices have direct bandgaps, which enables normal incident optical absorption, unlike the QWIP (quantum well infrared photodetector) structures. The merits of the Type II InAs/GaSb superlattices have brought much practical attention. However, high-quality epitaxial growth of this type of material has not been available until recently with the advent of state-of-the-art molecular-beam epitaxy (MBE) technology. The material growth has several major challenges, including: the GaSb substrate quality, the interface manipulation, superlattice mismatch reduction, *etc.* We shall review the progress that has been made to grow high-quality materials and devices for infrared detectors.

With the aid of a modern electron beam lithography system, one can create nanometer scale features in Type II superlattice material. The idea is to confine the carriers in the lateral direction leading to quasi quantum dot detectors. Once realized, Auger recombination may be significantly reduced as well, due to the formation of quasi-discrete energy levels. The application of a sidewall gate voltage would also enhance the lateral confinement and lead to an electrically tunable bandgap. Compared to self-assembled quantum dots, the implementation of e-beam technology can make regular arrays of quantum dots that have very low inhomogeneity, which is one of the key issues in controlling spectral broadening.

---

* Email: Razeghi@ece.northwestern.edu

The infrared focal plane array (IR-FPA) is one of the most important goals for development in an infrared material system. Existing technologies have been established within the industry for InGaAs, InSb, HgCdTe, QWIP, *etc.* Many processing techniques became proprietary and not available to the general public. With university-based equipment, we have pioneered FPA development for InAs/GaSb superlattices. Although it is still under active research, we will cover this topic up to the present state-of-the-art. There are several schemes for multicolor FPAs. The first is to use a sequential diode structure with a shorter cutoff wavelength diode grown before a longer cutoff wavelength one in the backside illumination scheme. This has been successfully implemented in the HgCdTe and QWIP FPAs. The second scheme is to use a nanopillar structure with different diameters. In this scheme, it is important to control the diameter of the nanopillars to be small enough so that the effect of quantum confinement is significant. The third one is to use a gated-nanometer scheme. We have initiated research using these concepts and the topics will be covered in later sections.

2 Theoretical Modeling

2.1 The Type-II InAs/GaSb Superlattice

In order to guide the design of the superlattice, it is important to have an effective model that can predict part, or most of, the important parameters. In the field of semiconductor engineering the **k.p** model has been widely used and has become practically the standard model. However, in order to model the Type-II InAs/GaSb superlattices, the **k.p** model has an intrinsic difficulty in dealing with interfaces and short-period superlattices. At the interface between the InAs and GaSb there is no common cation or common anion. This leads to different types of possible interfaces: InSb, GaAs, or mixed $Ga_xIn_{1-x}AsSb$. Although there has been some success using the **k.p** model for some of the superlattices, more accurate models need to be developed. A good candidate is the empirical tight-binding method (ETBM).

The empirical tight-binding method originates from the early work of J. C. Slater and G. F. Koster in 1954 [6]. The method was originally called Linear Combinations of Atomic Orbitals (LCAO). The non-orthogonality of the atomic orbitals posed a difficulty if one applies this method in an empirical frame. With the modification of the atomic orbitals to Löwdin orbitals [7], this problem was solved. However we assume that these orbitals still have the same symmetry as their corresponding atomic orbitals. The basis orbitals in this work should be considered only to be the Löwdin orbitals. As a balance between the modeling accuracy and the calculation load, we chose a basis of sp^3s^*, with nearest neighbor interactions, under the two-center approximation [8,9]. s* orbitals (the first excited s orbital) were used to better describe the conduction bands. It can be shown that a full superlattice Hamiltonian matrix cannot be block diagonalized into two equivalent parts for a general k-vector once the spin-orbital interactions are considered [10]. Therefore we used a full set of Löwdin orbitals with both spin-up and spin-down states in this work.

There has been a significant amount of work on the theory of ETBM for both bulk III-V materials and the III-V superlattices. Here we only briefly describe the

formalism to keep the integrity of this work. Assuming the construction Lowdin orbitals for the material is

$$\varphi_n^{\alpha}(\vec{r}-\vec{R}-\vec{\tau}_n) \tag{1}$$

where n runs through all the atoms in a unit cell, τ_n are the coordinate position vectors of the atoms in the cell, α runs though all considered types of Löwdin orbitals (s, p_x, p_y, p_z, and s*), and R represents the coordinate position vector of the unit cell in the entire material we are considering. The Bloch wavefunction of electrons for the entire material is described by

$$\Psi_{\vec{k}}(\vec{r}) = \sum_{\vec{R}_{SL}}\sum_{\alpha}\sum_{n=1}^{N}\exp\left(i\vec{k}\cdot(\vec{R}_{SL}+\vec{\tau}_n)\right)A_n^{\alpha}\varphi_n^{\alpha}(\vec{r}-\vec{R}_{SL}-\vec{\tau}_n) \tag{2}$$

where N is the total number of atoms in one unit cell, A_n^{α} are constants, and $\vec{R}_{SL}$ runs through all the unit cells that are involved in the nearest neighbor interactions.

The bandstructure calculation problem now reduces to an eigenvalue problem for $\Psi_{\vec{k}}(\vec{r})$

$$H\Psi_{\vec{k}}(\vec{r}) = E\Psi_{\vec{k}}(\vec{r}) \tag{3}$$

where H is the Hamiltonian operator, and E is the energy eigenvalues that are dependent on $\vec{k}$ values in the first Brillouin zone. We define an ideal Type II superlattice to be $[(AsIn)_m\text{-}AsGa_{x1}In_{1\text{-}x1}\text{-}(SbGa)_n\text{-}SbGa_{x2}In_{1\text{-}x2}]_N$ with perfect crystal structure, where each atomic symbol represents a layer of the same kind of atom without any exotic species, and the sequence of atoms is exactly as written down. Here we distinguish the interface compositions for later use. The particular atomic order we choose is only for our convenience in describing the formalism we developed, in which we start with an anion atom.

The antimony for arsenic or arsenic for antimony substitution is a growth characteristic particular to Type II InAs/GaSb superlattices. Because of the much lower antimony vapor pressure than that for arsenic at a growth temperature below 400 °C, we will focus on the former case that dominates. This refers to the antimony segregation in the InAs layers. The amount is directly determined by the initial antimony seeds which exist on the surface when each GaSb layer just finishes. It also depends on the growth temperature. It has been reported that there exists indium and arsenic segregation in the GaSb layers [11,12]. However, these imperfections are highly dependent on the details of the growth method used. We have minimized their existence. The x-ray data showed exact superlattice periods after taking into account the antimony segregation effects. In addition, within our modeling, without taking into account the indium segregation effects, we can get a good fit to a variety of superlattices grown at the same temperature but with different material compositions and amounts of segregation. All these facts are consistent with our assumption that the indium and arsenic segregation in the GaSb layers for the superlattices we have grown is negligible.

The physical presence of antimony segregation in the InAs layers has been reported using scanning tunneling microscopy (STM) [13]. We will follow the same notation in this work. We used the same method for the Sb segregation profile,

$$x(n) = x_i R^{n-1}(1-R) + x_0(1-R^n)\text{, n=1,2,3,…, N}_{\text{InAs}} \tag{4}$$

where x(n) stands for the composition of Sb in the nth $InAs_{1-x(n)}Sb_{x(n)}$ layer, x_i stands for the initial seeds of Sb, R represents the phenomenological segregation coefficient, x_0 is the Sb background incorporation ratio, and N_{InAs} stands for the number of InAs layers. The actual superlattice that we are modeling becomes

$$[As_{1-x(1)}Sb_{x(1)}In\text{-}As_{1-x(2)}Sb_{x(2)}In\text{-}\ldots$$
$$\text{-}As_{1-x(m)}Sb_{x(m)}In\text{-}AsGa_{x1}In_{1-x1}\text{-}(SbGa)_n\text{-}SbGa_{x2}In_{1-x2}]_N \quad (5)$$

The superlattice Hamiltonian matrix is thus modified to reflect this layer dependent InAs composition using a composition weighted arithmetic average of the material parameters for InAs and InSb. For example, for the atomic section of "-$InAs_{x(n)}Sb_{1-x(n)}$-GaSb-", the self-interaction energy for the $As_{x(n)}Sb_{1-x(n)}$ layer would be $\{x(n)\, E_{InAs}+[1-x(n)]\, E_{InSb}+x(n)\, E_{GaAs}+[1-x(n)]\, E_{GaSb}\}/2$, *etc.* The E_i (i = InAs, *etc.*) represents the general form for the corresponding fitting parameters. Since there exists neither common anion nor cation between InAs and GaSb, it is important to single out and use a realistic model for the interface layers in the calculations. Because the superlattice layers, including the interfaces, are under high stress and strain (~7%), Harrison's d^{-2} rule [14] was also applied to scale the interaction energies. The modeling results have shown that this scheme is accurate enough to predict the characteristics of the superlattices we grow.

The ETBM parameters under sp^3s^* nearest neighbor approximation are fitted to known bulk band structure parameters of GaAs, InAs, GaSb, and InSb at 77 K. The bandgap, the spin-orbit splitting energy, and the effective masses at conduction band minimum and for heavy holes along the (001) direction, were fitted to the exact experimentally measured values [15]. Other band energies at high symmetry points were fitted using a numerical least squares method. The band alignment energies between different compositional materials were obtained from [16].

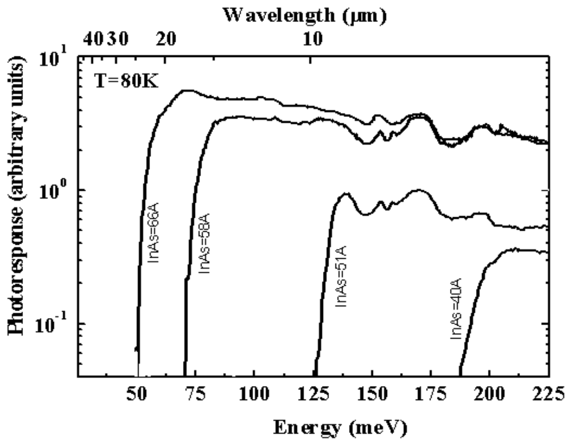


Figure 1. Photoresponse spectrum of a series of superlattices grown under the same condition with fixed GaSb layer thickness of 40 Å (~13 MLs) at a temperature of 80 K. The layer thickness of InAs changed from 40 Å (~13 MLs) up to 66 Å (~22 MLs). [From Y. Wei, and M. Razeghi, Phys. Rev. B v.69(8), 085316 (2004)].

We used the modified ETBM to calculate a series of superlattices we have grown. Details of the material growths have been reported elsewhere [3,4,5]. We fixed the GaSb layers at 40 Å (~13 monolayers (MLs)), and varied the thickness of

the InAs layers from 40 Å (~13 MLs) up to 66 Å (~22 MLs) [17]. InSb-only interfaces were attempted for these superlattices. Experimental photo-response spectra of the superlattices are shown in Figure 1. Calculations were done using fitting parameters of $x_i = 0.39$, $x_0 = 0.012$, and $R = 0.67$. These parameters were manually fitted to get a good visual satisfaction with reference to the values given in [13]. Their accuracy is good enough to describe the whole series of superlattices for design purposes. A comparison of the calculated bandgap and the experimental values is shown in Figure 2. The calculation is in good agreement with the experimental data within certain growth uncertainties. This indicates that ETBM is a reliable method to guide the superlattice design after considering the Sb segregation effects.

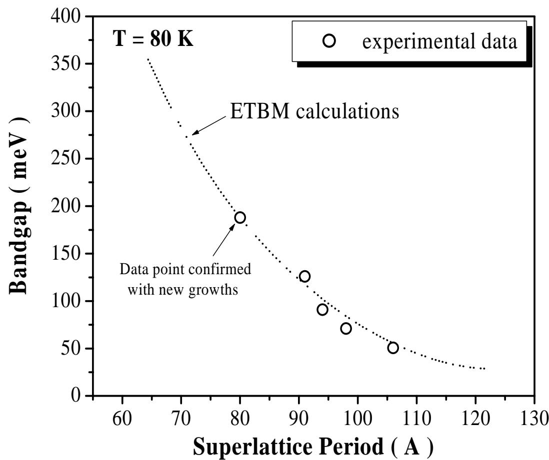
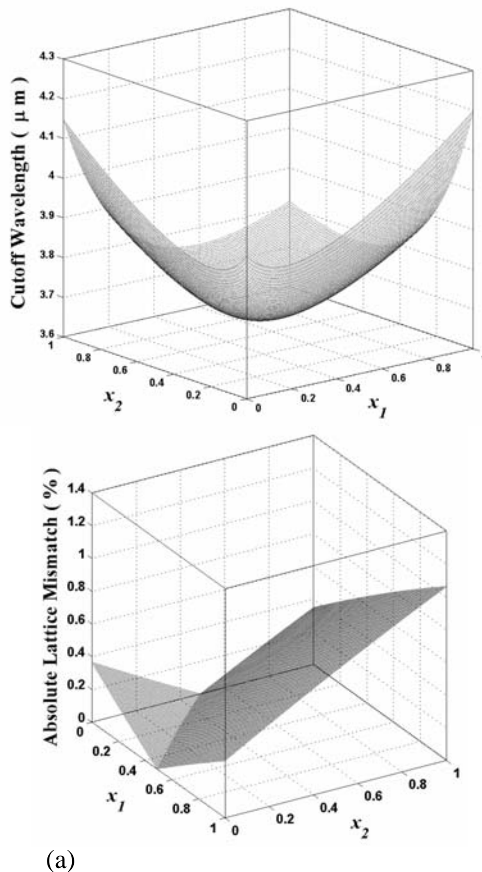


Figure 2. Comparison of the calculated bandgap with the experimental data. The experimental data points are scattered closely around the calculated curve using modified ETBM. The cut-off wavelength for a superlattice with InAs layer thickness of 40 Å has been confirmed with newer growths using the same growth conditions. [From Y. Wei, and M. Razeghi, Phys. Rev. B v.69(8), 085316 (2004)].

Although it is quite obvious that the properties of Type II InAs/GaSb superlattices should be strongly dependent on the interface type, only InSb and GaAs like types of interfaces have been experimentally exploited, and only Sb_xAs_{1-x} ($0<x<1$) mixed interfaces have been theoretically investigated using the pseudopotential method [10]. By calculation, we found out that there are limitations in applying only InSb or GaAs interfaces. In order to have almost zero lattice mismatch (well below 1000 ppm or 0.1%), there exist only a certain number of thickness combinations of InAs and GaSb (if only InSb interfaces are used) or one has to use alternating m-GaAs interfaces ($m \geq 0$) with n-InSb ($n \geq 0$, $m \cdot n \neq 0$) interfaces. For the former case, it is only valid for cut-off wavelengths greater than ~7 μm. Below this wavelength, there does not exist any combination of InAs and GaSb numbers of layers that can satisfy the zero lattice mismatch condition, except for unreasonably thick GaSb layers. For the latter case, since the interfaces are highly strained or stressed (~7% mismatch with GaSb (001) substrates), the superlattice may relax before growing the next type of interface.



(a) (b)

Figure 3. (**a**) The three-dimensional view of the cut-off wavelength for $[(AsIn)_6\text{-}AsGa_{x1}In_{1-x1}\text{-}(SbGa)_{10}\text{-}SbGa_{x2}In_{1-x2}]_N$. The maximum cutoff wavelength can be obtained when $x_1 = 1$ and $x_2 = 0$. The minimum cutoff wavelength can be obtained when $x_1 \approx 0.5$ and $x_2 \approx 0.5$. (**b**) The three-dimensional view of the absolute values for the lattice mismatch in between the superlattice of $[(AsIn)_6\text{-}AsGa_{x1}In_{1-x1}\text{-}(SbGa)_{10}\text{-}SbGa_{x2}In_{1-x2}]_N$ and the GaSb (001) substrate. We can clearly see the zero mismatch composition line. [From Y. Wei, and M. Razeghi, Phys. Rev. B v.69(8), 085316 (2004)].

For a cutoff wavelength below 7 μm, if InSb only interfaces were used, the lattice mismatch in the growth direction between the superlattice and the GaSb (001) substrate would be positive, *i.e.* the superlattice would begin to have a larger average lattice constant. This makes it possible to introduce GaAs species into the interface layers so as to decrease the average lattice constant and balance the strain in the superlattice. For the theoretically investigated Sb_xAs_{1-x} interfaces, growth is not very practical since the flux ratio of Sb and As is not the composition ratio of Sb and As in the material. In addition, this composition ratio is highly sensitive to the growth temperature; thus it is difficult to have good repeatability and reliability to implement these type of interfaces. We have investigated Ga_xIn_{1-x} type mixed interfaces both theoretically using ETBM and experimentally using a Varian Mod

Gen II Molecular Beam Epitaxy system. A clear advantage of using Ga_xIn_{1-x} type interfaces is that it is highly controllable and highly repeatable because of the non-volatility of the Ga and In species on the sample surface at a growth temperature close to 400 °C.

As an example, for a superlattice of $[(AsIn)_6\text{-}AsGa_{x1}In_{1-x1}\text{-}(SbGa)_{10}\text{-}SbGa_{x2}In_{1-x2}]_N$, Figure 3(a) shows the calculated three-dimensional map of cut-off wavelength using $x_i = 0.124$, $x_0 = 0.01$, and $R = 0.67$. These parameters were fitted to a series of superlattice growths that were done with lower antimony seeds. The entire range for x_1 and x_2 (from 0 to 1) has been calculated. The cutoff wavelength ranges from 3.68 μm to 4.20 μm, which corresponds to an energy variation of ~40 meV. The calculated three-dimensional map of mismatch (absolute values) is shown in Figure 3(b). We can see clearly that there exists a zero mismatch composition line. With almost zero lattice mismatch, superlattices of very high crystalline quality can be grown with significant thickness ~ microns. Once the strain is balanced in between the interfaces and the InAs layers, varying GaSb layer thickness has a negligible effect on the lattice mismatch between the superlattice and the substrate. However, increasing the GaSb layer thickness would lead to shorter cut-off wavelength. Therefore in order to maintain a similar cut-off wavelength, the InAs layer thickness would need to increase appropriately.

2.2 Nanopillar Structures

The leap from quantum wells/superlattices to quantum dots is a revolution in semiconductor design. Utilization of quantum dots would lead to potentially higher device operating temperature, especially for narrow bandgap infrared materials. In order to utilize quantum dots, there are two major schemes: one is using inter sub-band transitions, and the other is using interband transitions. In both cases, quantum dots have discrete energy levels. However the fundamental reason that leads to higher operating temperature is very different. For quantum dots utilizing inter sub-band transitions, it is the phonon bottleneck that prevents the phonon relaxation processes. Figure 4 illustrates this concept. The optical phonon process gets suppressed significantly due to non-matching energy levels for the transition. For quantum dots utilizing inter-band transitions, it is the "Auger bottleneck" that prevents the carrier recombination. Figure 5 illustrates the concept. Since the band energy levels can be engineered, appropriate material design is necessary.

When the physical dimensions of the material become small enough, why will it have discrete energy levels? From a quantum mechanics point of view, the energy levels inside semiconductor bulk/quantum wells are so crowded that we do not see the difference between them. However, for a nanometer-scaled object, the surrounding potential barrier would limit the amount of allowed energy levels and the difference between the energy levels becomes "visible". Ideal quantum dots are zero-dimensional structures, *i.e.* the momentum of electrons is quantized in all directions. A consequence of such confinement is that the energy levels are quantized, and produce a behavior similar to an artificial atom. Control of the dot dimensions leads to direct engineering of energy levels within the semiconductor crystal.

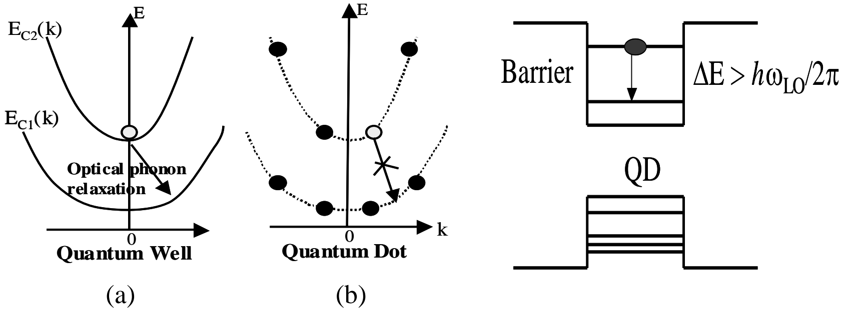


Figure 4. Illustration of the phonon bottleneck: (**a**) in k-space, (**b**) in real space

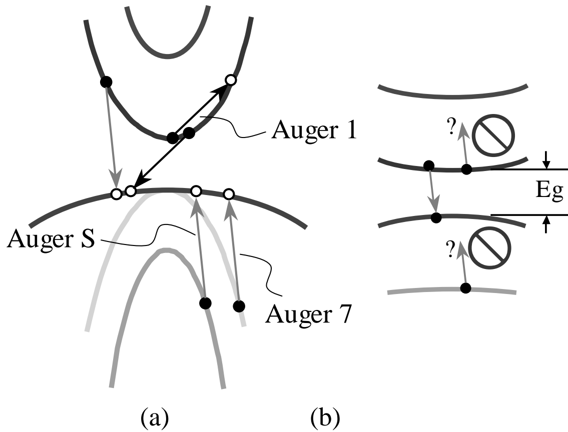


Figure 5. Illustration of the "Auger bottleneck" in k-space: (**a**) band-to-band Auger recombination processes, (**b**) the spacing of the energy levels does not match to the Auger recombination conditions

Type II superlattices utilize interband transitions. In order to realize the quantum dot structures for this material system, a topside etch-down process is necessary. This creates pillar-shaped dots. We call these structures Type II nanopillars, and cylindrical-shaped quantum dots are considered here in theoretical terms. A diameter of less than 100 nm for the nanostructure is important due to the onset of quantum size effects.

As the Type II superlattice period and pillar diameter approach the de Broglie wavelength of the electrons, the structure should begin to behave like a collection of stacked quantum dots. In this work, it is assumed that we may treat the device as a pseudo-quantum dot structure. Unlike quantum wells, which display one-dimensional carrier confinement, ideal quantum dots are zero-dimensional structures, *i.e.* the momentum of electrons is quantized in all directions. One consequence of such confinement is that the energy levels are discrete, and produce

behavior similar to an artificial atom. Control of the dot dimensions leads to direct engineering of energy levels within the semiconductor crystal. The interband transition energies can be thus designed. This enables researchers to change the detector cut-off wavelength by simply controlling the nanopillar diameter. This can be translated into multi-color detection by fabricating different diameter nanopillars on the same substrate. The top contact metal would be patterned so that like-sized pillars would be sensitive together. This concept is shown in Figure 6. Multicolor infrared detectors provide several advantages over their single-band counterparts. Multicolor detection generally improves target/background contrast, allows users the ability to detect a wider variety of infrared sources as well as perform spectrum analysis without the need for waveband filters. This added functionality can generally be packaged into a relatively smaller die, which is desirable for lightweight and space applications.

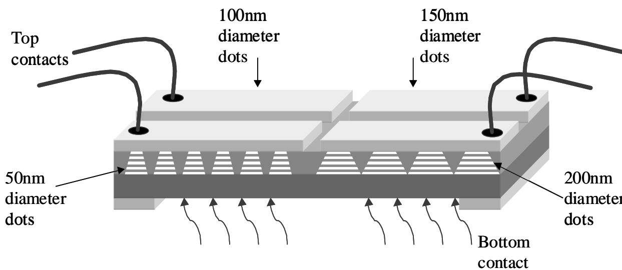


Figure 6. Schematic diagram of a possible multi-color infrared nanopillar detector array. [From M. Razeghi, A. Gin, Y. Wei, J. Bae, J. Nah, Microelectronics Journal v.34(5-8) 405-410 (2003)].

A single nanopillar structure will be simplified to a single particle cylindrical quantum dot and is considered here in theoretical terms [18]. We will use the effective mass approximation, which assumes a bulk-like electron and hole effective mass. Due to the localization of charge carriers, we will substitute the effective mass (and other important material parameters) of electrons and holes from InAs and GaSb bulk materials, respectively. Furthermore, we assume the parabolicity of the conduction and valence bands because we consider energy levels only near the Γ-point. As we do not apply any large magnetic fields to these devices, spin interactions are not taken into account and energy level splitting due to these interactions is considered negligible. The system is considered to be at equilibrium and is not time-dependent. Finally, for simplicity we will consider the quantum dot to have perfect cylindrical symmetry. The Hamiltonian for one particle (electron or hole) has the form

$$H = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(r,\theta,z) \tag{6}$$

In cylindrical coordinates, the Schrödinger equation for one particle is therefore

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\cdot\frac{\partial}{\partial r}+\frac{1}{r^2}\cdot\frac{\partial^2}{\partial \theta^2}+\frac{\partial^2}{\partial z^2}\right)\psi+V(r,\theta,z)\psi=E\psi \tag{7}$$

In the absence of any applied voltage or external strain, the Schrödinger equation for electrons can be decoupled with respect to r, θ and z. Analytical solutions then exist for the wave function as shown in Figure 7. Note the similarity with hydrogen atom wave functions. The time-independent Schrödinger equation in the radial direction is

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\cdot\frac{\partial}{\partial r}-\frac{p^2}{r^2}+V\right)\psi_r=E\psi_r \tag{8}$$

where $p = 0,1,2..$ is the angular momentum quantum number and V is an infinite potential barrier for $r > R_{max}$. The solution in the (r, θ) plane is

$$\psi(r,\theta)=J_p(kr)\cos(p\theta+\varphi) \tag{9}$$

where, $k=\sqrt{-\frac{2m^*E}{\hbar^2}}$, and J_p is the pth Bessel function of the first kind. If $p > 0$, two orthogonal solutions are obtained for $\varphi = 0$ and $\frac{\pi}{2}$. The energy levels are given by:

$$J_p(k\cdot R_{\max})=0 \tag{10}$$

Considering the one- and three-dimensional particle-in-a-box examples, we expect the same behavior in nanopillar structures. As we increase the confinement of carriers in nanopillars of decreasing diameter, it seems that we can assume that the energy levels will increase correspondingly.

The electron and hole energies have been calculated for Type II InAs/GaSb superlattice material for various confining radii. The shift of the transition energy levels can be significant with increasing radial confinement, as shown in Figure 8. Assuming the average energy gap is 100 meV, which can be easily achieved in the InAs/GaSb superlattice system, the cut-off wavelength for 40 nm diameter dots will have a blue shift of ~2 µm with respect to the one-dimensional superlattice. This is detailed in Table 1. Large energy shifts can be seen as nanopillar diameters shrink to below 50 nm. For instance, for an as-grown Type II InAs/GaSb superlattice with cut-off wavelength around 12 µm, we predict a cutoff wavelength of less than 9 µm for 30 nm diameter nanopillars in the same material. Since the quantum confinement is not dependent on the temperature, the wavelength red shift should not dependent on the temperature as well. Therefore, it seems advantageous to fabricate nanopillars with various diameters on the same substrate. Depending on their size, these devices should have different cutoff wavelengths, enabling multi-color detection.

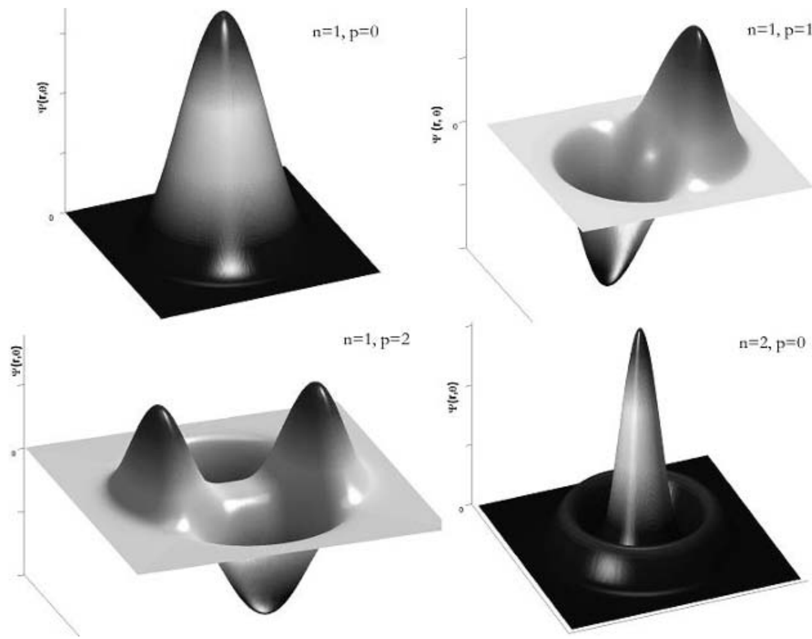


Figure 7. Radial wave functions of the lowest four energy levels of a cylindrical quantum dot; n=1,2… is the nth energy level at a given p. [From M. Razeghi, A. Gin, Y. Wei, J. Bae, J. Nah, Microelectronics Journal v.34(5-8) 405-410 (2003)].

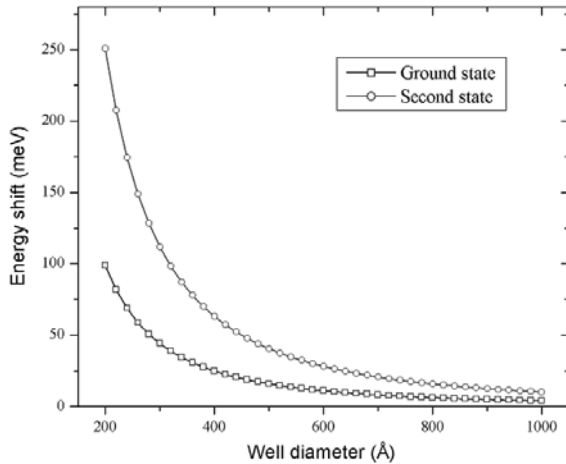


Figure 8. Effect of the cylindrical quantum dot diameter on the bandgap energy of InAs/GaSb superlattice. The mass of electrons and holes are 0.024 m_0 and 0.28 m_0 respectively and are assumed constant. [From M. Razeghi, A. Gin, Y. Wei, J. Bae, J. Nah, Microelectronics Journal v.34(5-8) 405-410 (2003)].

Table 1. Expected energy and wavelength shifts for Type II InAs/GaSb material with various pillar diameter and bulk superlattice cut-offs. This quantum effect is not temperature dependent.

Pillar diameter (nm)	Energy shift (meV)	Cut-off wavelength (μm), $\lambda_0 = 4$ μm	Cut-off wavelength (μm), $\lambda_0 = 12$ μm
100	4	3.95	11.6
50	16	3.8	10.4
40	25	3.7	9.7
30	44	3.5	8.4

3 Material Growth and Characterization

The material was grown using an Intevac Modular Gen II molecular beam epitaxy system equipped with As/Sb valved cracker sources, and In/Ga SUMO cells on *p*-type epi-ready GaSb (001) substrates. The photodiode structures were grown at 396 °C according to a calibrated pyrometer. First, a 0.75μm GaSb buffer layer doped with Be ($p \sim 1\times10^{18}$ cm^{-3}) was deposited. Then a 0.5 μm thick InAs/GaSb:Be ($p \sim 1\times10^{18}$ cm^{-3}) superlattice was grown, followed by a 2 μm thick nominally undoped superlattice. Finally, a 0.5μm thick InAs:Si/GaSb ($n \sim 1\times10^{18}$ cm^{-3}) superlattice was grown and capped with a 100Å thick InAs:Si ($n \sim 1\times10^{18}$ cm^{-3}) top contact layer. The growth rate was 0.5 monolayer/s for InAs layers and 0.8-0.9 monolayer/s for GaSb layers. The V/III beam-equivalent pressure ratio was about 4 for InAs layers and about 1.2 for GaSb layers. The cracker temperature for As was 800°C and that for Sb cells was 900°C.

Structural quality of the epitaxial layers was assessed using a Phillips high-resolution X-ray diffraction system. Figure 9 shows the typical X-ray diffraction pattern of the photodiode structures. The mismatch between the average lattice constant of the superlattice and the GaSb substrate can be controlled well within 0.1%, using either interface engineering or layer thickness tuning. For growths done on two-inch wafers, the typical full width at half maximum (FWHM) of the zeroth order peak is below 19 arcsec. The surface morphology of the samples was studied with a Digital Instruments Nanoscope IIIa atomic force microscope (AFM). The theoretical study [19] as well as experimental results [20] show the strong correlation between the surface roughness and the performance of InAs/$Ga_{1-x}In_xSb$ superlattice photodiodes. We have routinely achieved a root mean square (RMS) surface roughness between 1.2 and 1.4Å over an area of 20μm×20μm, which is the record for the Type II InAs/$Ga_{1-x}In_xSb$ material system. Figure 10 shows the surface morphology. Clear and wide atomic steps are visible over the 20μm×20μm scan area. Transmission electron microscopy (TEM) was used to analyze directly the cross-sectional superlattice quality with collaboration from Wright Patterson Air

Force Base (WPAFB). Figure 11 shows the TEM images of a sample grown with an AlSb buffer layer in three different magnifications. Interfaces are identified by the white straight lines. From the TEM images, we did not observe any dislocations, which is consistent with the narrow FWHMs in the X-ray diffraction and the small RMS roughness of the sample surface. Using a classical stress-strain theory, the superlattice period was calculated to be 55.0 Å, matching the 55±0.5 Å from the X-ray measurements. The XTEM measurement shows a period of 56.3±3 Å, with an individual layer thickness of 19.7±3 Å for InAs and 36.4±3 Å for GaSb.

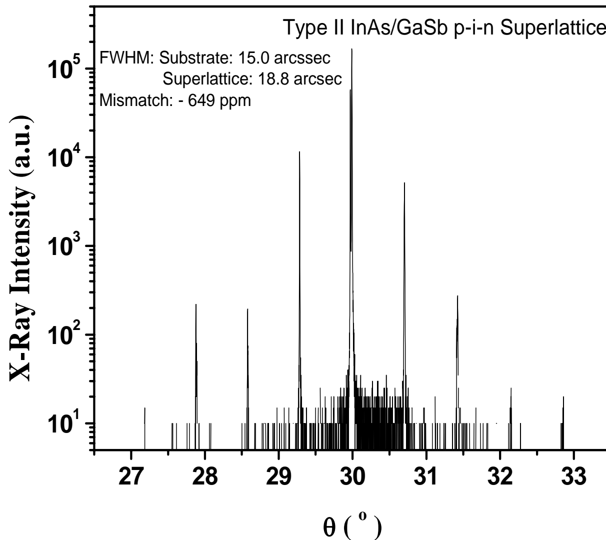


Figure 9. High-resolution X-ray diffraction for a typical Type II InAs/GaSb superlattice p-i-n photodiode structure grown on quarter two-inch wafers. The FWHM of the buffer layer was 15 arcsec and that of the zeroth order superlattice peak was only 18.8 arcsec. Both of these are the record for this material system. [From Y. Wei, J. Bae, A. Gin, A. Hood, M. Razeghi, G.J. Brown, M. Tidrow, J. Appl. Phys. v.94(7), 4720 (2003)].

For superlattices with cut-off wavelength below 5 μm, photoluminescence measurements can be performed at temperatures ranging from 11 K up to 300 K in a helium cryostat equipped with a sapphire window. The addition of AlSb in between the GaSb buffer layer and the superlattices would better confine the carriers and lead to higher luminescence efficiency. We have examined two different superlattices in the MWIR range: A: $[(AsIn)_6\text{-}AsGa_{0.34}In_{0.66}\text{-}(SbGa)_{10}\text{-}SbGa_{0.34}In_{0.66}]_n$, and B: $[(AsIn)_7\text{-}AsGa_{0.2}In_{0.8}\text{-}(SbGa)_{30}\text{-}SbGa_{0.2}In_{0.8}]_n$. The 50% cut-off wavelengths at liquid nitrogen temperature were calculated to be 3.7 μm and 3.58 μm respectively. These two superlattices were grown under the same conditions. We chose the superlattice compositions in order to satisfy simultaneously zero lattice mismatch and a given cut-off wavelength.

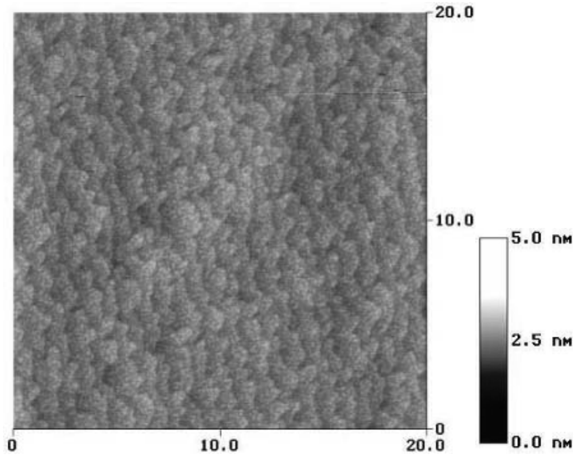


Figure 10. Surface morphology of 3 μm thick detector samples. The surface roughness was 1.57 Å for the shown area. [From Y. Wei, J. Bae, A. Gin, A. Hood, M. Razeghi, G.J. Brown, M. Tidrow, J. Appl. Phys. v.94(7), 4720 (2003)].

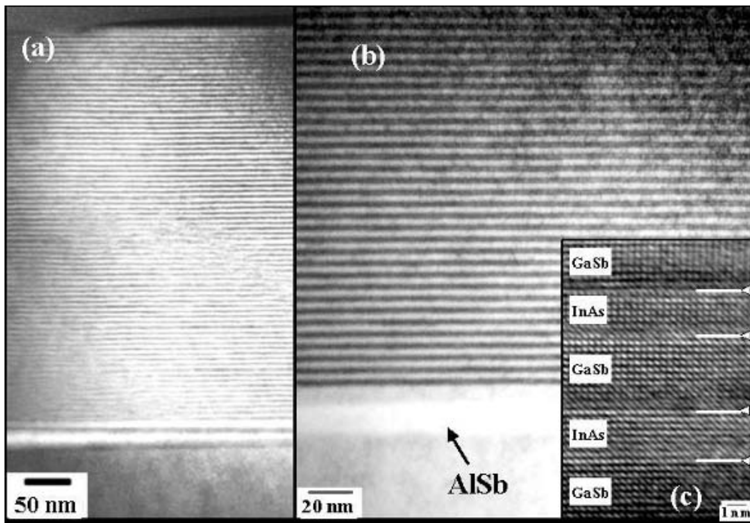


Figure 11. Transmission electron microscope images of the superlattice samples for photoluminescence measurement: (**a**) the entire grown structure (the irregular layer below the superlattice might be due to AlSb oxidation during sample preparation, or due to cleaving); (**b**) clear uniform atomic layers; (**c**) close-up view of individual layers and position of interfaces that we received from *WPAFB*. [From Y. Wei, J. Bae, A. Gin, A. Hood, M. Razeghi, G.J. Brown, M. Tidrow, J. Appl. Phys. v.94(7), 4720 (2003)].

For photoluminescence measurement, the excitation was provided by an Ar^+ ion laser emitting at 514 nm (green). Figure 12 shows a comparison of the photoluminescence (PL) intensity under the same measurement conditions at T = 11 K. The full width half maximum of these PL peaks does not represent the inhomogeneous broadening. We found that superlattice A, the sample with thinner GaSb layers, actually has at least 50% more integrated luminescence spectral intensity than superlattice B, which is opposite from conclusions drawn by other groups [21], who claimed that superlattices with thicker GaSb layers would exhibit stronger PL intensity. One probable explanation for our result is that electron and hole wavefunctions in smaller period superlattices overlap more than in a larger period superlattice material. Thus we should expect stronger PL for smaller period superlattices with the same overall thickness. Comprehensive temperature dependent and power dependent measurements were performed for the superlattice A. At 11 K under a fluent power of 140 mW, the luminescence FWHM was about 4.5 meV, as shown in Figure 13. This value is the upper bound of the inhomogeneous broadening for superlattice A. The power-dependent integrated PL intensity is shown in Figure 14 at 11 K. It shows a strong linear relationship, which indicates a band-to-band transition or excitonic behaviour.

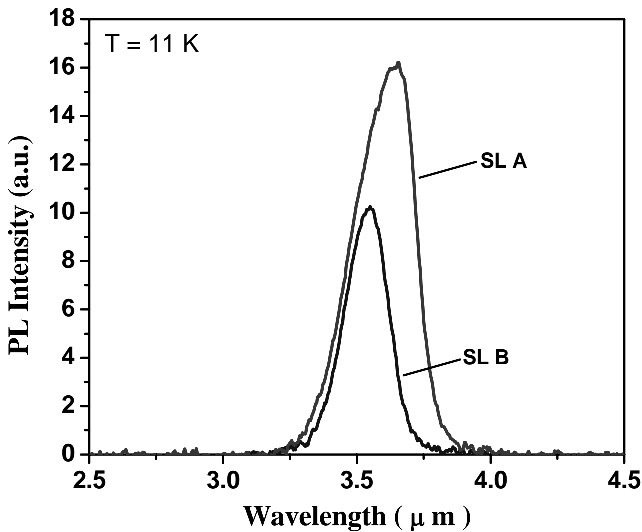


Figure 12. Photoluminescence measurement at 11 K for superlattices $[(AsIn)_6\text{-}AsGa_{0.34}In_{0.66}\text{-}(SbGa)_{10}\text{-}SbGa_{0.34}In_{0.66}]N$ (SL-A) and $[(AsIn)_7\text{-}AsGa_{0.2}In_{0.8}\text{-}(SbGa)_{30}\text{-}SbGa_{0.2}In_{0.8}]N$ (SL-B). Higher integrated luminescence intensity was observed for SL-A. [From Y. Wei, M. Razeghi, G.J. Brown, M. Tidrow, Proc. of SPIE v.5359, 301 (2004)].

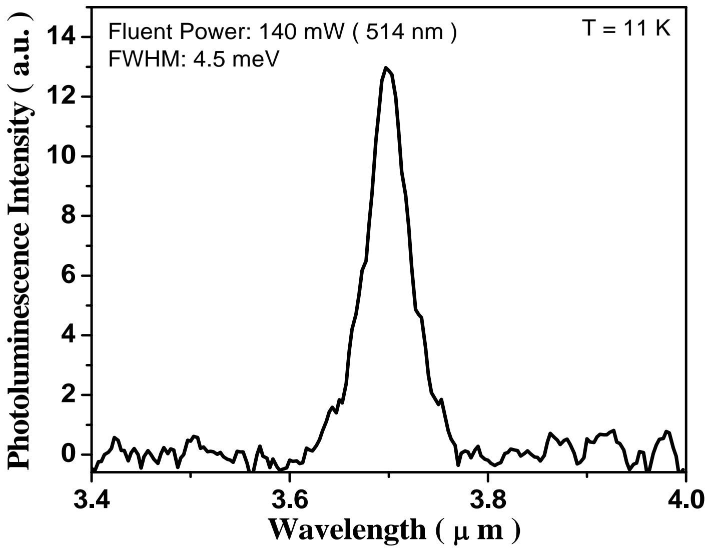


Figure 13. The photoluminescence spectrum for superlattice A at 11 K. The FWHM was about 4.5 meV. [From Y. Wei, J. Bae, A. Gin, A. Hood, M. Razeghi, G.J. Brown, M. Tidrow, J. Appl. Phys. v.94(7), 4720 (2003)].

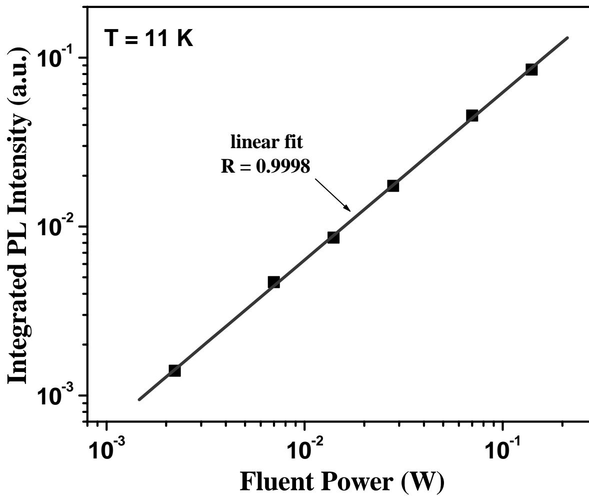


Figure 14. Power dependence of integrated luminescence intensity for superlattice A. The laser power ranges from 2.2 mW up to 140 mW. [From Y. Wei, J. Bae, A. Gin, A. Hood, M. Razeghi, G.J. Brown, M. Tidrow, J. Appl. Phys. v.94(7), 4720 (2003)].

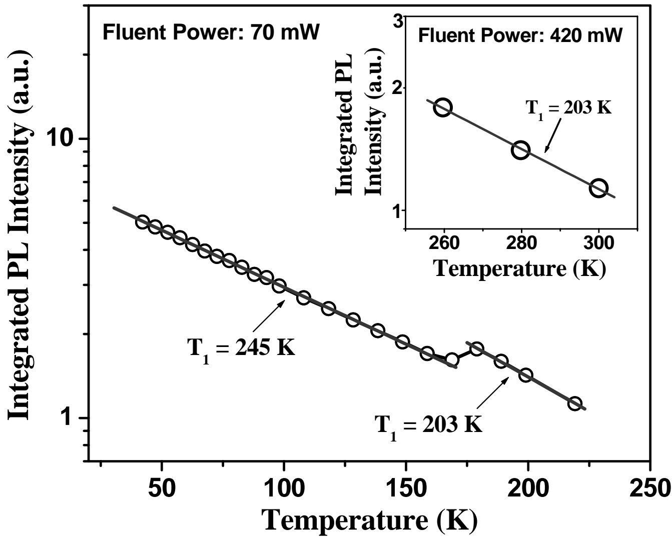


Figure 15. Temperature dependent integrated photoluminescence intensity for superlattice A. There was a camel's back feature between 170 K-180 K. Below 160 K, the characteristic temperature $T_1 = 245$ K; and above 180 K, $T_1 = 203$ K. [From Y. Wei, J. Bae, A. Gin, A. Hood, M. Razeghi, G.J. Brown, M. Tidrow, J. Appl. Phys. v.94(7), 4720 (2003)].

The temperature dependent luminescence intensity for superlattice A is shown in Figure 15. We found that an exponential decaying behavior was separated into two sections with a wiggling feature around 170~180 K. In the linear regions, the luminescence intensity $I(T)$ follows the following rule [21,22]:

$$I(T) = I_0 e^{-T/T_1} \tag{11}$$

where T stands for sample temperature, and T_1 is the characteristic temperature. Below 160 K, we found $T_1 = 245$ K, and above 180 K, $T_1 = 203$ K. These values are so far record highs, and ~4 times higher than comparable superlattices previously reported by other groups [21]. However, a distinction should be noted that other groups used different excitation conditions (laser wavelength and power). The trends of the temperature dependent photoluminescence intensity were similar to those recently reported in a Type II "W" laser structure [23].

4 Photodiodes with Cut-off Wavelength ~8 μm

For photodiodes with cutoff wavelength around 8 μm we used superlattice $[(InAs)_{11}\text{-}InSb\text{-}(GaSb)_6]_n$, whose period is around 55 Å. The growths were done on quarter two-inch or larger wafers. The (corrected) current responsivity and quantum efficiency are shown in Figure 16(a) and (b) respectively. A peak current responsivity of 2~3 A/W (~70 kV/W) and a *device* quantum efficiency of 40~50% have been achieved without any surface coating at zero bias. The 90% to 10% cut-

off energy width of these devices was only about 12 meV at 80K. Figure 17(a) shows the measured dark current *vs* voltage bias curve under 80 K background.

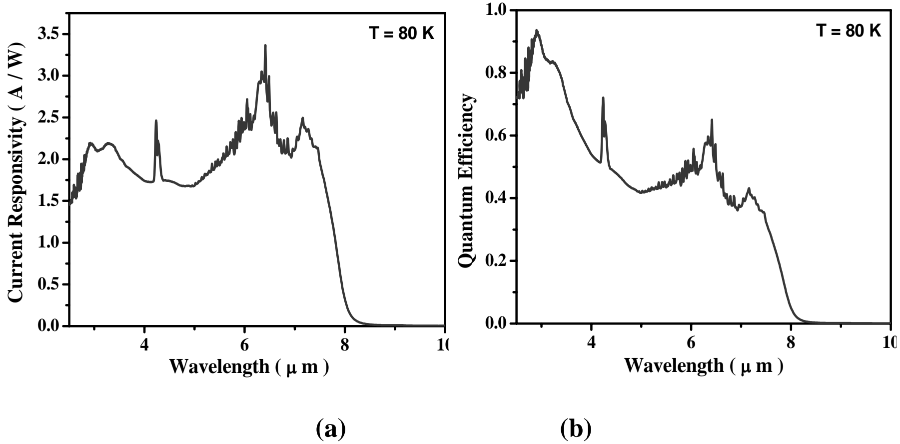


Figure 16. Calculated current responsivity (**a**) and *device* quantum efficiency (**b**) for the photodiodes under zero bias at liquid nitrogen temperature. A peak current responsivity of 2~3 A/W and a *device* quantum efficiency of 40~50% have been achieved without any surface coating. [From M. Razeghi, Y. Wei, J. Bae, A. Gin, A. Hood, J. Jiang, J. Nah, Proc. of SPIE, v.5246, 501 (2003)].

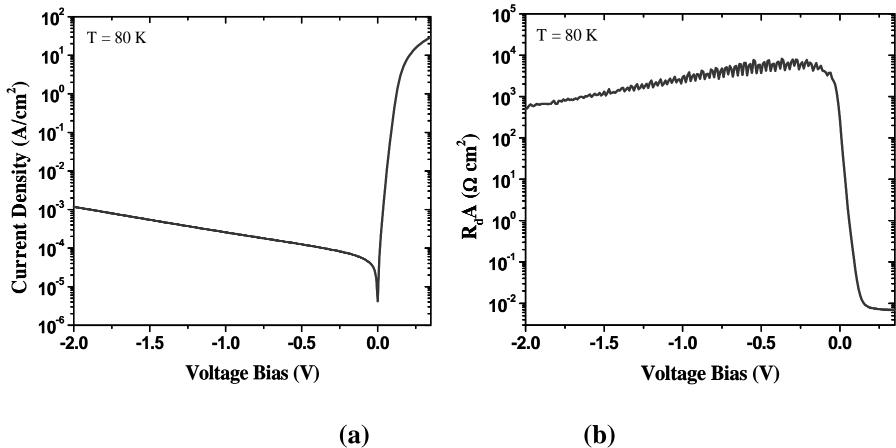


Figure 17. Measured dark current density (**a**) and R_dA product (**b**) for devices at ~80 K under 80 K background. The onset of tunneling current was not observed even up to –3 V bias. [From M. Razeghi, Y. Wei, J. Bae, A. Gin, A. Hood, J. Jiang, J. Nah, Proc. of SPIE, v.5246, 501 (2003)].

Despite significant dark current under small reverse bias, we did not observe the onset of tunneling current until –6 V reverse bias. This particular feature of the detectors has produced strong backing for these devices for focal plane array applications, where a reverse bias is always used. (Focal plane array results will be

discussed a little later). Figure 17(b) shows the measured R_dA product *vs* voltage bias curve. An R_0A of 280 Ω·cm^2, and R_dA of 7600 Ω·cm^2 at –0.2 V have been achieved. The corresponding detectivity *vs* wavelength relationship is shown in Figure 18. A thermally limited peak detectivity of 2~3×10^{11}cm·Hz$^{1/2}$/W has been obtained. This performance has made the focal plane array based on Type II InAs/GaSb superlattices realistic. Most recently, we have demonstrated an R_0A over 9000 Ω·cm^2 for diodes with a cutoff wavelength around 7.5 μm at 80 K, and detectivity beyond 10^{12} cm·Hz$^{1/2}$/W at zero bias.

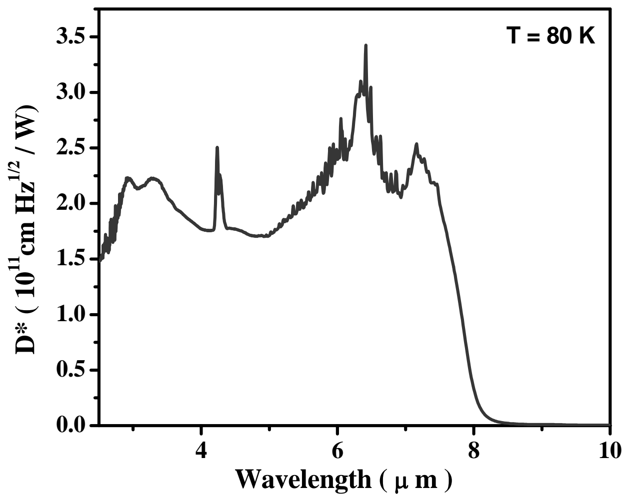


Figure 18. Detectivity *vs* wavelength curve at liquid nitrogen temperature. A thermally limited peak detectivity of 2~3×10^{11} cm·Hz$^{1/2}$/W has been achieved. [From M. Razeghi, Y. Wei, J. Bae, A. Gin, A. Hood, J. Jiang, J. Nah, Proc. of SPIE, v.5246, 501 (2003)].

5 Type II Focal Plane Arrays

One of the ultimate goals for detector development is the realization of high-quality imaging based on focal plane arrays. However, FPA technology is very complicated and the existing information is proprietary within the industry. In addition, the focal plane array based on the Type II InAs/Ga(In)Sb material system has only been demonstrated recently and there are few references. We are developing the technology in the academic area for the first time in the world. The processing of Type II FPAs is unique and faces many challenges compared with HgCdTe, QWIP, and InSb. All the up-to-date results reported here were obtained using university-based equipment.

There are three major challenges for the Type II FPA technology: surface passivation (both front and back sides), thermal mismatch, and substrate thinning. Since there is no established passivation for Type II materials, we skipped this step for the first attempts, but used a very thorough surface cleaning. The thermal

expansion coefficient of GaSb and InAs is significantly different from that of silicon, which is used for the read-out integrated circuits (ROIC). The GaSb crystal is highly brittle and extremely sensitive to the thinning process.

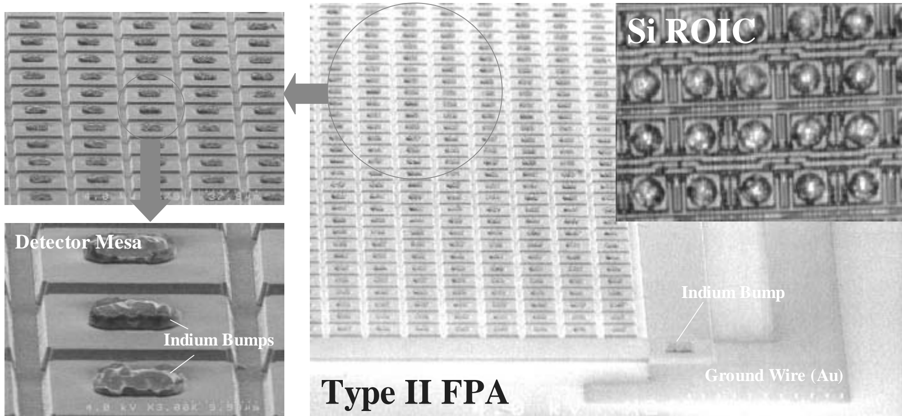


Figure 19. SEM images of FPA samples and optical image of the silicon ROIC, both with indium bumps thermally evaporated. The thickness of the indium bumps was ~2 μm. Dry etching using photoresist mask has created very nice pixel mesas. [From M. Razeghi, Y. Wei, J. Bae, A. Gin, A. Hood, J. Jiang, J. Nah, Proc. of SPIE, v.5246, 501 (2003)].

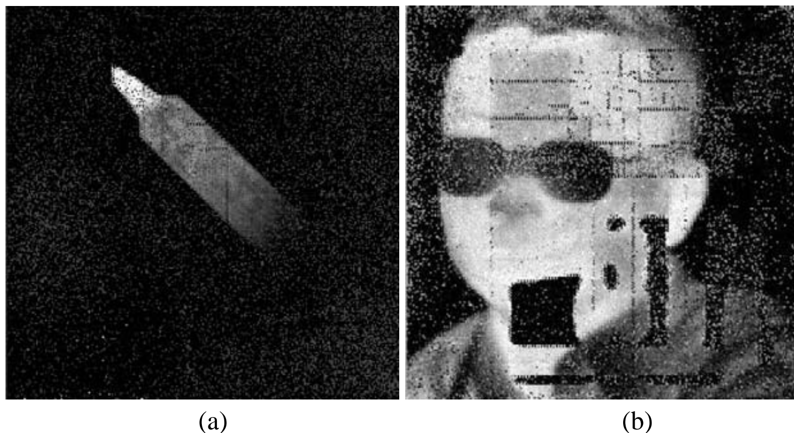


Figure 20. Images obtained from the world's first 256×256 FPA based on Type II InAs/GaSb superlattices p-i-n photodiodes: (**a**) image of a soldering iron; (**b**) portrait of author Yajun. [(a) is from M. Razeghi, Y. Wei, J. Bae, A. Gin, A. Hood, J. Jiang, J. Nah, Proc. of SPIE, v.5246, 501 (2003), and (b) is from Y. Wei, A. Hood, A. Gin, V. Yazdanpanah, M. Razeghi, M. Tidrow, Proc. SPIE, v.5732, 309 (2005)].

Figure 19 shows the result of the FPA and ROIC processing. Figure 20 shows the preliminary 256×256 FPA imaging results (λ_c~8μm). The minimum NEDT we have demonstrated is 1~2mK. This shows the potential of practical applications using Type II superlattice FPAs. Future work will identify the origin of the dead

pixels and demonstrate imaging with very high uniformity. The nanopillars based on Type II superlattices may lead to the potential of imaging applications at much higher operating temperature.

6 Nanopillar Fabrication

6.1 Electron Beam Lithography

Electron beam lithography, which uses accelerated electrons impinging on a charge-sensitive resist layer, is used in this work as a higher resolution alternative to photolithography. Similar to the scanning electron microscope, these systems have an inherent resolution advantage over photon-based imaging and lithography methods due to the much smaller de Broglie wavelength of electrons. Using a beam energy of 1 keV, the beam spot is approximately 5 nm. The Center for Quantum Devices uses a Leica LV-1 electron beam system.

6.2 Nanopillar Device Fabrication Process

III-V epitaxial material is grown using MBE or metal-organic chemical vapor deposition (MOCVD). In general, epilayers should display large atomic steps under the atomic force microscope and have a low defect density over a 20 µm × 20 µm scan size. No substrate cleaning is necessary if the sample is kept for a short period of time in nitrogen ambient after growth.

After an HMDS primer layer is applied, various electron beam resists can be spun on; for instance, a layer of (PMMA 950K C2) electron beam resist would be approximately 100nm thick. The sample is then soft baked on a hot plate at 180 °C for 90 s. The Leica LV-1 electron beam lithography system is used to expose the nanometer scale pattern in the resist. Doses ranging from 2-40 $\mu C/cm^2$ have been used to properly expose the electron beam resist, assuming features of around 50 nm in diameter and depending on resist type, thickness, *etc.*

The sample is developed in MIBK:IPA (1:2) for 10 s. The development is stopped using an IPA rinse for 10 s. The sample can now be examined under the optical microscope to identify large-area patterns and features. A hot plate bake at 100°C for 90 s is used to hard bake the developed PMMA.

After electron beam patterning with a Leica Lion LV-1 system, we evaporate gold, a combination of Ti/Au, or other metals suitable for etch masking. A layer with overall thickness of approximately 90-100nm is deposited using the metal evaporation system. A lift-off procedure in acetone removes the bulk of the metal, leaving the metal etch mask pattern. Lift-off generally takes less than 10 min in acetone soak and is followed by a 2-min water wash or bath. The sample is then blown dry with nitrogen to remove excess water. The sample may be examined with scanning electron microscope to confirm satisfactory metal lift-off.

Using an ECR reactive ion etching system, we remove the unmasked superlattice layers. This step transfers the electron beam defined pattern into the superlattice. Various dry etch recipes may be used depending on the material system and condition of the ECR chamber. Gas mixtures such as BCl_3:Ar, CH_4:H_2:Ar and Cl_2:Ar have been attempted for the material systems in this work. Numerous variables have been adjusted, including temperatures ranging from 300 ºC to

ambient, various power levels and chamber pressures. After etching, the sample can be examined by scanning electron microscope to determine etch rate, sidewall slope, *etc.*

Next we spin on a polyimide film that is much thicker than the pillar height to obtain a very flat, uniform surface. Next, we use the ECR-RIE system to etch back to just under the top of the pillar. In the ideal case, we would like to etch about 20-50 nm below the top of the pillar. This distance is feasible with current polyimide etch rates and can be determined accurately by atomic force microscope.

Large area photomasks are again used to define the top contact 'pixels' and provide bond pads for the contacts. The top contact is deposited using metal evaporation. Metal will also be deposited on the opened windows for gate contacts. Lift-off using acetone is again performed to leave the large top contact 'pixel' and gate and top contact bond pads. In practice, each group of differently sized diameter pillars will have different top contacts to allow for discrete addressing from a read out circuit, for example.

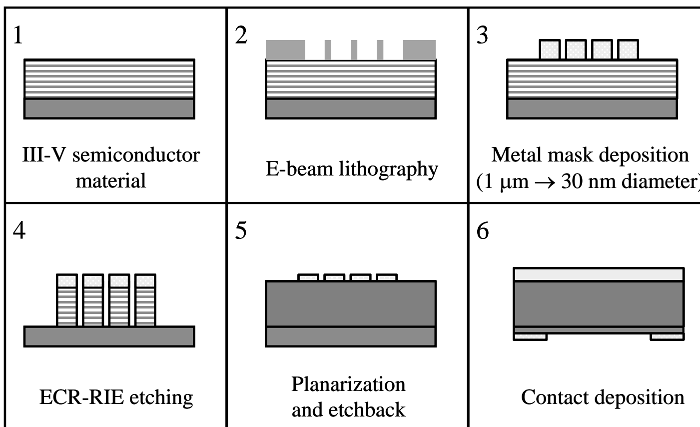


Figure 21. Nanopillar detector fabrication process. [From A. Gin, Y. Wei, A. Hood, A. Bajowala, Q. Nguyen, V. Yazdanpanah, M. Razeghi, G.J. Brown, M.Z. Tidrow, Proc. of SPIE, v.5361, p66 (2004)].

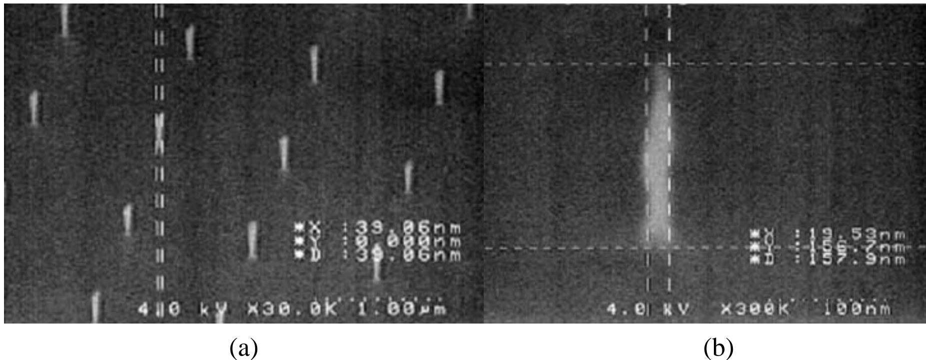
We use a chemical mechanical polishing (CMP) technique to remove the bulk of the substrate material (over 200µm). A further chemical treatment may be done to improve surface smoothness. Using backside alignment photolithography, we can deposit the backside metal contact in a ring around each detector pixel. Acetone lift-off will again be performed to define the backside contact and bond pads.

Wire bonds are connected to the top contacts. The bottom contact is simply a pressure contact with the sample holder or backside support. Incident light comes through the backside and illuminates the quantum dots. Each varying size of quantum dot detects a slightly different light spectrum, and addressing each different size group on the pixel allows us to obtain true multicolor detection. The overall nanopillar fabrication process is shown in Figure 21.

6.3 Nanopillars in GaSb Material

As a preliminary experiment, bulk GaSb material was patterned and etched using a reactive ion system. Due to the complexity of the fabrication process, it was important to begin our experiments with the processing of a homogeneous material. Furthermore, GaSb substrates have been etched successfully in the literature with reactive ion systems, so we were able to adapt similar etch recipes in our process. This experiment served as a proof of concept for the electron beam lithography and dry etching system for the realization of semiconductor nanopillars.

Epi-ready p-type GaSb (001) substrates were prepared with a spun-on HMDS resist primer followed by a 100 nm thick layer of polymethylmethacrylate (PMMA) 950K C2 resist. The samples were baked on a hotplate at 180 ºC for 90 s. Electron beam exposure was performed using a Leica LV-1 lithography system at doses ranging from 50 to 150 μC/cm^2 with beam currents between 10-20 pA. Circular dots with diameters from 2 μm to 30 nm were patterned in the resist. Additionally, the density of 50 nm dots was varied from 1 μm to 200 nm center-to-center spacing. On selected samples, arrays of dots as large as 500 μm × 500 μm were also exposed. Following exposure, samples were developed using methyl-isobutyl ketone:isopropyl alcohol (MIBK:IPA) with a volume ratio of 1:2 and baked on a hotplate at 100 ºC for 90 s. A 90 nm thick Au mask layer was deposited by electron beam evaporation and unwanted metal was subsequently removed with acetone lift-off.



(a) (b)

Figure 22. (**a**) Plasma-etched GaSb pillars in a loosely spaced array and (**b**) a close up of a pillar with diameter below 20 nm**.** [From M. Razeghi, A. Gin, Y. Wei, J. Bae, J. Nah, Microelectronics Journal v.34(5-8) 405-410 (2003)].

The ECR-RIE system was used for the pillar etch. The bulk GaSb samples were realized using BCl_3:Ar plasma chemistry. Etching pressures ranged from 1.2 to 1.8 mTorr. Microwave RF power of 325 W was used. The processes were initiated at room temperature with no active cooling. Final etch temperatures reached approximately 50 ºC. A ratio of 1:4 (BCl_3:Ar) gas flow mixture was found to give good sidewall profiles and a reasonably fast etch rate (300-400 nm/min). The pillars were inspected using a scanning electron microscope. We were able to achieve good etch uniformity across several square millimeters with nearly vertical pillar

sidewalls. Figure 22 is a scanning electron micrograph of bulk GaSb pillars in an array and one pillar with a diameter less than 20 nm.

We have performed initial experiments on GaSb substrates to achieve ~20 nm diameter nanopillar features using electron beam lithography and reactive ion etching technologies. This proof of concept exercise shows the feasibility of the fabrication method for nanostructures in III-V materials.

6.4 Nanopillars in InAs/GaSb Superlattice Material

After some preliminary experiments involving other III-V material systems, we demonstrated some quantum size effects in nanopillar structures. We fabricated nanopillars in the Type II InAs/GaSb superlattice in order to achieve similar effects for long wavelength infrared sensing materials. In this way, we hope to eventually realize multi-color infrared focal plane arrays using nanopillars with the Type II system.

Type II InAs/GaSb superlattice material was grown on GaSb substrates using molecular beam epitaxy. A one micron GaSb p+ buffer layer was grown followed by a p-i-n superlattice with 67-267-67 nm thickness respectively. Details of the epitaxial growth and superlattice design are outlined elsewhere [24]. Grown material was characterized using a Phillips high-resolution X-ray diffraction system and atomic force microscope (AFM). Root mean square (RMS) surface roughness was below 2.0 Å for 20 μm × 20 μm areas. All samples were cleaved and degreased before processing.

Lithography conditions, including pattern geometry and electron dose, were similar to those during bulk GaSb exposures. Various dot size, spacing, and doses were implemented. Samples were developed, 90 nm Au was evaporated and lift-off was performed. Au mask diameters smaller than 30 nm were common with less than 5 nm placement and stitching error. Recently, we have also used Ti/Ni etch masks with good results. Long lift-off times (~30-45m) are suggested for the best mask conditions.

6.5 Reactive Ion Etching Using BCl_3:Ar and CH_4:H_2:Ar

Sloped sidewalls in nanopillars can lead to inhomogeneous broadening of the transition energies due to the varying diameter of carrier-confining regions. This broadening can negatively influence the photoluminescence and absorption wavelength spectrum. To minimize these effects, we attempted different dry etch chemistries to fabricate superlattice nanopillars with vertical sidewalls.

Due to the success of the BCl_3:Ar etch with bulk GaSb substrates, we initially chose the same chemistry to etch the Type II superlattice material. Chamber pressures between 1 and 1.8 mTorr were used. However, these etching attempts using a BCl_3:Ar plasma were relatively unsuccessful. Semiconductor/mask etch selectivity remained greater than 5:1, but the sidewall profile was not vertical. The cone-shaped structures were most likely formed due to a non-volatile InClx product

that built up around the pillar sidewalls and base[25]. To minimize this by-product formation, we moved to a methane-based etch chemistry.

A CH_4:H_2:Ar gas flow was used to etch Type II superlattice structures in the ECR-RIE system with a ratio of 8:32:10. Microwave powers up to 325 W were attempted, but the samples were not anisotropically etched and had local surface roughness around the pillar structures. This was probably due to a combination of metal 'micromasking' and polymer deposition around the pillars. These effects can be seen in Figure 23. Polymer deposition seemed more prevalent around the superlattice material. To eliminate these issues, we developed a process which cycled between the CH_4:H_2:Ar superlattice etch and an O_2 polymer removal step.

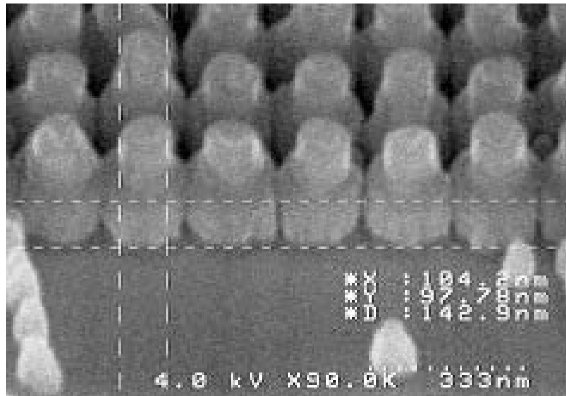


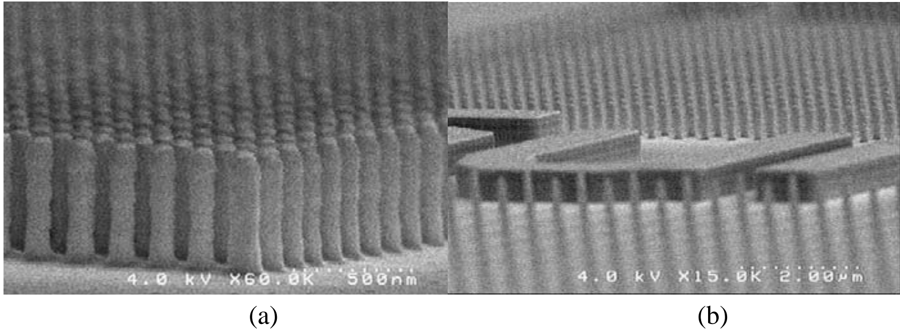
Figure 23. Type II superlattice etched with CH4:H2:Ar plasma. Note the polymer deposition on the sidewalls and metal mask. [From M. Razeghi, A. Gin, Y. Wei, J. Bae, J. Nah, Microelectronics Journal v.34(5-8) 405-410 (2003)].

6.6 Reactive Ion Etching Using Cyclic CH_4:H_2:Ar / O_2

The pillar etch was comprised of four cycles, each of which included a superlattice and polymer etching step. No microwave power was used. The etching steps were 10 and 3 min for CH_4:H_2:Ar and O_2, respectively. Between each step, the chamber was purged with nitrogen gas and evacuated. Initial substrate temperatures ranged from 18.8 to 150 ºC.

At low temperatures, a large amount of polymer deposition was evident around the pillar structures. Polymer removal seems to be proportional to substrate temperature, due to the increased thermal energy of reactants near the sample surface. At 150 ºC, there was a slight undercut and removal of the 30 nm metal masks. The optimal etch conditions were 250 W oxygen plasma, 200 W CH_4:H_2:Ar plasma, and 150 ºC substrate temperature. Using these parameters, the polymer and superlattice etch steps were well balanced and the increased sample temperature assisted the removal of polymer from the nanopillar surface. The maximum aspect ratio achieved was approximately 10:1. The smallest Type II InAs/GaSb superlattice nanopillars created thus far using this technique have a diameter of less than 50 nm

and height of approximately 480 nm. Some examples of Type II nanopillars are shown in Figure 24.



(a) (b)

Figure 24. (a) Oblique SEM image of nanopillars formed by reactive ion etching. (b) Large area oblique view of etched sample, nanopillars and large features have been etched with vertical sidewalls. [From A. Gin, Y. Wei, A. Hood, A. Bajowala, Q. Nguyen, V. Yazdanpanah, M. Razeghi, G.J. Brown, M.Z. Tidrow, Proc. of SPIE, v.5361, p66 (2004)]

More recently, we have optimized the type II nanopillar etch by modifying the dry etch recipe. Keeping the gas composition the same, we have increased the number of etch cycles while keeping the ratio of CH_4:H_2:Ar to O_2 cycle times equal. For instance, we doubled the number of cycles and halved the cycle time for CH_4:H_2:Ar and O_2 steps. Arrays of ~50 nm diameter nanopillars with 80 nm spacing in 500 μm × 500 μm fields have recently been fabricated. Assuming 100% yield, these arrays are composed of more than 39 million individual pillars.

6.7 Device Fabrication

6.7.1 Polyimide Planarization

To planarize the surface, HD Microsystems Pyralin PI2555 polyimide was applied by spin coating. Following a curing bake, the sample was cleaved to examine the interaction between the polyimide and nanopillars. As shown in Figure 25, the planarization above the nanopillar structures is very good. Furthermore, the polyimide was able to flow between pillars and fill in the nanometer-scale spaces. The overall polyimide thickness was approximately 3.6 μm. The damage seen on the pillars in the figure represents the stress caused from the sample cleaving process. We do not believe the cracks are formed during polyimide curing.

6.7.2 Polyimide Etchback

The polyimide was etched with a reactive ion etching system using an oxygen plasma. The material was successfully removed to just below the nanopillar tops (approximately 50-100nm down) as seen in Figure 26.

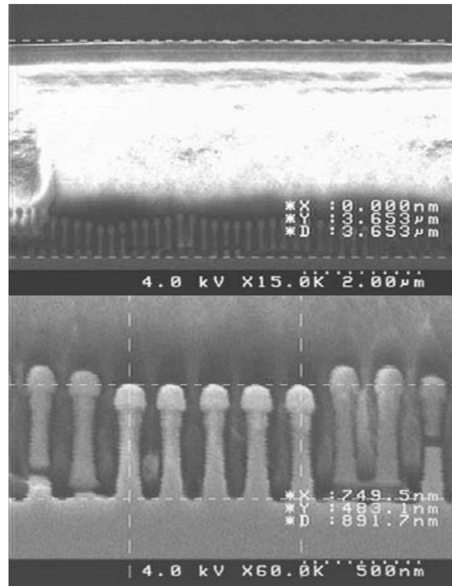


Figure 25. Polyimide coating of the nanopillar sample. Damage to nanopillars has occurred due to the cleaving process.

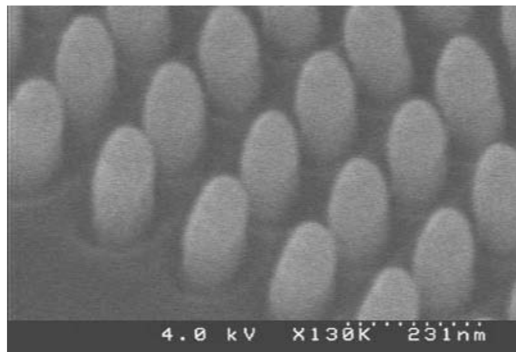


Figure 26. Oblique SEM image of polyimide etched back to reveal the tops of nanopillars

6.7.3 Top Contact Deposition

After a photolithography step, electron beam metal evaporation was used to deposit the top metal contact. The layer is a Ti/Au metal with ~200 nm thickness. From SEM images it was evident that the metal layer covers the tops of the nanopillars conformally. The top contact metal 'bridges' down to a non-patterned area to minimize the wire bond damage to the nanopillar structures. The samples were

indium bonded to leadless chip carriers and individual nanopillar areas were then wire bonded to pads located on the carrier. The conductive GaSb substrate was used as the ground contact.

6.7.4 Dark Current Measurements

Dark current measurements were performed at room temperature using a Karl Suss probe station. Current density and differential resistance of the nanopillar areas were measured using an HP 4155A semiconductor parameter analyzer. Furthermore, several areas with metal contacts over polyimide were measured for comparison. As seen in Figure 27, while the nanopillar areas exhibited a diode-like behavior with relatively high dark current, electron conduction through the polyimide layer was significantly lower, with no diode characteristic.

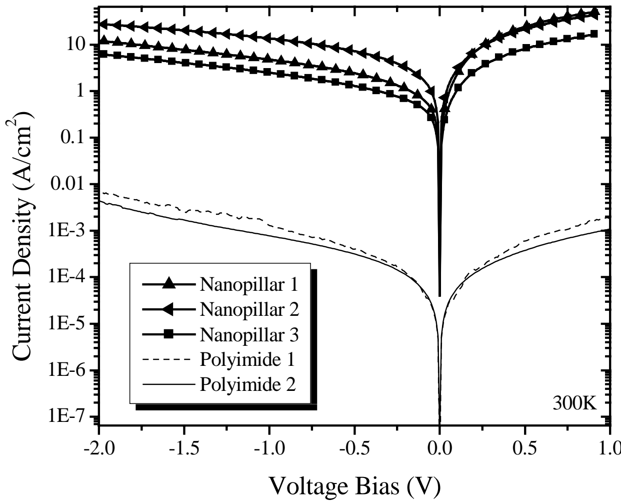


Figure 27. Room temperature dark current measurements of nanopillar and polyimide areas

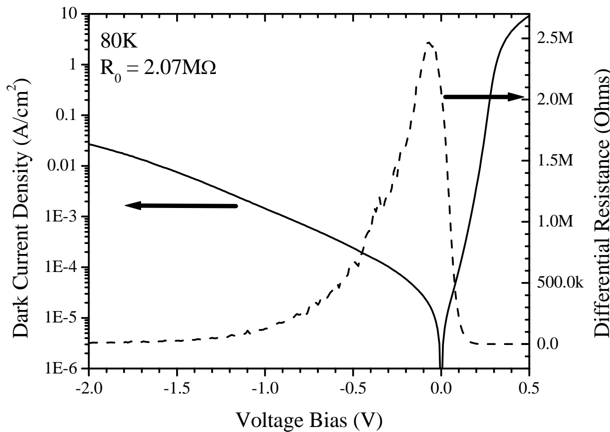


Figure 28. Dark current density and differential resistance of nanopillar area at 80K

Further I-V measurements were taken from the samples while housed in a Janis cryostat at 80K under ~4×10^{-5} Torr with cold shield. The dark current density and typical differential resistance curve for a nanopillar area is shown in Figure 28. The zero-bias resistance in this case is 2.07 MΩ. Further electrical testing is underway at this time. The detectivity measurement would require dedicate backside thinning process that is still in development.

7 Conclusion

In summary, we have successfully developed the empirical tight-binding modeling for the design of Type II InAs/GaSb superlattices. We have demonstrated very high quality superlattice material growths using state-of-the-art MBE. With our established device processing techniques, we demonstrated high performance photodiodes and the world's first infrared focal plane array in the LWIR range based on this type of superlattice. We also calculated the wavelength variations based on quantum confinement effects in the nanopillars and demonstrated the initial results for the processing of nanopillar photodiodes based on Type II superlattices.

Acknowledgement

The research was sponsored partially by AFOSR, AFRL, MDA,. The authors would like to acknowledge support and encouragement from Dr. Gail J. Brown (WPAFB), Lt. Col. Todd Steiner (AFOSR), Dr. Allan Hahn (AFRL), Dr. Vaidya Nathan (AFRL), Dr. Meimei Tidrow (MDA), and Dr. Len LaCroix (Lockheed Martin Company).

Appendix. Superlattice Hamiltonian Matrix

The values of the energy terms should be those for the corresponding atoms or adjacent atomic layers. The format of the superlattice Hamiltonian is

$$\mathrm{H_{SL}} = \begin{bmatrix} H_a & H_{ac} & 0 & 0 & \cdots & 0 & H_{ca}^{+} \\ H_{ac}^{+} & H_c & H_{ca} & 0 & \cdots & 0 & 0 \\ 0 & H_{ca}^{+} & H_a & H_{ac} & 0 & \cdots & 0 \\ 0 & 0 & H_{ac}^{+} & H_c & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \vdots & \vdots & 0 & H_a & H_{ac} \\ H_{ca} & 0 & 0 & 0 & 0 & H_{ac}^{+} & H_c \end{bmatrix}; \mathrm{H}_a = \begin{bmatrix} h_{aa} & h_{a,so} \\ h_{a,so}^{+} & h_{aa}^{*} \end{bmatrix};$$

$$h_{aa} = \begin{bmatrix} E_{sa} & 0 & 0 & 0 & 0 \\ 0 & E_{pa} & -i\frac{\Delta_a}{3} & 0 & 0 \\ 0 & i\frac{\Delta_a}{3} & E_{pa} & 0 & 0 \\ 0 & 0 & 0 & E_{pa} & 0 \\ 0 & 0 & 0 & 0 & E_{ssa} \end{bmatrix}; \; h_{a,\,so} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\Delta_a}{3} & 0 \\ 0 & 0 & 0 & -i\frac{\Delta_a}{3} & 0 \\ 0 & -\frac{\Delta_a}{3} & i\frac{\Delta_a}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

Similar forms for H_c. "*" represents complex conjugate operation, and "+" represents Hermitian conjugate operation.

$$H_{ac} = \begin{bmatrix} h_{ac} & 0 \\ 0 & h_{ac} \end{bmatrix}; \; H_{ca} = \begin{bmatrix} h_{ca} & 0 \\ 0 & h_{ca} \end{bmatrix};$$

$$h_{ac} = \beta \cdot \begin{bmatrix} f_1 E_{sasc} & \sqrt{3} f_2 \cos\theta_x E_{saxc} & \sqrt{3} f_2 \cos\theta_y E_{saxc} & \sqrt{3} f_1 \cos\theta_z E_{saxc} & 0 \\ -\sqrt{3} f_2 \cos\theta_x E_{xasc} & f_1\left[E_{xaxc} + (3\cos^2\theta_x - 1)E_{xayc}\right] & 3 f_1 \cos\theta_x \cos\theta_y E_{xayc} & 3 f_2 \cos\theta_x \cos\theta_z E_{xayc} & -\sqrt{3} f_2 \cos\theta_x E_{xassc} \\ -\sqrt{3} f_2 \cos\theta_y E_{xasc} & 3 f_1 \cos\theta_y \cos\theta_x E_{xayc} & f_1\left[E_{xaxc} + (3\cos^2\theta_y - 1)E_{xayc}\right] & 3 f_2 \cos\theta_y \cos\theta_z E_{xayc} & -\sqrt{3} f_2 \cos\theta_y E_{xassc} \\ -\sqrt{3} f_1 \cos\theta_z E_{xasc} & 3 f_2 \cos\theta_z \cos\theta_x E_{xayc} & 3 f_2 \cos\theta_z \cos\theta_y E_{xayc} & f_1\left[E_{xaxc} + (3\cos^2\theta_z - 1)E_{xayc}\right] & -\sqrt{3} f_1 \cos\theta_z E_{xassc} \\ 0 & \sqrt{3} f_2 \cos\theta_x E_{ssaxc} & \sqrt{3} f_2 \cos\theta_y E_{ssaxc} & \sqrt{3} f_1 \cos\theta_z E_{ssaxc} & 0 \end{bmatrix}$$

$$h_{ca} = \beta \cdot \begin{bmatrix} g_1 E_{sasc} & -\sqrt{3} g_2 \cos\theta_x E_{xasc} & \sqrt{3} g_2 \cos\theta_y E_{xasc} & \sqrt{3} g_1 \cos\theta_z E_{xasc} & 0 \\ \sqrt{3} g_2 \cos\theta_x E_{saxc} & g_1\left[E_{xaxc} + (3\cos^2\theta_x - 1)E_{xayc}\right] & -3 g_1 \cos\theta_x \cos\theta_y E_{xayc} & -3 g_2 \cos\theta_x \cos\theta_z E_{xayc} & \sqrt{3} g_2 \cos\theta_x E_{ssaxc} \\ -\sqrt{3} g_2 \cos\theta_y E_{saxc} & -3 g_1 \cos\theta_y \cos\theta_x E_{xayc} & g_1\left[E_{xaxc} + (3\cos^2\theta_y - 1)E_{xayc}\right] & 3 g_2 \cos\theta_y \cos\theta_z E_{xayc} & -\sqrt{3} g_2 \cos\theta_y E_{ssaxc} \\ -\sqrt{3} g_1 \cos\theta_z E_{saxc} & -3 g_2 \cos\theta_z \cos\theta_x E_{xayc} & 3 g_2 \cos\theta_z \cos\theta_y E_{xayc} & g_1\left[E_{xaxc} + (3\cos^2\theta_z - 1)E_{xayc}\right] & -\sqrt{3} g_1 \cos\theta_z E_{ssaxc} \\ 0 & -\sqrt{3} g_2 \cos\theta_x E_{xassc} & \sqrt{3} g_2 \cos\theta_y E_{xassc} & \sqrt{3} g_1 \cos\theta_z E_{xassc} & 0 \end{bmatrix}$$

$$f_1 = \frac{1}{4}\left[\exp(i\vec{k}\cdot\vec{\tau}_1) + \exp(i\vec{k}\cdot\vec{\tau}_2)\right]; \; f_2 = \frac{1}{4}\left[\exp(i\vec{k}\cdot\vec{\tau}_1) - \exp(i\vec{k}\cdot\vec{\tau}_2)\right]$$

$$g_1 = \frac{1}{4}\left[\exp(-i\vec{k}\cdot\vec{\tau}_3) + \exp(-i\vec{k}\cdot\vec{\tau}_4)\right]; \; g_2 = \frac{1}{4}\left[\exp(-i\vec{k}\cdot\vec{\tau}_3) - \exp(-i\vec{k}\cdot\vec{\tau}_4)\right]$$

$$\vec{\tau}_1 = \frac{1}{4} a_i [1+\varepsilon_{xx}, 1+\varepsilon_{yy}, 1+\varepsilon_{zz}]; \; \vec{\tau}_2 = \frac{1}{4} a_i [-1-\varepsilon_{xx}, -1-\varepsilon_{yy}, 1+\varepsilon_{zz}];$$

$$\vec{\tau}_3 = \frac{1}{4} a_i [1+\varepsilon_{xx}, -1-\varepsilon_{yy}, -1-\varepsilon_{zz}]; \; \vec{\tau}_4 = \frac{1}{4} a_i [-1-\varepsilon_{xx}, 1+\varepsilon_{yy}, -1-\varepsilon_{zz}];$$

i = material made of two adjacent atomic layers, such as InAs, InSb, GaSb, GaAs, GaInAs, GaInSb, *etc.*

$$\beta = \frac{3}{(1+\varepsilon_{xx})^2 + (1+\varepsilon_{yy})^2 + (1+\varepsilon_{zz})^2}; \; \cos\theta_j = \sqrt{\frac{\beta}{3}}(1+\varepsilon_{jj}); \; j = x, y, \text{ or } z;$$

$$\varepsilon_{xx} = \varepsilon_{yy} = \frac{a_{sub}}{a_i} - 1; \; \varepsilon_{zz} = -D_i^{001}\varepsilon_{xx};$$

a_{sub}: substrate (GaSb) lattice constant; D_i^{001} strain constant of the corresponding material i.

References

1. W. Herschel, Phil. Trans. 255 (1800)
2. W. Herschel, Phil Trans. 284 (1800)
3. G. A. Sai-Halasz, R. Tsu, and L. Esaki, Appl. Phys. Lett. **30**, 651 (1977)
4. G. A. Sai-Halasz, L. L. Chang, J. M. Welter, and L. Esaki, Solid State Commun. **27**, 935 (1978)
5. G. A. Sai-Halasz, L. Esaki, Phys. Rev. B **18**, 2812 (1978)
6. J. C. Slater and G. F. Koster, Phys. Rev., **94**(6), 1498-1524 (1954)
7. P. Löwdin, J. Chem. Phys., **18**(3), 365-375 (1950)
8. P. Vogl, H. P.Hjalmarson, J. D. Dow, J. Phys. Chem. Solids, **44**(5), 365-378 (1983)
9. D. N. Talwar and J. P. Loehr, B. Jogai, Phys. Rev. B, **49**(15), 10345-10353 (1994)
10. L. C. Lew Yan Voon, Dissertation, WPI, Worcester, MA (1993)
11. J. Steinshnider, M. Weimer, R. Kaspi, and G. W. Turner, Phys. Rev. Lett. **85**(14), 2953-2956 (2000)
12. R. Magri, and A. Zunger, Phys. Rev. B, **65**(16), 165302 (2002)
13. J. Steinshnider, J. Harper, M. Weimer, C. H. Lin, S. S. Pei, D. H. Chow, Phys. Rev. Lett. **85**(21), 4562-4565 (2000)
14. W. A. Harrison, "*Electronic Structure and the Properties of Solids: the Physics of the Chemical Bond*", Dover Publications Inc., ISBN: 0-486-66021-4 (1989)
15. O. Madelung, W. von der Osten, U. Rössler, "*Landolt-Börnstein New Series*", Group III, Vol 22(a), ISBN: 0-387-16609-2, (1987)
16. C. G. Van de Walle, Phys. Rev. B **39**(3) 1871-1883 (1989)
17. H. Mohseni, Y. Wei, M. Razeghi, Proc. SPIE, **4288**, 191-199 (2001)
18. M. Razeghi, Y. Wei, A. Gin, A. David, Proc. of Electrochemical Society, 2002-**14** 88-97 (2002)
19. G. Bastard, Phys. Rev. B **25**, 7584 (1982)
20. F. Fuchs, L. Burkle, W. Pletschen, J. Schmitz, M. Walther, H. Gullich, N. Herres, and S. Mueller, Proc. SPIE **3794**, 41 (1999).
21. A. P. Ongstad, R. Kaspi, and C. E. Moeller; M. L. Tilton, D. M. Gianardi, and J. R. Chavez; G. C. Dente, J. Appl. Phys. **89**(4), 2185 (2001)
22. M. Razeghi, *MOCVD Challenge*, Volume I, Institute of Physics Pub (1995)
23. C. L. Canedy, W. W. Bewley, C. S. Kim, M. Kim, I. Vurgaftman, and J. R. Meyer, J. Appl. Phys. **94**, 1347 (2003)
24. M. Razeghi, Y. Wei, A. Gin, G. J. Brown, and D. K. Johnstone, Proc. SPIE, **4650,** 111-116 (2002)
25. F. Ren, W. S. Hobson, J. R. Lothian, J. Lopata, J. A. Caballero, S. J. Pearton and M. W. Cole, Appl. Phys. Lett., **67**(17), 2497-2499 (1995)

High-speed Avalanche Photodiodes for the 2-5 μm Spectral Range

M.P. Mikhailova, I.A. Andreev
Ioffe Physical-Technical Institute RAS
25 Polytechnicheskaya
St. Petersburg
Russia

1 Introduction

Impact ionization processes in III-V semiconductors and their alloys have for a long time been a subject of intense research [1-8]. This is partly because of the fundamental interest in the behaviour of hot carriers in high electric fields. But, there is also a practical requirement to create avalanche photodiodes (APDs) *i.e.* low-noise photodetectors with internal gain and fast response, for use in optical fiber communications, laser range-finder systems and so on.

In the past, great attention was paid to photodetectors for the spectral range of 1-1.6 μm [9-11] for which there are silica optical fibers with low losses [12] and for which long-lifetime semiconductor lasers and highly efficient light emitting diodes have been developed [13-15]. APDs based on Ge and Si were originally used for optical communication systems in the range of 1.06-1.55 μm [5]. But, later III-V materials, such as GaAs, AlGaAs, InP, InGaAs/InP, InGaAsP/InP, InAlAsSb were studied as promising candidates for the design of APDs [16-20].

Photodetectors for optical communication should simultaneously satisfy some important requirements: high efficiency, fast response or high gain-bandwidth product, and very low noise within the working bandwidth. It is well known that to fabricate low-noise, high speed APDs it is necessary to utilize semiconductor materials with large asymmetric charge carrier ionization coefficients [21-23]. But in most semiconductors (GaAs, InP, InGaAs) the electron and hole ionization coefficient ratio is about 1. This is the main reason why up to now there are practically no high performance APDs for the spectral range 1.06-1.55 μm based on III-V alloys.

Recently optical fibers based on chalcogenide and fluoride glasses were created [24, 25]. Very low optical losses (up to 10^{-2} to 10^{-3} dB/km) and small dispersion were predicted in the spectral range of 2–4 μm where the first atmosphere transmission window lies [25]. There are other important areas of applications in the mid– infrared region. High resolution laser diode spectroscopy and laser range-finding with solid state lasers such as Ho:YAG (λ=2.06 μm) or Er:YAG:(λ=2.94 μm) also need high speed MID-IR photodiodes [26-28]. Furthermore, a free-space optical link in the mid–IR spectral range is very promising for high frequency communications [29].

It has been established that there is a class of materials based on GaSb and InAs and their alloys where high asymmetry of electron (α) and hole (β) ionization

coefficients can be achieved [4, 6, 30-36]. Many of these investigations were made in the Ioffe Institute RAS, Russia. In this paper impact ionization properties of III-V semiconductors and their connection with band structure will be considered. The multiplication factor and ionization coefficient ratio of electrons and holes in APDs will be discussed. Anisotropy of electron ionization coefficients in multi-valley semiconductors and monopolarity of ionization coefficients in GaSb will also be considered (Section 2). A new physical approach will be examined to design an APD for the spectral range 2-5 μm based on materials with high ionization coefficient ratio, in particular, InAs and GaSb solid solutions with energy gap E_g equal to the valence band spin-orbit splitting, Δ_0 where "resonant" impact ionization takes place. Experimental and theoretical studies of impact ionization processes in several III-V semiconductors (such as GaSb, GaInAs, GaInAsSb, GaAlAsSb, InAs, InGaAs, InSb, InAsSb, GaInSb) and experimental data on ionization coefficients will be described (Section 3). High speed APDs for the spectral range 2-2.5 μm based on GaInAsSb/GaAlasSb heterostructures are presented in Section 4, including low-noise APDs with separate absorption and multiplication regions (SAM APD). The dark current problem in GaSb-based APDs will also be discussed. Receiver sensitivity for long-wavelength optical communications with GaInAsSb/GaAlAsSb APDs will be evaluated. Section 5 is devoted to low-noise narrow gap APDs (λ=3.0-5.0 μm) based on InAs and its alloys (InAsSb, InGaAs, InAsSbP) with a "resonant" composition. Ionization coefficient dependence on band structure peculiarities will also be considered. Finally, in Section 6 superlattice APDs will be briefly described as well as some methods of ionization coefficient separating. Some applications of high-speed APDs for the 2-5 μm spectral range will also be reviewed.

2 Impact Ionization in III-V Semiconductors

2.1 Interband Ionization in Semiconductors

Interband ionization is the process of excitation of electron-hole pairs caused by the Coulomb interaction of electrons. If pair generation occurs due to the energy of a hot electron from the conduction band we speak of electron impact ionization (see Figure 1a). The process, in which the ionizing particle is a hole, is referred to as hole impact ionization (see Figure 1b). Impact ionization can be observed when heating the charge carriers using light with sufficiently high-energy photons as well as carrier heating in a high electric field. The latter is most important, as a large number of semiconductor devices work in conditions of impact ionization in high electric fields. In the present review we confine ourselves to just this case.

Impact ionization is characterized by the ionization coefficient, *i.e.* the number of additional electron-hole pairs generated by a charge carrier per unit length in the electric field. The electron ionization coefficient α (E) is related to the distribution function $f(\varepsilon, E)$ as

$$\alpha(E)=\frac{1}{v_{drift}}\int_{\varepsilon_i}^{\infty} w_i(\varepsilon)f(\varepsilon, E)\rho(\varepsilon)d\varepsilon \quad (1)$$

Here v_{drift} is the electron drift velocity, f(ε,E) is the distribution function, $w_i(\varepsilon)$ is the impact ionization probability per unit time, ρ(ε) is the density of states; ε is the threshold energy of the impact ionization which is usually larger than, or of the order of, the energy gap E_g. This relation is given for the case when the impact ionization probability depends only on energy. In the general case of the bands being anisotropic, the impact ionization probability can depend both on the magnitude and the direction of the electron momentum. This situation takes place, *e.g.* in Si [37]. The hole impact ionization coefficient is customarily marked as β(E) and defined by a formula identical to Equation (1).

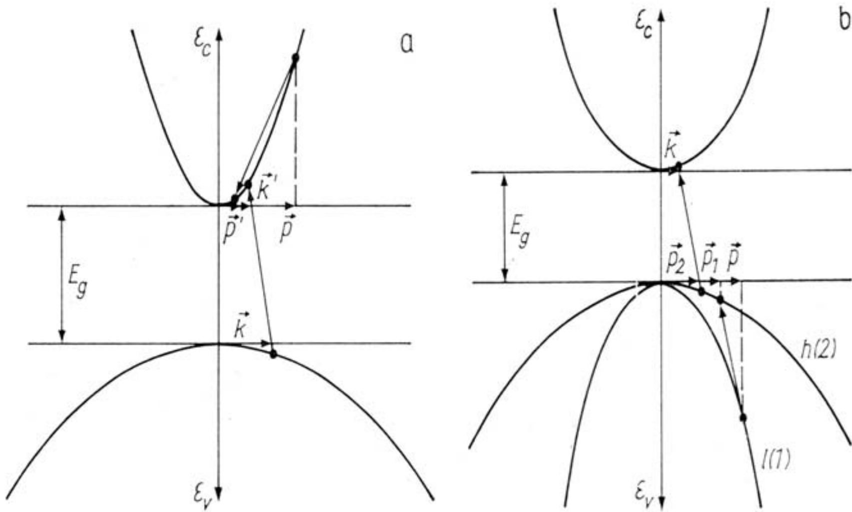


Figure 1. a) Scheme of the electron impact ionization. ***p, p'*** are the initial and final momentum of the ionizing electron ***k, k'*** – the momentum of the generated hole and electron respectively. **b**) Scheme of the hole impact ionization. ***p*** is the initial momentum of the ionizing light hole. In the final state, there are two holes in the heavy subband with momentum $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ as well as the electron with momentum k in the conduction band.

2.2 Threshold Energy of Impact Ionization

A threshold energy for impact ionization must exist as the electron-hole pair generation requires an energy of not less than E_g. Indeed, the threshold ionization energy is somewhat larger. An electron with a high kinetic energy owing to the Coulomb interaction generates an electron-hole pair and drops towards the bottom of the conduction band. Here the conservation law of not only energy but also momentum should be fulfilled (see Figure1a):

$$\varepsilon_c(p)=\varepsilon_c(p')+\varepsilon_v(k)+\varepsilon_c(k')+E_g \quad (2)$$

$$p=p'+k+k' \quad (3)$$

where $\varepsilon_c(p)$, $\varepsilon_v(k)$, $\varepsilon_c(k')$ are the kinetic energies of ionizing electron and produced particles respectively; $\varepsilon_c(p)$ is the electron energy after ionization; p, p', k, k' are the corresponding momentum of the particles generated. It follows from Equations (2) and (3) that, in the case of simple parabolic bands [38,39], the lowest energy necessary for the electron to generate a pair, *i.e.* ionization threshold energy ε_{ie}, equals

$$\varepsilon_{ie}=E_g(1+m_e/(m_e+m_h)) \quad (4a)$$

and for heavy holes

$$\varepsilon_{ih}=E_g(1+m_h/(m_e+m_h) \quad (4b)$$

where, m_e, m_h are the effective masses of an electron and heavy hole respectively. It is clear from Equation (4) that ε_{ie}~E_g if m_e<<m_h and ε_{ih}~$2E_g$ which is the case in III-V semiconductors. Indeed, the bands are non-parabolic and the effective mass increases as the energy grows. Therefore, the threshold energy will be larger than the bandgap; *e.g.* in GaAs ε_{ie} is close to $1.5E_g$. Impact ionization can also be accompanied by emission or absorption of phonons. In this case its probability will be lower. However the phonon contribution removes the limitation due to momentum conservation, and then for the threshold energy we have

$$\varepsilon_{ie}=E_g \pm h\omega \approx E_g \quad (5)$$

where ω is the phonon frequency. Here, sometimes an impurity can play the part of the third body in the process of ionization and ionization processes of this kind are called indirect.

Impact ionization is usually studied in electric fields for which the average energy of the charge carriers is not large $\varepsilon \ll E_g$. In this case, the number of electrons with large energies, which are capable of performing the impact ionization, is small and falls exponentially with rising energy. Therefore, the impact ionization in an electric field is usually controlled (exceptions are very strong fields when the electron energy $\varepsilon \approx E_g$) by the process associated with the lowest threshold. In the following considerations, the threshold ionization energy will be always taken equal to E_g.

If one considers hole ionization threshold taking into account the spin-orbit splitting band Δ_0, we obtain for the semiconductors with $\Delta_0<E_g$ (GaAs, InP *etc.*) a lower hole ionization threshold close to E_g [39]:

$$\varepsilon_{ih}=(1+\frac{m_{so}(1-\Delta_0/E_g)}{2m_h+m_e-m_{so}})E_g \approx E_g \qquad \text{at} \quad m_h>m_{so}\gg m_e \quad (6)$$

2.3 The Dependence of the Ionization Coefficients on the Electric Field

When treating experimental results of the field dependence of the ionization coefficient the following relation is usually applied:

$$\alpha(E)=\alpha_0\times\exp\left(-(E_0/E)^m\right) \tag{7}$$

α_0, E_0 and m being constants. The experimental character of the dependence of ionization coefficients on electric field is caused by the exponential dependence of the 'tail' of the distribution function in the high energy range. Let us consider the high energy distribution function in a simple example, when the energy losses are due to electron interaction with optical phonons of energy $h\omega$. Here we shall consider that $h\omega \gg kT$, so the phonon absorption is negligible. In moderately high fields when the following inequality is met:

$$eE\lambda \ll h\omega \tag{8}$$

where λ represents the mean free path between collisions, then electron energy accumulation in the electric field takes place in the ballistic regime. This case was qualitatively considered by Townsend for gases [40] and Shockley for semiconductors [41]. Their basic ideas were the following. The high-energy distribution function 'tail' is formed by some 'lucky' electrons, which avoided collisions and accumulated an energy sufficient for impact ionization. The probability for an electron for traveling along the length x to avoid collision is proportional to $\exp(-x/\lambda)$. Such an electron in the electric field acquires an energy $\varepsilon=eE\lambda$. Then we have for the high-energy distribution function

$$f\sim\exp(-\varepsilon/(eE\lambda)) \tag{9}$$

It is natural that in this case the electron distribution in momentum space should be anisotropic and stretched along the electric field direction. Equation (9) holds only when the electron mean free path is energy independent. In the simplest case of scattering on deformations the optical phonon distribution function can be written as (see [6])

$$f(\varepsilon, E)\sim\exp(-3h\omega\varepsilon/(eE\lambda)^2) \tag{10}$$

Substituting Equations (9) and (10) in Equation (1) for the electron ionization coefficient, we get for the two limiting cases

$$\alpha(E)=\alpha_{01}\exp(-E_{01}/E); \qquad E\lambda \ll h\omega \tag{11}$$
$$\alpha(E)=\alpha_{02}\exp(-(E_{02}/E)^2); \qquad eE\lambda \gg h\omega \tag{12}$$

The pre-exponential factors α_{01} and α_{02}, have a power dependence on the electric field. This dependence is caused by the inter-relationship of the impact ionization probability and the energy near the ionization threshold, as well as by the dependence of the distribution function pre-exponential factor on electric field.

Experimental data on the field dependence of electron and hole ionization coefficients in semiconductors indicate the diffusional regime (of Davidov-Wolf type) for fields other than those meeting the criterion at Equation (8) [42,43].

If for example, the mean free path $\lambda \approx 5$ nm and $h\omega \approx 0.03$eV, then, Equation (8) is fulfilled for fields up to $E < 6 \times 10^4$ V/cm only, while the ballistic regime is observed in GaAs up to fields of $E \approx 10^5$ V/cm. This observation triggered off a number of theoretical works in which attempts were made to obtain the ballistic regime from the kinetic equation and to establish its exact criterion.

Earlier in Baraff's work [44] a numerical calculation of the distribution function was performed on the basis of the kinetic equation taking into account the scattering on lattice deformations, the mean free path being constant. Here, in the limiting case of weak fields a relation of the type at Equation (9) was obtained. The constant mean free path and the threshold ionization energy were considered in Baraff's theory as fitting parameters. The numerical nomograms of the ionization coefficient calculated by Baraff for a long time served as a basis for the comparison of experiment with theory. This was performed by determining independently the direct impact ionization threshold and by fitting the experimentally measured coefficients with Baraff's nomograms. This enables one to obtain the mean free path λ. Such a fitting has been made even for III-V semiconductors with essentially non-parabolic energy spectrum and energy dependent mean free path which is characteristic of polar semiconductors. We will consider Baraff's nomograms in Section 4 for analysis of experimental ionization coefficients in some III-V materials.

2.4 Two-valley Model

Gribnikov [45] was the first to consider the heating of charge carriers by the electric field in a semiconductor with two valleys (sub-bands) with essentially different (light and heavy) mass. In this model, the ballistic regime appeared to be prolonged up to electric fields considerably higher than required by Equation (8). In [46] a quasi-drift carrier heating mechanism was suggested which also leads to a field dependence of the ionization coefficient $\alpha(E)$ that is ballistic as in Equation (11). Such a mechanism of heating is realized in comparatively weak fields. We shall now consider briefly the two-valley model.

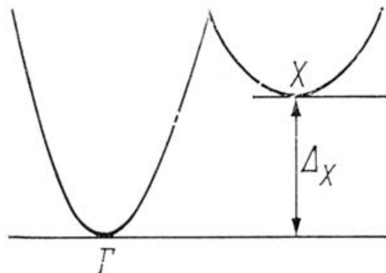


Figure 2. Model of a semiconductor with two valleys

Consider a semiconductor (GaAs type) with two valleys Γ and X (see Figure 2), the second valley (X) being displaced in energy with respect to the first one by the value Δ. According to [45] in a multi-valley semiconductor the ballistic heating mechanism changes into the diffusional one at appreciably higher electric fields as compared with the single valley case. It should be emphasized that owing to the connection between the valleys, the energy dependence of the distribution functions f_Γ and f_X is determined by that of the valleys in which it is easier for an electron to reach higher energy. In valley Γ only the ballistic mechanism of energy accumulation can be realized, since when emitting a phonon an electron passes into valley X and is able to return into valley Γ only after many collisions (of the order of density of states in both valleys, where $g_X/g_\Gamma >> 1$). In weak fields, satisfying the criterion $eE\lambda_X<3h\omega_X$, in ballistic heating also takes place in valley X. However, according to Equation (8), in such fields the high-energy distribution function is defined just by valley Γ. Consequently, the energy dependence of functions f_Γ and f_X is proportional to the factor

$$\exp(-\int\frac{d\varepsilon'}{eE\lambda_\Gamma(\varepsilon')}) \qquad (13)$$

The distribution in a 'light' valley (Γ) is strongly anisotropic and expands along the electric field direction. The dependence of the distribution function on energy in both valleys is the same and determines the valley where electrons can acquire the higher energy. It was shown in [6, 48] that in the two-valley case ballistic heating of electrons is prolonged up to higher electric fields:

$$eE\lambda_\Gamma < 3h\omega_X\,(\lambda_\Gamma/\lambda_X)^2 \qquad (14)$$

where λ_Γ and λ_X are the mean free paths in the Γ and the X valleys, respectively. In very strong fields when the Equation (14) is not valid, the asymptotic of the distribution function is controlled by the diffusion heating in the heavier valley (X). This case is important for understanding anisotropy of impact ionization coefficients in multi-valley semiconductors such as GaAs, InP, GaSb, and their alloys [6,31,39,48].

2.5 Inter-relationship between Multiplication Coefficients and Ionization Coefficients of Electrons and Holes

Let us consider the connection between multiplication coefficients and ionization coefficients (see for example, [1, 21, 22]). Impact ionization coefficients are usually found in structures with abrupt p-n junctions, heterostructures, or Schottky diodes. With a sufficiently large reverse bias applied to these structures, electron-hole pairs are generated due to impact ionization by charge carriers, which have been heated in the electric field. Owing to this fact the device current rises considerably. This phenomenon is called avalanche multiplication. Ionization coefficients are determined as the ratio of the output multiplied current (or photocurrent) to the input

initiating current: for electrons $M_n=I_n(W)/I_n(0)$, for holes $M_p=I_p(W)/I_p(0)$, W being the width of the space charge region.

Experimentally it is customary to measure the coefficients of avalanche photocurrent multiplication under separate illumination of p-n structures from p- or n-sides with strongly absorbed light, thus obtaining the current multiplication initiated by electrons or holes, respectively. Earlier, experiments were made by Vul and Shotov [50]. The corresponding relation between the multiplication and ionization coefficients established in [51] has the following form:

$$1-\frac{1}{M_n}=\int\alpha(x)\exp(-\int(\alpha-\beta)dx')dx \tag{15}$$

$$1-\frac{1}{M_p}=\int\beta(x)\exp(-\int(\alpha-\beta)dx')dx \tag{16}$$

From Equations (15) and (16), assuming that ionization coefficients are locally available (*i.e.* the ionization coefficients at a point are determined by the magnitude of the electric field at the same point) and assuming that the electric field in the p-n junction is determined by

$$E(x)=E_m(W-x)/W \tag{17}$$

which is correct for abrupt p-n junctions, one can obtain the following equations:

$$\beta(E_m)=\frac{1}{M_n}\frac{dM_p}{dW} \tag{18}$$

$$\alpha(E_m)=\beta(E_m)+\frac{d}{dW}\ln\frac{M_n(W)}{M_p(W)} \tag{19}$$

E_m being the maximum electric field in the p-n junction. The problem of local ionization coefficients was considered in [52]. When studying the processes of impact ionization in inhomogeneous structures, in order to apply the developed theoretical approach it is also necessary that other characteristic lengths (mean free path λ, transport length λ_T, impact ionization hot carrier path λ_i) are essentially smaller than the characteristic scale of the electric field variation. In abrupt p-n junctions based on narrow-gap semiconductor compounds of InAs and InSb type, in which the above-mentioned path lengths are especially large, this condition is often violated. In InSb, $\lambda_i \approx 10^{-2}$ cm, and electrons can pass the narrow space charge region without producing any ionization. Generation of electron-hole pairs occurs in a quasi-neutral region and consequently it has been shown in [53] that an anomalous dependence of breakdown voltage on free carrier concentration in the base of a p-n junction can be observed.

2.6 Noise and Response Speed of APDs

Maximum gain, response speed, multiplication coefficient (M) and bandwidth (B) product M×B, as well as noise of APDs depends strongly on the electron/hole ionization coefficient ratio α/β [21–23]. Statistical parameters of avalanche

multiplication processes were considered independently by McIntyre [21] and Tager [22]. The noise of an APD per unit bandwidth can be described by [21]

$$\langle i^2 \rangle = 2eI_{ph} \langle M^2 \rangle F \tag{20}$$

where e is electron charge, I_{ph} the un-multiplied photocurrent (signal). <M> is the average avalanche gain and F is the excess noise factor. If the avalanche process was deterministic, *i.e.* every injected photo-carrier undergoes the same gain M, the excess factor F would be unity (F=1) and the resulting noise would be only multiplied input shot noise due to the random arrival of photons. The avalanche process is intrinsically statistical in nature, so that individual carriers in general have different gains characterized by a distribution with an average <M>. Tager [21] showed that the spectral density of avalanche noise is M times more in the case of equal ionization coefficients than for the case when one coefficient is considerably larger than other. If $\alpha=\beta$ there is a positive feedback effect in an avalanche which amplifies noise fluctuations and causes additional noise. McIntyre [22] deduced simple relations between the excess noise factor F and the ionization coefficient ratio $k=\beta/\alpha$ for a p-i-n diode with a uniform electric field. The excess noise factor depends strongly on the ratio of ionization coefficients for holes and electrons β/α.
For electron injection

$$F_n = M_n \left[1 - (1-k) \left(\frac{M_n - 1}{M_n} \right)^2 \right] \tag{21}$$

and for hole injection

$$F_p = M_p \left[1 - (1 - \frac{1}{k}) \left(\frac{M_p - 1}{M_p} \right)^2 \right] \tag{22}$$

where M_n and M_p are the electron and hole multiplication factors, respectively. Figure 3 demonstrates calculated spectral density of excess noise factors for various relations between ionization coefficients (using multiplication coefficient *vs* electric field density for various ratios of ionization coefficients from [21]). It is evident that low avalanche noise is obtained when the ratio of ionization coefficients β/α is either very large or very small and when the multiplication process is initiated by the carrier with the highest ionization coefficient.

To create very low-noise APDs the case of monopolar gain is most important. The breakdown condition in the monopolar case differs from the case of $\alpha=\beta_w$. For $\alpha=\beta$

$$\int \alpha dx = \alpha w = 1, \tag{23}$$

where w is the width of the space charge region, and when $\alpha/\beta >> 1$

$$M_n = \exp\left(\int \alpha dx \right) = \exp(\alpha w) \tag{24}$$

This means that there are not any breakdown in this case and the multiplication coefficient rises exponentially with increasing electric field. The case of monopolar gain is interesting because it leads to better stability of the avalanche processes in APDs (See Figure 4) [23].

Response speed of APDs was discussed for example in [21, 61]. It was shown that the product of multiplication coefficient and bandwidth, M×B, depends on carrier transit time through the space charge region $\tau_i = w/v_{drift}$, drift velocity and electron-hole ionization coefficient ratio $k=\alpha/\beta$:

$$M \times B = \frac{1}{\tau_i} \sim \frac{1}{(w/v_{drift}) \times (\beta/\alpha)} \qquad \text{for } \alpha/\beta >> 1 \tag{25}$$

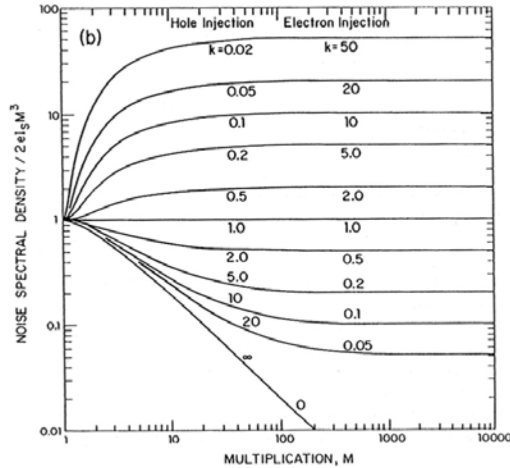


Figure 3. The dependence of excess noise factor on average multiplication coefficient at various ratios $k=\beta/\alpha$ (electron injection) and $k=\alpha/\beta$ (hole injection) [22]

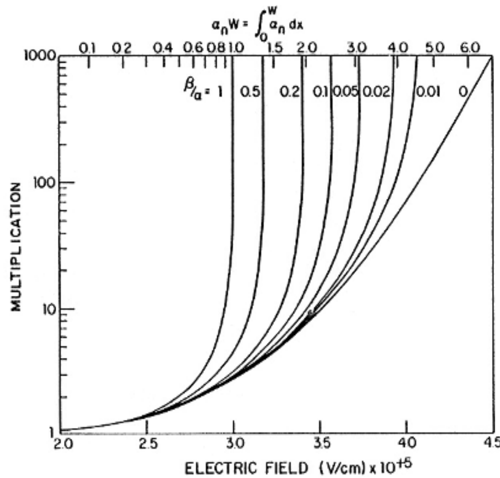


Figure 4. Multiplication coefficients *vs* electric field for various ratios of ionization coefficients $k=\alpha/\beta$ [21]

Note that in most III-V semiconductors and their solid solutions (GaAs, InP, InGaAs(P) and others) electron and hole ionization coefficients are similar in contrast to Si [37,55]. This prevents the realization of low-noise APDs based on these materials. For that reason during the last two decades, many efforts were made to research and study semiconductors with large ionization coefficient ratios [1-6]. As was demonstrated in [4], the most appropriate materials are GaSb, InAs, and their solid solutions which allow the design of APDs for the spectral range 2-5 μm.

3 Ionization Coefficients in III-V Semiconductors and their Alloys

3.1 Electron Impact Ionization

3.1.1 Anisotropy of Ionization Coefficients in Multi-valley Semiconductors of GaAs, InP Type

In several III-V semiconductors, the band structure at high energies is strongly anisotropic [56]. There are subsidiary X and L valleys in GaAs, InP and GaSb lying higher than the bottom of Γ-valley in the (100) and (111) directions. The main carrier scattering mechanism is intervalley scattering leading to redistribution of electrons between valleys with increasing electric fields. It was shown in [6,48] that in GaAs like semiconductors the X-valley plays an important role in high electric fields. The dominant scattering of electrons with energy $\varepsilon > \Delta_{\Gamma X}$ into the X-valley (see Figure 2) leads to ballistic heating in the Γ-valley and diffusion heating in the X-valley.

Pearsall *et al.* [39] were the first to observe the orientation dependence of electron ionization coefficients in GaAs p-n junctions which were prepared on substrates oriented along (100), (110) and (111). In the region of weak and moderate fields, the ionization coefficients were different for different orientations, and their functional dependencies on electric field were not adequately described. For E|| [111] the values of the ionization coefficient were lower than for E||[100]. For the sake of convenience, experimental points taken from [39] have been represented in Figure 5 where the solid lines are the result of theoretical calculations from [6]. As seen from Figure 5, the character of the dependence of ionization coefficient on electric field is in accordance with theory. For E||[111] the beginning of the square law dependence $\ln\alpha \sim (E_{02}/E)^2$ is already observed in fields $E \approx 3.6 \times 10^5$ V/cm, whereas for E||[100] in the same fields $\ln\alpha \sim (E_{01}/E)$. It is interesting that Pearsall *et al.'s* work [39] on experimental determination and interpretation of the anisotropy of the electron ionization coefficient in GaAs stimulated theoretical investigations of the process of
impact ionization in this semiconductor.

Hess *et al.*[49] calculated the distribution function using the Monte-Carlo method taking account of the real band structure of GaAs. These authors did not obtain any anisotropy of the electron ionization coefficient in GaAs. Their

calculations were based on the assumption that the electron mean free paths in Γ and X valleys were nearly the same, $\lambda_{\Gamma} \approx \lambda_X$. It followed automatically from this assumption that the diffusional regime of electron heating and an isotropic distribution function must be observed. The weak anisotropy of the ionization coefficient of electrons and holes in InP was obtained, *e.g.*, in [47]. The orientation dependence of electron ionization coefficients for two directions E||[100] and E||[110] was considered. In a wide range of electric fields (up to 6.6×10^5 V/cm) for both electrons and holes a linear dependence of lnα, lnβ on 1/E is observed. In the framework of the theoretical approach the situation for the electrons in InP should not noticeably differ from that in GaAs.

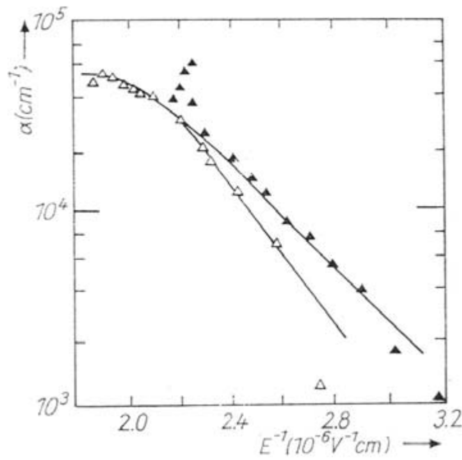


Figure 5. Anisotropy of ionization coefficients in GaAs. Experimental points E||(100); E||(111), T=300 K. After Pearsall *et al.* [3]. *Solid curves* — theoretical calculation [6].

3.1.2 Anisotropy of Ionization Coefficients in Multi-valley Semiconductors of GaSb and their Alloys

A peculiarity of the GaSb band structure is that the energy of the fourfold degenerate L valley is very close to the minimum of the central Γ valley (Δ_L=0.085 eV, Δ_X=0.5eV [56]; see Figure 6. The electron effective mass at the L-valley bottom is far larger than the effective mass in the Γ-valley ($m_L=0.5m_0$, $m_\Gamma=0.043m_0$, m_0 being the mass of a free electron). So, a considerable number of electrons will be in the L valley at room temperature [6]. The availability of the L-valley, in which there are no states at energies $\varepsilon \approx E_g$, does not influence the character of the exponential dependence, like that in GaAs, but leads to an essential reduction in the pre-exponential factor α_{01}, as a portion of charge carriers is present in those valleys, and thus cannot take part in the process of impact ionization. Estimations show only a few per cent of the total number of electrons remain in the Γ-valley: (using $n_\Gamma=1/(1+4(m_L/m_\Gamma)^{3/2}\exp(-\Delta_{\Gamma L}/kTe)$, where $kTe\approx3h\omega$, $h\omega_0=0.029$ eV for GaSb [4]).

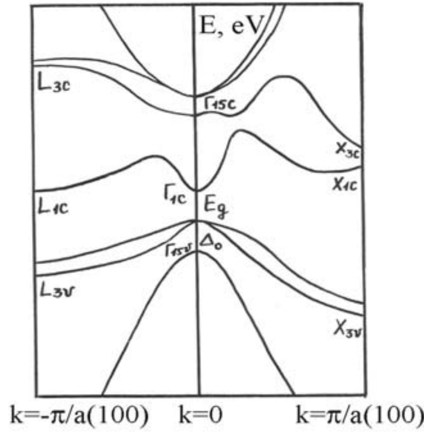


Figure 6. Band structure of GaSb [56]

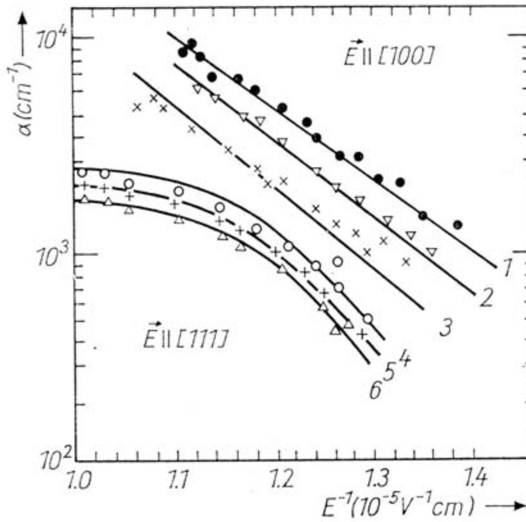


Figure 7. Electron ionization coefficient orientation dependence in $Ga_{1-x}Al_xSb$, T=77 K. (after Zhingarev *et al.* [32] 1-3 E||(100), 4-6 E||(111)

This results in a strongly decreasing electron ionization coefficient for such materials and thus in more pronounced anisotropy. The dependence of electron ionization coefficient on orientation and composition in $Ga_{1-x}Al_xSb$ alloys was studied in [31], and there an anisotropy of the electron ionization coefficient was found. Figure 7 shows the field dependencies of the ionization coefficient in these materials for two directions of the electric field with regard to the crystallographic axes. These curves have quite different shapes, the electron ionization coefficient for E||[111] lying considerably lower than for E||[100]. In $Ga_{1-x}Al_xSb$ electron ionization

coefficient ratio for the two orientations was found to be $\alpha_{[100]}/\alpha_{[111]} \approx 3.5$, while in GaAs this ratio was close to 1.5 [39].

3.2 Electron Impact Ionization in Semiconductors of InAs, InSb Type

In III-V semiconductors of InSb, InAs type and their alloys the energy from the minimum of the central Γ valley to the bottom of the nearest subsidiary valley in the conduction band considerably exceeds the energy gap E_g [56] (see Figure 8). In these cases, the presence of the higher valleys does not affect the form of the distribution function for energies lower than E_g, and consequently, the impact ionization processes take place in the central Γ-valley.

Dumke [59] was the first to calculate generation rates and impact ionization coefficients in InAs and InSb for the ballistic regime. The solution of the kinetic equation with Kane's spectrum under polar scattering, and the transition from ballistic heating to diffusional heating, were given in [6]. Figure 9 shows a calculated curve for the electron ionization coefficient in InSb from [6]. In the weak field range $\ln \alpha \sim E_{01}/E$, where $E_{01} \approx 5\times10^3$ V/cm. The same value of E_{01} was obtained in [59]. In fields $E \geq 5\times10^2$ V/cm the transition to $\ln \alpha \sim (E_{02}/E)^2$ must be observed, where $E_{02}=1.6\times10^3$ V/cm.

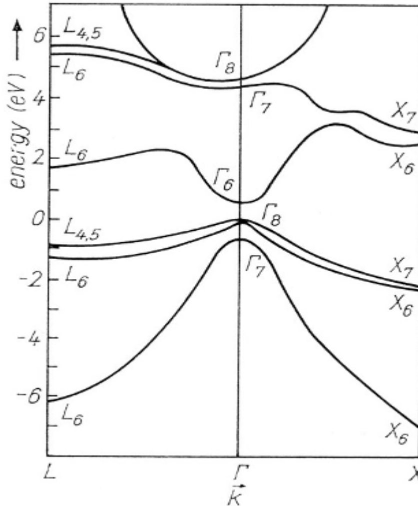


Figure 8. Band structure of InAs [56]

Electron and hole ionization coefficients have been investigated in InAs and InSb in p-n junctions [30,31,62,63]. In InSb at fields $E \geq 10^4$ V/cm saturation of the dependence of the electron ionization coefficient on electric field was observed [48] (see Figure 9). At the same time, the hole ionization coefficient could not be

determined. It was only reported to be considerably lower than that for electrons [62]. In InAs p-n junctions the electron and hole ionization coefficients depend linearly on electric field ($\ln(\alpha,\beta) \sim E_{01}/E$) in a wide range of electric fields ($E=(2\text{-}8) \times 10^4$ V/cm) [30]. Hole ionization coefficients turned out to be much higher than those for electrons (see Section 3.2.3). Impact ionization in InAs and InSb was studied in pulsed electric fields as well as in the presence of a transverse magnetic field [64-66].

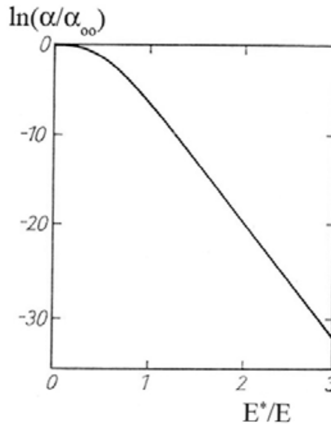


Figure 9. Calculated dependence of electron ionization coefficient in InSb on the reduced electric field E*/E, where E*=539 V/cm [48]

The characteristic pulse duration in those experiments was 10^{-7} to 10^{-9}s which considerably exceeds the electron transit time in the space charge region in p-n junctions. Hence noticeable generation of electron-hole pairs is observed in considerably weaker electric fields as compared with p-n junctions, *e.g.* in InSb at $E \approx (2 - 3) \times 10^2$ V/cm and in InAs at $E \approx (0.8 - 1) \times 10^3$ V/cm [64, 66].

3.3 Hole Impact Ionization

Experimental data indicate that in quite a few $A^{III}B^{V}$ semiconductors (GaAs, InP, GaAsSb, GaInAs, InGaAsP) and in a wide range of electric fields the hole ionization coefficients $\beta(E)$ depend on electric field obeying the ballistic law, $\ln\beta \sim E_{01}/E$. Here it is interesting that the values of the hole ionization coefficients in different materials are similar. This makes it possible to assume that the behavior of hot holes in $A^{III}B^{V}$ semiconductors is controlled by a general scattering mechanism. The valence bands of $A^{III}B^{V}$ semiconductors consist of three twice-degenerate sub-bands with the same centre at the Γ point (see Figure 4). There are two sub-bands of heavy and light holes and a spin-orbit split sub-band separated by energy Δ. Depending on the ratio of E_g and Δ, $A^{III}B^{V}$ semiconductors can be classified into three types:

$\Delta << E_g$, (GaAs, InP), $\Delta >> E_g$ and $\Delta \approx E_g$(InAs, GaSb). The dependence of the hole ionization coefficient on electric field in GaAs, InP and their alloys was experimentally studied in [39, 47]. In a wide range of electric fields the field dependence of hole ionization coefficient appeared to obey the ballistic law $\ln\beta \sim E_{01}/E$. The magnitude of the hole coefficient did not depend on the electric field direction with respect to crystallographic axes [47]. The absence of hole ionization anisotropy in GaAs and InP is well consistent with the idea that the distribution function is formed as a result of the hole travelling along the light spin-orbit split sub-band [6].

3.4 Hole Impact Ionization in Semiconductors with Band Gap "Resonance", $E_g=\Delta$ (InAs, GaSb)

Let us consider the most interesting case of hole impact ionization in semiconductor compounds with band "resonance", *i.e.* in which the energy gap is equal to the value of the spin-orbit splitting of the valence band, $E_g=\Delta$ (InAs, GaSb, etc). From the first investigation of avalanche multiplication in an InAs p-n junction, the hole multiplication coefficient M_p was found to be considerably larger than that of electrons, M_n. In 1976, on the basis of experiments [30, 67, 68] devoted to the dependence of the charge carrier ionization coefficient on field and the noise in InAs and InGaAs avalanche photodiodes, an idea (confirmed later by theoretical calculation of [69]) was suggested that in materials of InAs, GaSb type the direct impact ionization is performed by holes from the spin-orbit split off valence band

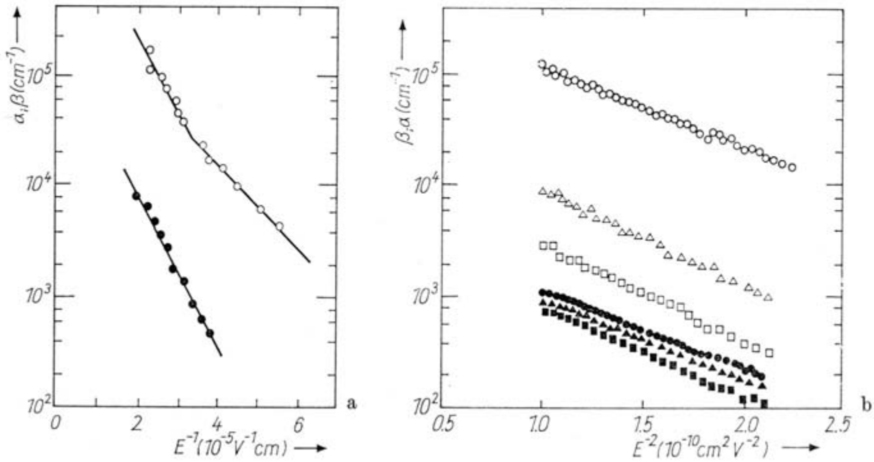


Figure 10. Electron and hole ionization coefficients in **a**) InAs [30] ○–β, ● –α; **b**) $Ga_{1-x}Al_xSb$ (E||(111)) [32]; *filled symbols* - α, *empty symbols* - β; ○, ● - x = 0; △, ▲ - x = 0.07; □, ■ - x = 0.23

with the minimum threshold energy $\varepsilon_{th}=\Delta=E_g$. The Coulomb interaction of the charge carriers responsible for impact ionization is connected, in this case, with small momentum transfer, which essentially increases the probability of such a process. In materials with $E_g\approx\Delta$ the hole ionization coefficient increases sharply. Figure 10a,b shows the dependence of the hole and electron ionization coefficient on the electric field in InAs and GaAlSb at 77 K. It is seen that here the ionization coefficient ratio is large and β exceeds α by more than one order of magnitude. In GaSb and its alloys two characteristic features of band structure produce a lucky combination: the presence of the low-lying heavy L valley "pumping" hot electrons from the Γ valley, and, on the other hand, the band "resonance" which facilitates hole impact ionization. This makes it possible to reach practically monopolar hole multiplication [31,32]. As illustrated by Figures 10b and 12 the maximum value of the ratio $\beta/\alpha\approx100$ was obtained at T=77 K in $Ga_{1-x}Al_xSb$ compounds similar to GaSb with the electric field orientation E||[111]. As the composition of $Ga_{1-x}Al_xSb$ changes, the values of E_g, Δ_X, Δ_L, Δ are also changed. The L-valley minimum slightly lowers, while the difference $|E_g-\Delta|$ increases. In this case the situation for electron ionization differs insignificantly whereas for hole ionization it becomes worse due to band "resonance" violation which results in decreasing β/α ratio. At electric field orientation E||[100], when the L valley is of no importance, a high ratio of the ionization coefficients β/α can be attained only due to band "resonance".

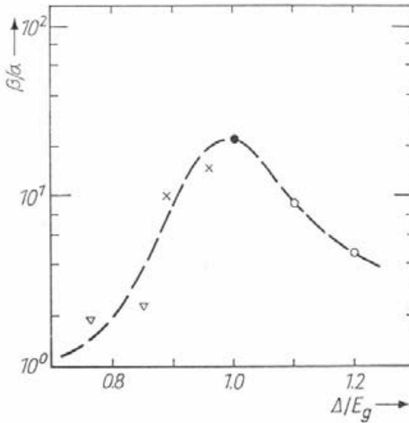


Figure 11. Hole and electron ionization coefficient ratio *vs* Δ/E_g in InAs and its alloys [36]: ● - InAs, ✕ — $In_{1-x}Ga_xAs$ (x=0.02, 0.04); ○ — $InAs_{1-x}Sb_x$ (x=0.06, 0.08); ▽ — InAsSbP

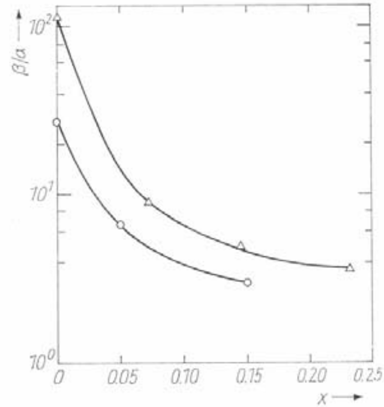


Figure 12. Orientation dependence of β/α on composition of $Ga_{1-x}Al_xSb$ [32].
△- $E_{||}$ (111), ○- $E_{||}$ (100)

There have been a number of works dealing with the "resonant" impact ionization in solid solutions based on GaSb [31-33] and InAs [34-36]. It is clear that impact ionization probability has its maximum value at band "resonance" $E_g=\Delta$, and should decrease as Δ/E_g changes. Such results were obtained in [34-36] where the

ionization coefficient ratio β/α, was studied as a function of the value Δ/E_g for various compositions, orientation and temperatures in some $A^{III}B^V$ solid solutions. Figure 11 represents a "resonant" curve β/α *vs* Δ/E_g for InAs and its alloys (InGaAsSb, InAsSb, InAsSbP) [36].

Figure 12 shows the dependence of β/α on x in $Ga_{1-x}Al_xSb$ alloys of different orientations [33] and Figure 13 gives Hilderbrandt *et al.*'s data on "resonant" impact ionization in such materials [33]. It should also be mentioned that in the semiconductors in question the probability of inverse processes, *i.e.* Auger recombination with participation of spin-orbit splitting, is also large. Attention was drawn to this fact for the first time in [69-73].

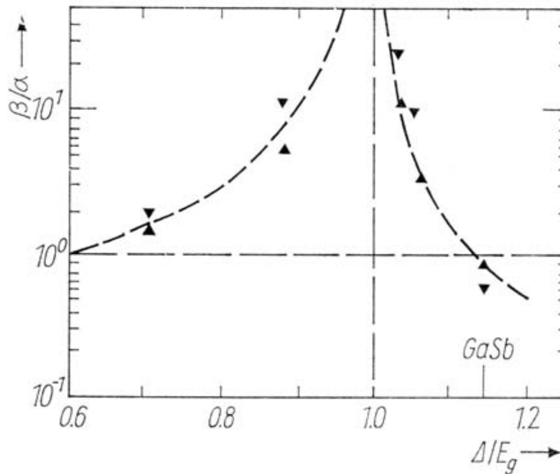


Figure 13. "Resonant" impact ionization of $Ga_{1-x}Al_xSb$ after Hilderbrandt [33]

Next we consider in detail experimental data on the investigation of impact ionization coefficients in GaAlSb, GaInAsSb as well as design and parameters of low-noise avalanche photodiodes including APDs with separate absorption and multiplication region, (SAM) APD, for the spectral range of 1.5-2.5 μm. Sensitivity of the receiver with GaInAsSb/GaAl(As)Sb APD will be also discussed.

4 Avalanche Photodiodes for the 2-2.5 μm Spectral Range

There have been many experimental measurements of ionization coefficients in III-V compounds and their alloys reported in the literature. Since the ratio of the electron and hole ionization coefficients is very important for determining the ultimate performance of avalanche photodiodes, the results of some of these measurements will be discussed here.

Avalanche photodiodes have some advantages over p-i-n photodiodes:
- High quantum efficiency in the working spectral range
- High speed response, or product of multiplication coefficient and bandwidth
- Minimum noise in this bandwidth (low level of excess noise factor)

It is known that high speed and low-noise avalanche photodiodes can only be designed based on semiconductor materials with a high ionization coefficient ratio for electrons (α) and holes (β) [5,21-23]. Preliminary studies of avalanche multiplication in diode structures based on GaInAsSb solid solutions demonstrated their availability for creation of both p-i-n and avalanche photodiodes for the spectral range of 1.6-2.4 μm [80,82]. In this Section field dependencies of ionization coefficient of GaAl(As)Sb and GaInAsSb materials are presented. Results of the noise characteristic study of APDs and the parameters of the optical receiver for high bit rate transmissions at the 2.1μm wavelength are also demonstrated.

4.1 Experimental Ionization Coefficients in Solid Solutions Based on GaSb

4.1.1 Ionization Coefficients in GaSb, GaAlSb and GaAlAsSb

Zhingarev *et al.* [32,75] have carried out a systematic study of ionization coefficients of electrons (α) and holes (β) in AlGaSb alloys. They studied the spin-orbit split-off band resonance effect in AlGaSb alloys and found the same effect in GaSb at 77 K, since at this temperature E_g=Δ=0.8 eV. In a more recent paper,

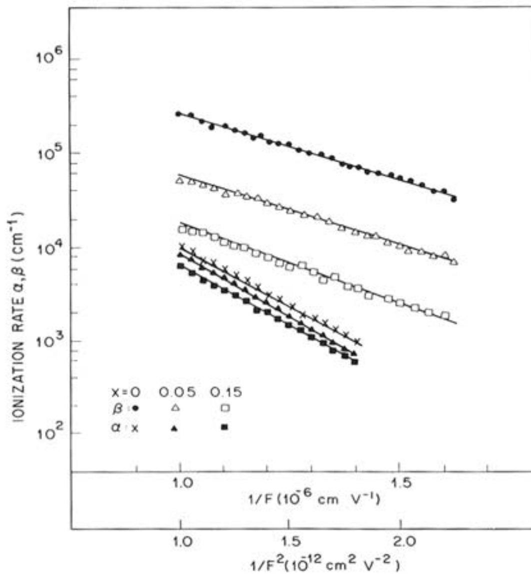


Figure 14. Electric field dependence of the ionization coefficients in <100> $Al_xGa_{1-x}Sb$ at 77K; *the upper horizontal scale* is for holes; *the lower* for electrons (from [75])

Zhingarev *et al.* [75] have investigated the hole and electron ionization rates in GaAlSb at 77 K, their dependence on <100> and <111> orientation and alloy composition. They investigated multiplication in p-n heterostructures (p-GaSb/n-Ga(Al)Sb LPE layers) and obtained pure electron or hole injection by illuminating the samples from the n- and p-sides. Figures 10b and 14 show the electric field dependence of α and β for the <111> and the <100> directions, respectively. The ratio of ionization coefficients was observed to be higher for the <111> direction, with a maximum value β/α=115.

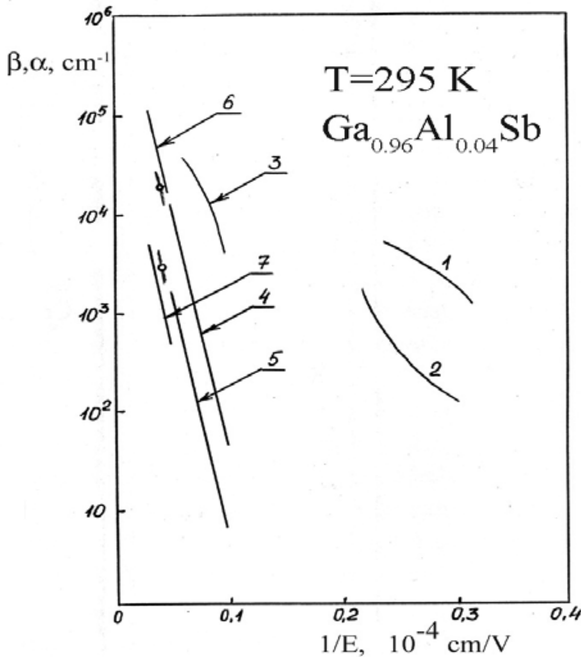


Figure 15. Electron (α) and hole (β) ionization coefficients in $Ga_{1-x}Al_xSb$ of "resonant" composition (x ≈ 0.04) as a function of the electric field: 1,2 β,α (after Hildebrand *et al.* [76]) 3 β (after Zhingarev *et al.*. [75]) 4,5 β,α (after Mikava *et al.* [77]) 6,7 β,α (after Liquet *et al.* [78])

Hildebrand *et al.* [33, 76] have independently carried out a study of α and β in GaSb and $Al_XGa_{1-X}Sb$ alloys for various values of x. Their work provides convincing evidence of the resonant enhancement of impact ionization for an aluminum content x~0.05 at 300 K. Hildebrand *et al.* realized Liquid-Phase Epitaxial (LPE) p-i-n junctions and derived α and β from a study of the multiplication of the photocurrent observed in the cases of pure hole and electron injection; this carrier injection was produced by a focused electron beam on the p and n sides of the junction. This is the ideal geometry for the determination of α, β.

More recently Mikawa *et al.* [77] investigated Zn-diffused $Ga_{0.947}Al_{0.053}Sb$ mesa diodes realized on LPE grown n-GaAlSb layers. Ionization coefficients were

deduced from the multiplication of photocurrents by frontal illumination at two wavelength values ensuring pure hole and electron injection, respectively. The variations of β and α in $Ga_{1-X}Al_XSb$(x~0.05) obtained by several research groups are shown in Figure 15 [78]. They are very different, but they confirm the fact that β/α is high:

β/α =20 for the electric field $E=3.3\times10^4$ V/cm (after Hildebrand [76])
β/α =5-7 for $E=1.0\times10^5$ V/cm<E<2×10^5 V/cm (after Mikawa [77])
β/α =60 (after Zhingarev [75])

In principle, this enables AlGaSb/GaSb based APDs to achieve a lower excess noise factor and a higher gain-bandwidth product than for InGaAs/InP APDs.

4.1.2 Experimental Ionization Coefficients in GaInAsSb

Multicomponent $Ga_{1-X}In_XAs_YSb_{1-Y}$ solid solutions lattice-matched to GaSb (y/x~0.9) are of much interest as promising materials for the fabrication of light sources and photodetectors operating in the 2-5 μm spectral range. It has been established earlier that the characteristics of the energy band structures of InAs and GaSb, namely, the similarity of the band gap and the spin-orbit splitting Δ_O (E_g~Δ_O), provide an opportunity for reaching a high value of the ratio of hole (β) and electron (α) ionization coefficients and for making avalanche photodiodes with low excess noise (see Section 3). The fabrication of the photodiodes including ultrafast p-i-n and avalanche photodiodes based on GaInAsSb/GaAlAsSb heterostructures operating in the spectral range 1.6-2.4 μm was reported in [79,82,83]. The detailed investigation of the avalanche multiplication of photocurrent and the field dependence of the ionization coefficients in $Ga_{0.80}In_{0.20}As_{0.17}Sb_{0.83}$ was carried out. The experimentally determined ionization coefficients were compared with the modified Baraff theory [81].

Authors [36] investigated the impact ionization coefficients employing heterostructures grown by LPE on <111> n-GaSb substrates doped with Te (5-7×10^{17} cm^{-3}). The active region consisted of an un-intentionally doped narrow-gap n-type $Ga_{0.80}In_{0.20}As_{0.17}Sb_{0.83}$ layer 2.3 μm thick (E_g=0.54 eV at T=300K). The wide-gap ('window') layer of p^+ -$Ga_{0.66}$ $Al_{0.34}As_{0.025}Sb_{0.975}$ (E_g=1.20 eV at T=300K) was 2.0 μm thick and doped with Ge up to $(1\text{-}2)\times10^{18}$ cm^{-3}. The carrier density in the narrow gap active layer was estimated as 2×10^{16} cm^{-3}. Mesa photodiode samples were illuminated with monochromatic light through the wide-gap GaAlAsSb layer. Pure hole or pure electron injection into the GaInAsSb layer was obtained by using illumination with radiation at wavelengths of λ=2.1 μm and λ=1.1 μm, respectively.

The electric field dependence of α and β in the GaInAsSb/GaAlAsSb heterostructure at T=230 K is illustrated in Figure 16. The experimental dependences of the ionization coefficients on the electric field ($(1.5\text{-}2.3)\times10^5$ V/cm) were exponential and could be described by the following relationship:

$$\alpha,\beta = \alpha,\beta_0 \exp[-(E_{0\alpha,\beta} /E)^2] \qquad (26)$$

where α,β_0 are ionization coefficients in an infinitely high electric field and $E_{0\alpha,\beta}$ are characteristic fields depending on the threshold ionization energy ε_i and on the mean free path of carriers λ. Estimates indicated that in the case of impact ionization by electrons and holes the strong field criterion, $eE\lambda > 3\ \varepsilon_{ph}$ ($\varepsilon_{ph} = \hbar\omega$ is the energy of an optical phonon, ε_{ph}=0.029 eV) was obeyed (see Section 2) [5]. These dependencies were used to estimate the values of the impact ionization parameters α,β_0 and $E_{0\alpha,\beta}$, which are listed in Table 1.

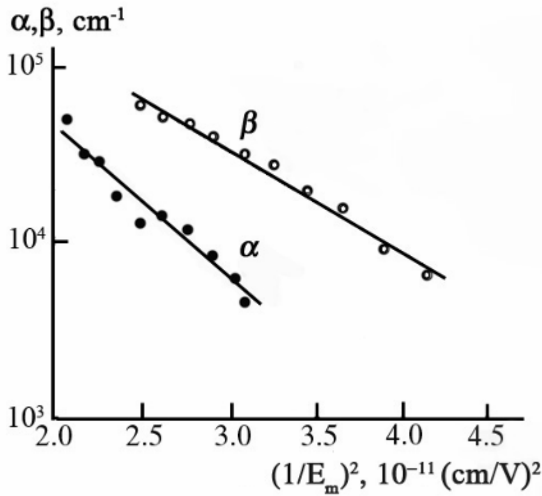


Figure 16. Dependence of the hole and electron coefficients on the square of the reciprocal maximum electric field in the $Ga_{0.8}In_{0.2}As_{0.17}Sb_{0.83}$ solid solution. T=230K

Table 1. Estimated values of impact ionization parameters

$Ga_{0.8}In_{0.2}As_{0.17}Sb_{0.83}$	E_g, eV	Δ_0, eV	α_0, cm^{-1}	$E_{0\alpha}$, V/cm	λ_e, Å
T=230K	0.57	0.72	2.41×10^6	4.45×10^5	52.3
			β_0, cm^{-1}	$E_{0\beta}$, V/cm	$\lambda_{hs\text{-}o}$, Å
			1.98×10^6	3.69×10^5	68.0

It is clear from Figure 16 that the hole ionization coefficient was greater than the electron ionization coefficient and their ratio β/α~4-5. The ionization by holes from the spin–orbit splitting valence band predominated in the range of electric fields $E=(1.5\text{-}2.3)\times10^5$ V/cm. The impact ionization threshold $\varepsilon_{i\ s\text{-}o}$ of these holes was minimal and close to Δ_O. In the case of the $Ga_{1-X}In_XAs_YSb_{1-Y}$ solid solution with x=0.2 there was a shift away from the "resonant" band. However, under these conditions the contribution of the holes from the spin-orbit-split band could still be large. On the other hand, in the case of heavy holes the direct impact ionization threshold is much higher $\varepsilon_{i\ hh}=2E_g$ (see Equation (6) in Section 2.2). The subsidiary

L-valley in GaInAsSb should be higher on the energy scale than the Γ minimum and the difference $\Delta_{\Gamma L}> 0.7$ eV, exceeds the threshold energy for electrons ε_{ie}=0.62 eV. This should increase the role of electrons in this alloy compared with GaSb and correspondingly raise the impact ionization coefficient of electrons. In the case of GaSb the electron ionization coefficient was $\alpha \sim 10^3$ cm^{-1}, whereas the GaInAsSb solid solution was characterized by $\alpha > 10^4$ cm^{-1} at the same electric field values.

The field dependence of $Ga_{0.8}In_{0.2}As_{0.17}Sb_{0.83}$ ionization coefficients was compared with the modified Baraff theory [81]. Figure 17 represents the graphical dependences of the product of the impact ionization coefficient and the mean free path $\alpha\lambda$ on the ratio $e_i/eE\lambda$. The key parameter of these curves is $<e_{ph}>/\varepsilon_i$, which is the ratio of the average energy of an optical phonon to the threshold ionization energy. The average energy lost as a result of phonon scattering $<\varepsilon_{ph}>$ and the mean free path of a carrier after the emission of an optical phonon λ are given by the following relationships:

$$\lambda = \lambda_O \tanh (\varepsilon_{ph}/2kT) \quad (27)$$

$$<\varepsilon_{ph}>= \varepsilon_{ph} \tanh (\varepsilon_{ph}/2kT) \quad (28)$$

where ε_{ph} is the optical phonon energy and λ_O is the mean free path at absolute zero.

The experimental values of the ionization coefficients of $Ga_{0.8}In_{0.2}As_{0.17}Sb_{0.83}$ solid solution correspond to Baraff curves quite well if we select the electron mean free path to be λ_e =32.5 Å and for holes in the spin–orbit split valence band λ_{hs-o}= 42.5 Å.

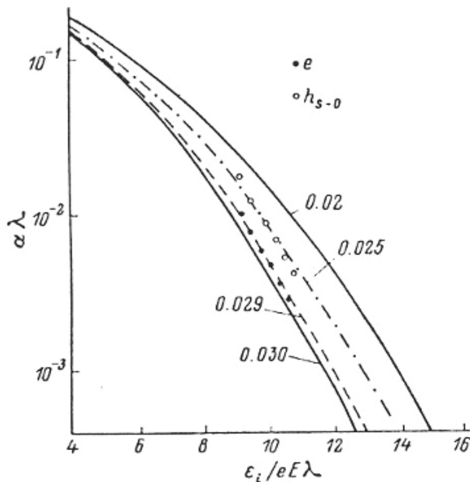


Figure 17. Dependence of the product of the ionization coefficient and mean free path $\alpha\lambda$ on $\varepsilon_i/eE\lambda$ in the $Ga_{0.8}In_{0.2}As_{0.17}Sb_{0.83}$ solid solution. T=230K. *Solid curves* are calculated after Baraff [81] with a parameter $<\varepsilon_{ph}>/\varepsilon_i$. *Dark and open circles* are experimental data (after [36])

The mean free paths found by comparison with the Baraff curves differ by a factor of approximately 1.5 which is in satisfactory agreement with experiment using the characteristic field E_0 and the expression $\lambda=[3\varepsilon_{ph}\,\varepsilon_i/(eE_O)^2]^{1/2}$ (see Table 1).

4.2 Dark Current in APDs with a 'Resonant' GaAl(As)Sb Composition

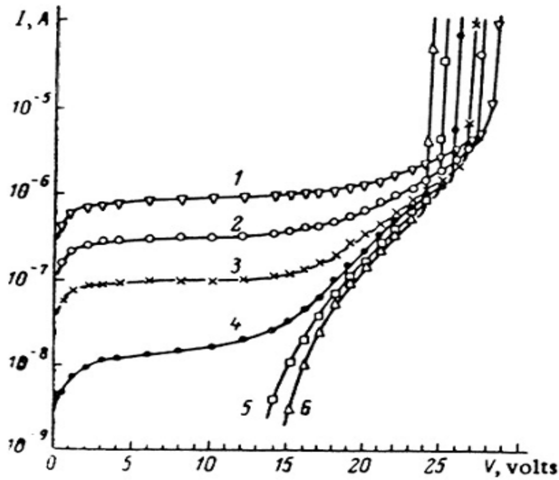
Low dark current in an APD is important because the multiplication process is associated with excess noise. However, extremely low dark current is not essential for applications of p-i-n detectors in wide bandwidth detection systems [1,5]. Low dark current requires a low defect-density material and a high quality defect-free p-n junction. The GaAl(As)Sb quaternary alloys are lattice-matched to GaSb and it should be possible to fabricate a low defect material and suitable p-n junction as well as a heterojunction with low leakage current. However, it has been shown by several research groups working with GaAl(As)Sb that, even if a high-quality defect-free p-n heterojunction is formed in narrow band-gap material, large dark current can still occur due to Zener breakdown or tunneling [84,85].

Surface leakage currents have been blamed for the large dark currents often observed in AlGa(As)Sb p-n junctions [86]. Recent improvements in crystal growth, processing techniques including sulfide passivation [87], and AlGa(As)Sb APD device structure improvements have sufficiently reduced the surface leakage component so that it is no longer dominant [88].

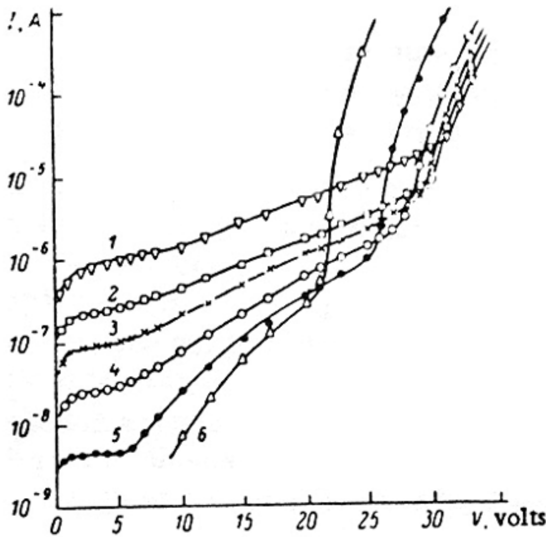
An investigation was made of the nature of the dark current in p-n GaAlSb(As) diode structure with the "resonant" composition ($E_g=\Delta_0$) [85]. Measurements at low temperature show that the reverse current for the APDs are dominated by tunneling. Figure 18 demonstrates the change of the dark current with reverse bias voltage of typical diodes based on (a) a p-n GaAlSb homostructure and (b) a heterostructure at different temperatures. The device structures used for these measurements are presented in Table 2.

Table 2. Device parameters of homo- and heterostructure APDs

Type of structure / Parameters	Type 1 n-GaSb<111>- n-GaSb- n-$Ga_{0.94}Al_{0.06}Sb$- p - $Ga_{0.94}Al_{0.06}Sb$	Type 2 n-GaSb<111>- n-GaSb- n-$Ga_{0.94}Al_{0.06}Sb$- p- $Ga_{0.7}Al_{0.3}Sb_{0.975}As_{0.025}$
E_g,eV	0.8	0.8/1.1
A, cm^2	1.6×10^{-3}	1.3×10^{-4}
N_B, cm^{-3}	1.5×10^{16}	4.5×10^{16}
V_B, V	24.0	12.0
$j_{0.5VB}$, A/cm^2	1.3×10^{-3}	8×10^{-3}
$j_{0.5VB}$, A/cm^2	4×10^{-3}	2.6×10^{-2}
τ_{eff}, ns	3.0	0.6



a)



b)

Figure 18. Current-voltage characteristics of (**a**) $Ga_{0.94}Al_{0.06}Sb$ p-n structure and (**b**) $Ga_{0.94}Al_{0.06}Sb/Ga_{0.66}Al_{0.34}Sb_{1-y}As_y$ heterostructures at various temperatures; T,K: 1-293; 2-273; 3-253; 4-233; 5-173(a), 213(b); 6-77(a), 193(b)

At low voltages ($V<0.5V_B$, where V_B is the breakdown voltage) the reverse current was due to generation in the space charge layer. Analysis shows that initially the dark current increases slowly with applied bias ($I \sim W$, where W is the thickness of the space charge region). Its value is given by

$$I_{g\text{-}r} = q\, n_i\, A\, W / \tau_{eff} \tag{29}$$

Here, n_i is the intrinsic carrier density in the selected material; A is the diode area; q is the electronic charge and τ_{eff} is the effective lifetime of the minority carriers. Using this relationship and the room temperature values of the current, the lifetimes τ_{eff} were found in the range of 0.6-3 ns (see Table 2). In diodes in which the concentration of un-intentional acceptors was reduced (as in Type I structures grown at a lower temperature of 400ºC) the reverse current density was several times smaller and the effective lifetime longer.

At higher voltages ($V \sim (0.5\text{-}0.9)V_B$) the excess dark current was found to be due to the interband tunneling *via* deep centres in the space charge layer. The magnitude of the interband tunnel current depends on the electric field E and on the tunneling energy ε (see, for example, [89])

$$I_t = I_O \times \exp\left(-\frac{4(2m_e)^{1/2}}{3q\hbar E}\, \varepsilon^* \right) \tag{30}$$

where m_e is the mass of the tunneling particle and ε^* is the tunneling energy or the height of the tunnel barrier, which for the interband case should be equal to the band gap E_g, whereas in the case of tunneling involving impurities it should be equal to the energy of the impurity. The interband tunneling of carriers *via* local levels in the band gap is considered in [90]. Expressions were obtained for the tunnel currents allowing for the complex structure of the valence band of III-V semiconductors and for the Kane nature of the spectrum. It was shown that the effective tunneling energy differs from the real energy level position. The value of effective tunneling energy was found to lie within the range of 0.3-0.45 eV, which is close to $E_g/2$. The pre-exponential factors I_O should be different for tunneling via centres of different types of symmetry: l-c center (Γ_6 symmetry) or h center (Γ_8 symmetry).

The breakdown in p-n GaAl(As)Sb structures was due to the avalanche process [85]. The temperature coefficient of the breakdown voltage was determined in the temperature range of 80-400 K and its value was found to be $\gamma = (1.1\text{-}1.5)\times10^{-3}\ K^{-1}$. The breakdown voltages were within the range V_B=10-50 V and the photocurrent multiplication coefficients were M=20-80. The reverse current density in p-n structures based on GaAlSb(As) was $j=1.3\times10^{-3}\ A/cm^2$ for V=0.5 V_B and $j=4.0\times10^{-3}\ A/cm^2$ for V=0.9 V_B, respectively. It was found that a reduction in the concentration of the unintentional acceptors reduced the generation current and narrowed the range in which the excess tunnel current was observed.

4.3 GaInAsSb/GaAlAsSb Avalance Photodiode with Separate Absorption and Multiplication Region (SAM APD)

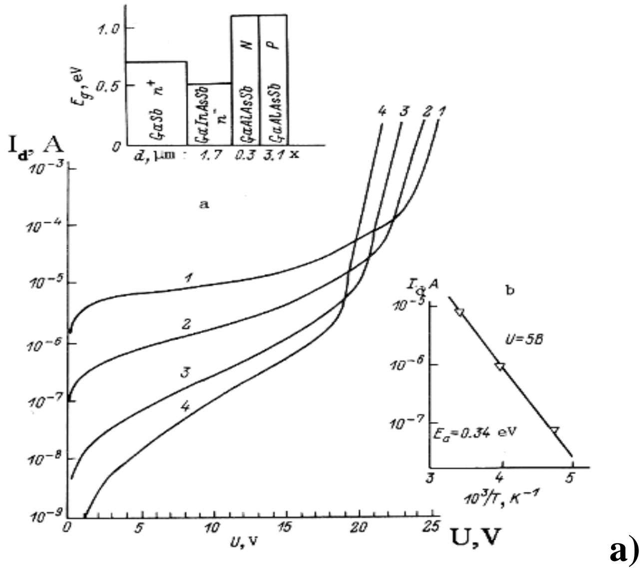
In this Section we consider the design and some parameters of GaInAsSb /GaAlAsSb avalanche photodiodes with separate absorption and multiplication regions (SAM APD). In [93] Susa *et al.* suggested for the first time a structure with

separate absorption and multiplication regions. The construction of the SAM APD was first implemented in a InGaAs/InP APD for the spectral range of 1-1.6 μm [93,94]. The location of the multiplication region with the p-n junction in a wide-gap material makes it possible to reduce the electric field at the hetero-boundary with the narrow gap region, and to avoid significant tunneling currents in this region when avalanche breakdown is achieved in the multiplication region. However, to achieve this, the carrier concentration and the layer thickness must meet strict requirements [94].

The SAM APD concept has been used to develop an APD on the basis of GaInAsSb/GaAlAsSb solid solutions [95]. A narrow-gap absorption region was made from an n-type (Te doped) $In_{0.22}Ga_{0.78}As_{0.18}Sb_{0.72}$ layer (Eg=0.52 eV), while the wide-gap n-p junction lies in the $Ga_{0.66}Al_{0.34}As_{0.025}Sb_{0.975}$ alloy layer of indirect-gap composition (Eg=1.2 eV) separated from the heteroboundary by 0.3 μm. The structures were fabricated by liquid-phase epitaxy on a GaSb(111) substrate (Te doped up to $n=2\times10^{17}cm^{-3}$). The carrier concentration in the narrow-gap region was $n=(5\text{-}7)\times10^{15}cm^{-3}$, and in the 0.3 μm thick wide-gap region was $n=7\times10^{16}cm^{-3}$; while the p^+ layer of wide–gap GaAlAsSb was doped with Ge up to $5\times10^{18}cm^{-3}$. A study of the current–voltage characteristics of the diodes over the temperature range 296-78 K (Figure19a) showed that the temperature coefficient of the breakdown voltage was positive: $\gamma=(\Delta U_B/U_B)(1/\Delta T)>0$, $\gamma=2\times10^{-3}$ K^{-1}. That is, the breakdown has an avalanche nature. The breakdown voltage was determined by the wide-gap material and was U_B=21-24 V at T=296K and 18-20 V at 78 K, respectively. The dark current density at $U=0.5U_B$ was $j=(1\text{-}2)\times10^{-2}$ A/cm^2, which is lower by a factor of two or three than that in ordinary APDs.

The dark current in the voltage interval 5-12 V was determined by generation and recombination in the narrow-gap material. The activation energy was E_a=0.34 eV according to the temperature dependence of the reverse current. The electric field in the space-charge region of the wide-gap N-P junction was $E=3.5\times10^5$ V/cm, falling to 8×10^4 V/cm at the boundary with the narrow-gap layer. The field penetrated into the narrow-gap layer when the voltage across the structure was about 4-5 V, as is quite evident from the step in the reverse branch of the current-voltage characteristic under illumination (see Figure 19b). A significant multiplication of the photocurrent was observed at U> 6 V. According to estimates, the field required for the beginning of significant tunneling in the wide-gap material was $E\sim10^6$ V/cm, while that in the narrow-gap GaInAsSb region was above 2.5×10^5 V/cm.

Figure 19b shows the dependencies of the multiplication coefficients as a function of the reverse bias voltage for the SAM-APD illuminated through the wide–gap p-GaAlAsSb layer with monochromatic light of various wavelengths. Maximum values of the multiplication factor M=10-30 (T=296K) and M=100-400 (T=78K) at an initiating photocurrent $I_{in}=5\times10^{-9}$A and at a wavelength of λ=2.0 μm were achieved in a SAM APD of this type. It can be seen from Figure 19b that the multiplication coefficient increases with increasing wavelength of the light. The space charge region lies in the n-type narrow-gap GaInAsSb absorption layer, with maximum electric field near the p-n GaAlAsSb junction and results in predominant multiplication of holes.



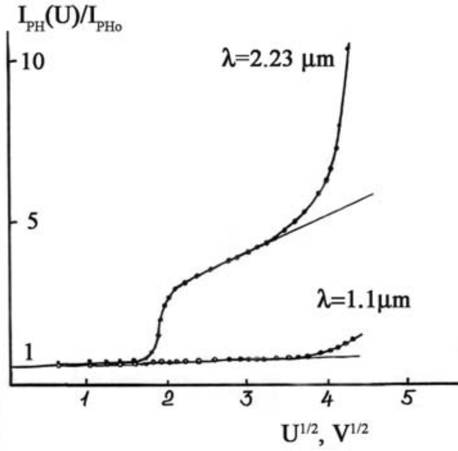


Figure 19. a) Current voltage characteristic of SAM APD. GaInAsSb/GaAlAsSb at various temperatures: 1-296; 2-253; 3-213; 4-173. *Inserts;* a) Schematic structure of SAM APD, b) Dark current vs 1/T at V=-5V. **b**) Normalized photocurrent *vs* reverse bias under illumination λ=2.23 μm (hole injection) and λ=1.1 μm (electron injection).

Figure 20 shows the photosensitivity spectrum at various applied bias. Since the diode was illuminated through the wide–gap p-type layer, the photosensitivity at zero bias is determined primarily by the GaAlAsSb material (curve 1). When the bias was raised above 4 V, and the electric field penetrated into the narrow-gap

layer, the photosensitivity spectrum of the SAM APD looks like that of a heterostructure photodiode.

Further improvement of these APDs requires optimization of several parameters of the structure, including carrier concentration in the wide-gap layer and the layer thickness, which determine the electric field at the heteroboundary. Furthermore, 'resonant' or quasi-resonant composition of the wide-gap GaAlAsSb solid solution (with $E_g \sim \Delta_O$) could be used for the multiplication region, in order to reduce the excess-noise coefficient.

Recently the authors of [95] have suggested an avalanche photodiode structure with (InGaAsSb) separate absorption and (n-GaAlSb/p-AlGaAsSb) multiplication regions. This SAM APD is sequentially composed of a 2.2 μm thick Te compensated $In_{0.22}Ga_{0.78}As_{0.18}Sb_{0.82}$ layer with an electron concentration of n=(5-7)$\times 10^{15}$ cm^{-3}; a 0.3 μm thick n-$Ga_{0.96}Al_{0.04}Sb$ 'resonant' composition layer with n=8$\times 10^{16}$ cm^{-3}; and a 1.5 μm thick $Al_{0.34}Ga_{0.66}As_{0.014}Sb_{0.986}$ window layer with p=5$\times 10^{18}$ cm^{-3}. The location of the p-n junction coincides with the hetero-boundary between two wide gap materials. Figure 21 shows the multiplication factor as function of the reverse bias voltage for the SAM-APD illuminated using light of 2.1 μm wavelength. The maximum values of the multiplication factor were measured to be M=30-40 (T=296K). The space – charge region lies in the n-type GaInAsSb absorbing layer with a maximum electric field near the n-$Ga_{0.96}Al_{0.04}$Sb/p-$Al_{0.34}$GaAsSb heterointerface and results in predominant multiplication of holes in the n-$Ga_{0.96}Al_{0.04}$Sb multiplication region. The breakdown voltage was determined by the wide-gap material and was 10-12 V at T=296K.

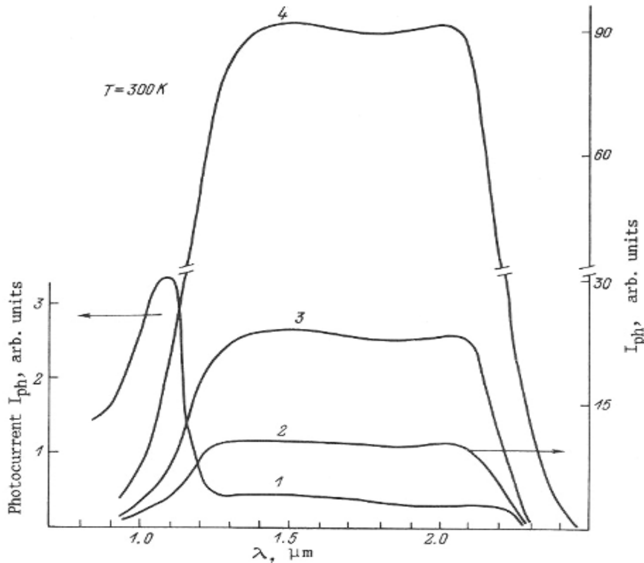


Figure 20. Photocurrent spectrum of GaInAsSb/GaAlAsSb SAM APD at reverse bias U,V: 1-0; 2-5; 3-10; 4-20. T=296 K

As was shown in [95], the multiplication layer in Al(Ga)AsSb might also be beneficial for reduction of excess noise arising due to the multiplication. As a band resonance condition takes place in $Al_{0.04}Ga_{0.96}Sb$ at 0.76 eV, very high values of β/α up to 60 are achieved for the ratio of the hole and electron ionization coefficients. Thus, an essentially unipolar multiplication by holes is provided that reduces the excess noise problem in these APDs. The ratio of the ionization coefficients in APDs based on GaInAsSb and GaAl(As)Sb alloys were also determined from the noise measurements. The results of these measurements will be presented and discussed in Section 4.4.

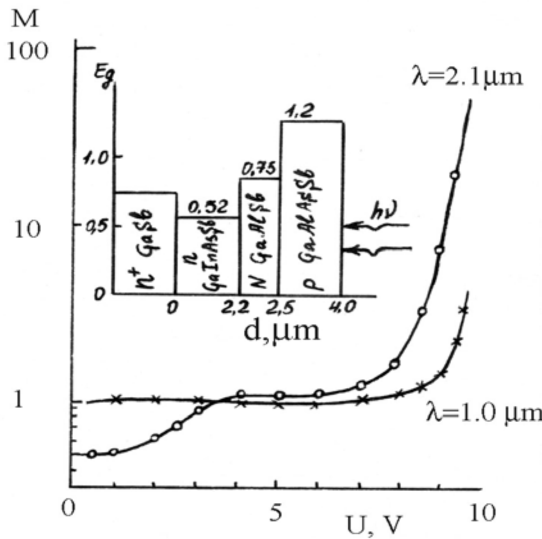


Figure 21. Multiplication coefficient M *vs* reverse bias for the SAM APD GaInAsSb/ $Ga_{0.96}Al_{0.04}Sb$ with "resonant" composition in the avalanche region

4.4 Noise and Response Speed of GaInAsSb/GaAlAsSb APDs

In this section the results of noise measurements in APDs based on GaInAsSb and GaAlAsSb alloys and their heterostructures will be discussed.

Fluctuations in the course of avalanche multiplication create a noise current whose spectral power density is equal to the shot noise of the photodiode current, multiplied by the mean-square value of the multiplication coefficient [5]:

$$S_i = 2q(I_d + I_{ph}) \langle M^2 \rangle \tag{31}$$

Where $\langle M^2 \rangle = \langle M \rangle^2 F(M)$, where $\langle M \rangle$ is the average value of the multiplication coefficient and F(M) is the factor representing the excess noise of the avalanche process (noise factor). It is shown in [21,22] that the noise factor is related to the multiplication coefficient and to the ratio of the hole and electron ionization

coefficients. For example, in the case of pure hole injection in the multiplication region, we have;

$$F_p(M) = M_p\left[1-(1-\frac{1}{k})\left(\frac{M_p-1}{M_p}\right)^2\right] \quad \text{(here } k=\beta/\alpha) \tag{32}$$

If an avalanche is initiated by electrons, we have $k=\alpha/\beta$. In the limiting case when $\beta>>\alpha$ and $\beta<<\alpha$ the excess noise factor is F(M)=1-2. In the other limiting case when the ratio of the ionization coefficients is k=1, we obtain F(M)=M. Therefore, the avalanche noise should be minimal in the case of very different α and β, and an avalanche is ignited by the carriers with the higher ionization coefficient.

The spectral density of the noise power due to avalanche multiplication in GaInAsSb/GaAlAsSb APDs [36,95] was determined using apparatus based on a noise factor meter KH5-30. Measurements were carried out using a noise generator calibrated at f=3 MHz in a frequency band of 0.3 MHz. The spectral density of the noise power was determined both during illumination of the APD samples with radiation of wavelength λ=2.0 μm, and also in the absence of illumination. It was assumed that the noise associated with dark current was not correlated with the noise due to photocurrent. Figure 22a shows the dependence of the spectral density of the noise power of the photocurrent (S_{iph}) on the value of the photocurrent (I_{ph}) for a n-GaSb/n-GaInAsSb/p-GaAlAsSb APD structure. It can be described by the following empirical law:

$$S_{iph} = 2e\, I_{pho}\, M^y \tag{33}$$

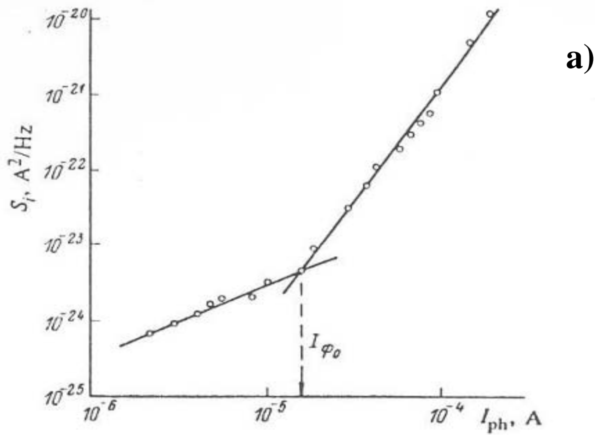
where I_{pho} is the photocurrent initiating an avalanche (deduced from a kink in the characteristic) and $M= I_{ph}(V)/I_{pho}$. The spectral density of the noise power can also be written in the form suggested in [21]:

$$S_{iph}=2e\, I_{pho}\, M^2 F(M) \tag{34}$$

We then have

$$F(M) = M^{y-2} = M_p\left[1-(1-\frac{1}{k})\left(\frac{M_p-1}{M_p}\right)^2\right] \tag{35}$$

Figure 22b shows the dependence of the noise factor on the multiplication coefficient of the APD under investigation. It also includes the theoretical curves of McIntyre [21]. Fitting parameters could be k_{eff} representing the effective ratio of the hole and electron ionization coefficients. It is clear from this figure that the best agreement is obtained for k_{eff} = 6-7. This result is in agreement with the data, which were obtained in an investigation of the field dependences of the ionization coefficients (see Section 4.1.2). The experimentally determined value of the noise factor F(M) was F=3.2 for a multiplication coefficient of M=10. The spectral density of the noise power due to avalanche multiplication of the three types of APDs based on GaInAsSb/GaAl(As)Sb were determined [95]. The results of the measurements are presented in Table 3, together with a description of the different APD structures.



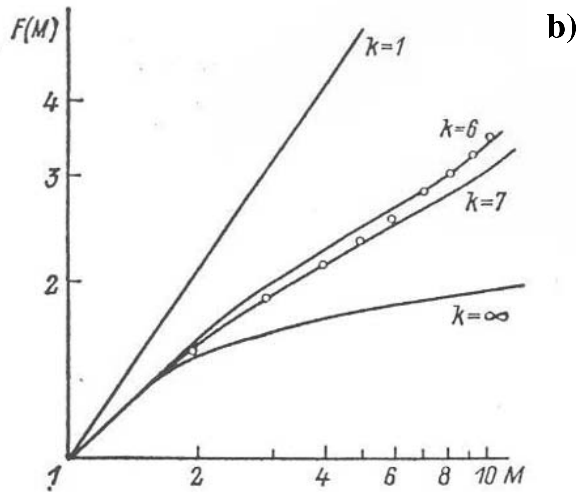


Figure 22. **a**) Spectral noise density (S_i) *vs* photocurrent value for GaSb/ GaInAsSb /GaAlAsSb APD structure. **b)** Excess noise factor F vs multiplication coefficient (M) for GaSb/ GaInAsSb/ GaAlAsSb APD structure. *Solid curves* – after McIntyre [21] at various k_{eff}=β/α; *circles* – experimental data

Table 3. The excess noise factor and ratio of the hole and electron ionization coefficients for different SAM APDs

SAM APD type		F~M ^{y}, y	F(M), M=10	β/α
Absorption region	Multiplication region			
1. $In_{0.22}Ga_{0.78}As_{0.18}Sb_{0.82}$	n-$In_{0.22}Ga_{0.78}As_{0.18}Sb_{0.82}$ p-$Al_{0.34}Ga_{0.66}As_{0.014}Sb_{0.986}$	0.5	3.2	7
2. $In_{0.22}Ga_{0.78}As_{0.18}Sb_{0.82}$	n-p $Al_{0.34}Ga_{0.66}As_{0.014}Sb_{0.986}$	0.7	5.0	3
3. $In_{0.22}Ga_{0.78}As_{0.18}Sb_{0.82}$	n-$Ga_{0.96}Al_{0.04}Sb$/ p- $Al_{0.34}Ga_{0.66}As_{0.014}Sb_{0.986}$	0.2	1.6	60

Low-noise avalanche photodiodes for the spectral range of 1.6-2.4 μm were first created using the GaInAsSb solid solution in the absorption region and the wide-gap $GaAl_x(As)Sb$ alloy of resonant composition (Eg=0.8 eV at T=300K for x=0.04) in the multiplication region [95]. This APD has a very high ratio of ionization coefficients, β/α>30 and low excess noise factor, F~1.6 (M=10), which is close to the theoretical limit (See Table 3).

Table 4 presents some values of the excess noise factor F(M) as a function of the ionization coefficient ratio for a number of avalanche photodiodes on the basis of III-V semiconductors and their alloys As seen from the Table, APDs based on GaSb, InAs and their alloys with practically monopolar hole multiplication have the lowest excess noise factor as compared with other known APDs with ionization coefficient ratio β/α ~1.

Table 4. Excess noise factor as a function of the ionization coefficient ratio for different avalanche photodiodes

Material	β/α	F(M), at M=10	Ref.
$Ga_{0.47}In_{0.53}As$/InP (100)	2-3	3.3	[4,95]
InGaAsP/InP (100)	0.3	3.5	[4, 114]
GaAlAs/GaAs (100)	1	5.9	[115]
$Ga_{0.72}Al_{0.28}Sb$/GaSb (100)	2	5	[115, 116]
$Ga_{0.05}Al_{0.95}Sb$/GaSb (111)	16	1. 2	[4,116]
GaSb (111)	100	1.9 (77K)	[4, 116]
InAs	20-50	1.1 (77K)	[69]
Ge	2	7	[116]
Si (111)	50	5 (M=100)	

A further step in the enhancement of the system capacity for high bit rate, long-wavelength communication systems can be expected from the use of very low loss (10^{-2}- 10^{-3} dB/km) mid-IR (2-4 μm) fibers based on fluoride or chalcogenide glasses

[24,25]. For such applications high speed and sensitive photodiodes (especially avalanche photodiodes APD) are required.

A suitable APD construction for the spectral range 1.5-2.5 μm is a SAM APD with an $In_{0.22}Ga_{0.78}As_{0.18}Sb_{0.82}$ absorption region and a $Ga_{1-X}Al_X(As)Sb$ (x=0.04-0.05) resonant composition multiplication region. As shown in Section 4.1, the $Ga_{0.96}Al_{0.04}Sb$ alloy has an hole/electron ionization coefficient ratio $\beta/\alpha>20$.

The sensitivity for a direct detection receiver using an SAM APD may be calculated according to the treatment of Personick *et al.* [100] at bit rate B=500 Mbit/s, using the NRZ coding format and BER=10^{-9}. (In these calculations we assume that the APD is followed by a transimpedance FET preamplifier with trans-conductance, g_m=40 ms; total input capacitance, C_T =1.5 pF and using a feedback resistance - bit rate product R×B =750 kOhm×Mbit/s). The dependence of minimum detectable power $\eta\langle P_{APD}\rangle$ on M for the SAM APD based on GaInAsSb/GaAl(As)Sb for the wavelength λ=2.1 μm and for a standard Ge APD operating at λ=1.55 μm were calculated and are presented in Figure 23. The SAM APD parameters for a device of 200 μm diameter and 1.1 A/W sensitivity at 2.1 μm wavelength, and the corresponding Ge APD parameters are listed in Table 5. The calculations were performed for a hole and electron ionization coefficient ratio k=60 for the SAM APD. The expression of McIntyre [21] for the excess noise factor is used (see also Section 4.4). When $M=M_{opt}$ (here M_{opt} is the optimum value of the avalanche gain), $\eta\langle P_{APD}\rangle$ is minimized, thereby yielding the highest attainable receiver sensitivity for a given $k=\beta/\alpha$ and amplifier noise current

.

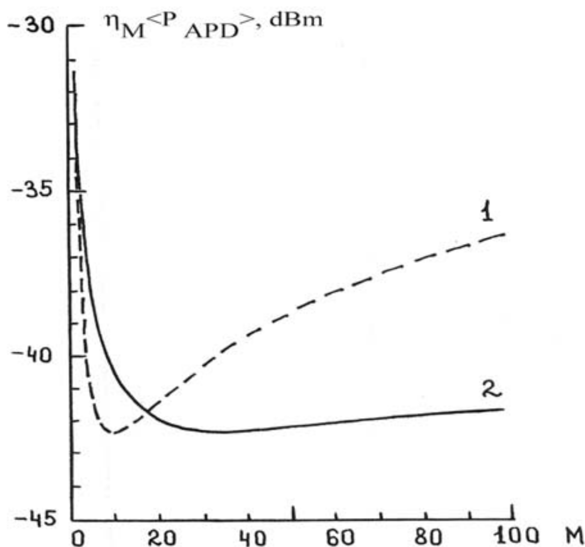


Figure 23. The dependence of the minimum detectable power $\eta\langle P_{APD}\rangle$ on M for SAM APDs based on GaInAsSb/ GaAl(As)Sb (λ=2.1 μm)-*curve 2*, and standard Ge APD (λ=1.55 μm)-*curve 1*

Table 5. A comparison of Ge APD and III-V based SAM APD parameters

Parameter	APD type	
	Ge APD	GaInAsSb/GaAl(As)Sb SAM APD
Wavelength λ, μm	1.55	2.1
Capacitance C, pF	1	7
Dark current I_{DM}, A (multiplied)	10^{-7}	10^{-6}
Dark current I_D, A (non-multiplied)	10^{-6}	10^{-5}
Hole and electron ionization coefficient ratio $k=\beta/\alpha$	~1-2	~60

It can be pointed out that in spite of higher dark current values, the receiver with the GaInAsSb/GaAl(As)Sb SAM APD has a minimum detectable power level $\eta\langle P_{APD}\rangle$ = –42.3 dBm at M_{opt}=34-39. As shown in Figure 23 sensitivity $\eta\langle P_{APD}\rangle$ is close to the minimum detectable power of the receiver with the Ge APD ($\eta\langle P_{APD}\rangle$ = –41.8 dBm and M_{opt}=10), which operates at the wavelength λ=1.55μm. This result is a consequence of the fact that the SAM APD based on GaInAsSb/GaAl(As)Sb has a higher value of the ionization coefficient ratio β/α=60 and a much lower value of spectral noise density.

Similar results were obtained by Benoit J. *et al.* [100], who made a theoretical evaluation of the sensitivity of the optical receiver with a SAM GaAlAsSb/GaInAsSb/GaSb APD for the 2.5 μm wavelength. Receiver sensitivity as low as –50.5 dBm at a bit rate of 0.66 Gbit/s and –46.5 dBm at a bit-rate of 2 Gbit/s can be obtained by thermoelectric cooling these devices down to 190K. At minimum detectable power optimum gain values higher than 100 can be reached.

These results demonstrate the potential of an optical receiver with a GaInAsSb/GaAl(As)Sb SAM APD for use in a mid-IR wavelength optical communication system. Fast, low-noise APDs for the spectral range of 1.6-2.5 μm are also of great interest because of their potential applications in laser range-finding, particularly with the development of compact solid-state lasers based on ILF:Ho (λ=2.06 μm) and YAG-Cr,Tm:Ho (λ=2.09 μm). The applicability of SAM APDs based on GaInAsSb/GaAl(As)Sb in a laser range-finding system for detecting pulsed laser signals at λ=2.06-2.09 μm was demonstrated in [26].

5 Avalanche Photodiodes for the 3-5 μm Spectral Range

In this section the performance of APDs based on InAs, InGaAs, and InAsSb for the spectral range 3-5 μm are considered in respect of ionization coefficient, field dependence, noise and response speed characteristics.

Avalanche gain in InAs photodiodes at 300 K was first reported by Lukovsky and Emmons [101], and gains of ~50 were measured at a frequency of 125 MHz. However, they did not study processes of avalanche multiplication and ionization in these materials in detail. The first experimental data on electron and hole ionization

coefficients in InAs photodiodes were published in [30, 67]. Below we consider the field ionization coefficient dependencies and the influence of band structure on the threshold energy of carrier ionization.

5.1 Experimental Investigation of Ionization Coefficients in InAs, InGaAs and InAsSb.

An investigation of the avalanche multiplication and impact ionization processes in p-n junctions formed in InAs, $In_{1-X}Ga_XAs$ and $InAs_{1-X}Sb_X$ were carried out by the authors in [30,67,35]. P-n structures were prepared by liquid phase epitaxy on InAs substrates with carrier concentration $n=(4\text{-}8)\times10^{16}$ cm^{-3} and were doped with Zn during the epitaxial growth process or by diffusion of Cd into an n-type film. The samples were constructed in such a way that p-n junctions could be illuminated by radiation either from the p-side or the n-side and to investigate separately the multiplication caused by electrons or holes in the same sample. InAs p-n junctions with breakdown voltages 15-30 V, InGaAs p-n junctions with breakdown voltages in the range 1-4 V, and InAsSb p-n junctions with breakdown voltage near 8 V were investigated at 77 K.

The electric field dependence of α and β in InAs and InGaAs at T=77 K is illustrated in Figure 24a. It is clear from the figure that the values of β were an order of magnitude higher than the values of α. The values of β were within the range 4×10^3–2×10^5 cm^{-1} and those of α in the range 4×10^2–3×10^4 cm^{-1}. The experimental dependences of the ionization coefficients on the electric field were exponential and could be described by

$$\alpha,\beta = \alpha,\beta_0 \exp\left[-(E_{O\alpha,\beta} /E)\right] \tag{36}$$

in a wide range of electric fields ($E=(2\text{-}8)\times10^4$ V/cm). Here, α,β_0 are ionization coefficients in an infinitely high electric field and $E_{O\alpha,\beta}$ are characteristic fields depending on the threshold ionization energy ε_i and on the mean free path of carriers λ. These dependences were used to estimate the values of the impact ionization parameters α,β_0 and $E_{O\alpha,\beta}$, listed in Table 6. In the case of holes, two linear regions on the curves of the hole ionization coefficient for both the InAs and the InGaAs were observed. The first region (in weak fields) corresponded to the ionization by holes from the spin-orbit-split band and the second one (in high fields) was due to holes from the heavy-hole band [30]. In this work the first important evidence was obtained that in impact ionization of InAs holes from the spin-orbit split band take part with the least possible threshold energy $\varepsilon_i=E_g=\Delta_O$.

Threshold ionization energy is related directly to the energy band structure of a semiconductor and depends on the effective mass of the carrier participating in the impact ionization process as well as on the band gap (see Section 2.2). In the case of InAs the threshold energy of conduction-band electrons is $\varepsilon_{ie}=1.05E_g=0.43$ eV, whereas, for the holes in the heavy band $\varepsilon_{ih} = 2E_g= 0.82$ eV and for the holes in the split-off band subject to the condition $\Delta_O\sim E_g$, $\varepsilon_{ih\ s\text{-}o}=0.43$ eV. It has been found that details of the structure of the InAs valence band, namely the similarity of the band

gap E_g and of the spin-orbit splitting Δ_O, results in an active contribution of holes in the spin-orbit band to the processes of impact ionization in weak fields. The spin-orbit band is filled as a result of the scattering of carriers from the heavy hole band. The ratio of the hole densities in the two bands (for the same energy) is governed by the ratio of the densities of states. The majority of the holes are concentrated in the heavy hole band. The ratio of the hole densities in these bands at a level near 0.43 eV amounts to 0.2. However, since the impact ionization threshold of the heavy holes is high, they cannot be responsible for the observed impact ionization processes. Impact ionization from the light-hole band may compete with the impact ionization from the spin-orbit split valence band. However, the threshold energy for the light holes is high. Therefore, only holes with minimum impact ionization threshold (holes from the spin-orbit split valence band $\varepsilon_{ih,s\text{-}o}$ =0.43 eV) participate in the impact ionization process. Experimentally it has been established that the ionization coefficient of the holes β is over an order of magnitude greater than electron ionization coefficient α. However, in the case of InSb p-n junctions, for which Δ_O is considerably greater than Eg (Eg=0,24 eV, Δ_O=0.80 eV at T=77K) the holes from the split-off band cannot participate in the ionization near the threshold. It was experimentally found the electron ionization coefficient is considerably greater than the hole ionization coefficient α/β>>1 [62].

Table 6. Estimated values of impact ionization parameters

Parameters \ APD material	InAs	$InAs_{0.88}Sb_{0.12}$
E_g, eV	0.41	0.30
Δ_O, eV	0.43	0.36
Electrons		
e_{ie} ,eV	0.43	
α_0, cm^{-1}	1.00×10^5	1.20×10^5
$E_{O\alpha}$, V/cm	1.57×10^5	1.00×10^5
s-o holes		
$e_{ih,s\text{-}o}$,eV	0.43	0.36
β_0, cm^{-1}	4.65×10^5	1.10×10^5
$E_{O\beta}$, V/cm	8.80×10^4	1.20×10^5
Heavy holes		
$e_{ih,h}$,eV	0.82	0.60
β_0, cm^{-1}	1.60×10^6	7.00×10^5
$E_{O\beta}$, V/cm	1.50×10^5	1.55×10^5

Avalanche photodiodes have also been fabricated in InSb [102]. Multiplication measurements in these devices indicates that α is much greater than β, and analysis of the noise measurements using McIntyre's theory agreed with this conclusion; that is, the noise increased only slightly faster than the signal. Avalanche gain values as

high as 50, breakdown voltages of 18-23 V and long-wavelength cut-offs of 5.4 μm at 77K were attained.

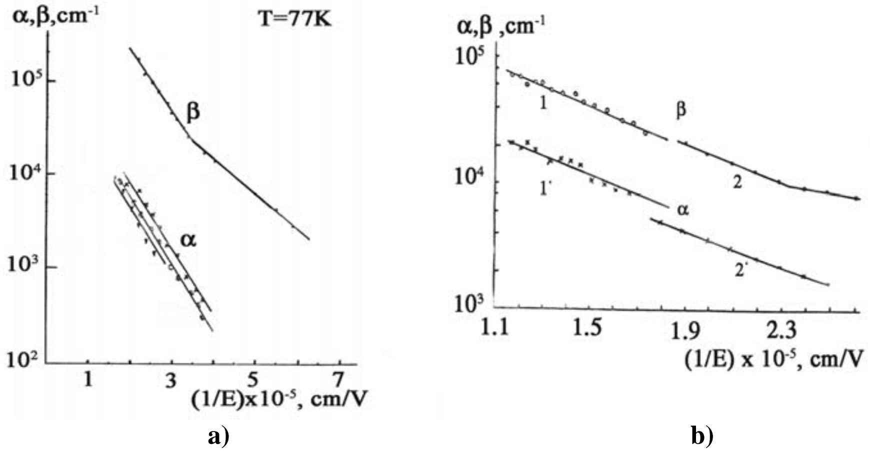


Figure 24. Dependence of the electron (α) and hole (β) ionization coefficients on the inverse electric field; **a)** for InAs, $In_{1-x}Ga_XAs$ (x=0.02) **b)** for $InAs_{1-x}Sb_x$ (x=0.12) 1,1 – epitaxial structure, 2,2 – diffusion p-n junction

In this respect it is of interest to investigate avalanche multiplication and impact ionization in $InAs_{1-X}Sb_X$ solid solutions because this should make it possible to study some intermediate cases between the two mentioned above and to find changes in the impact ionization parameters due to changes in the energy band structure with variations in the solid solution composition. An investigation was made of the avalanche multiplication and impact ionization processes in p-n junctions formed in $InAs_{0.88}Sb_{0.12}$ solid solutions. The experimental results of photocurrent multiplication were used to calculate the electron and hole ionization coefficients. The electric field dependence of α and β in InAsSb at T=77 K is illustrated in Figure 24b. In fields of $(3–5)\times10^4$ V/cm the ionization coefficients were described by the expressions $\alpha,\beta = \alpha,\beta_0 \exp[-E_{O\alpha,\beta}/E]$, where $\alpha_0= 1.2\times10^5$ cm^{-1} and $E_{O\alpha}=1\times10^5$ V/cm apply to electrons, while $\beta_0=1.1\times10^5$ cm^{-1} and $E_{O\beta}= 1.2\times10^5$ V/cm apply to heavy holes from the spin-orbit–split valence band. It was found that β>α because of the considerable contribution of the spin-orbit–split band to ionization of holes. Estimates were obtained for the threshold ionization energies [35] of electrons $\varepsilon_{ie} =1.05E_g=0.32$ eV and for holes from the heavy band $\varepsilon_{ih} = 2E_g=$ 0.60 eV An analysis was made of the influence of composition of $InAs_{1-X}Sb_X$ on the main role played by the spin-orbit-split valence band in the impact ionization processes under conditions corresponding to $\Delta_O> E_g$. In the case of $InAs_{1-X}Sb_X$ with composition x=0.12, the threshold ionization energy value of holes from spin-orbit-split band ε_{ihso} was close to Δ_O~0.36 eV, and the band gap was E_g=0.30 eV. For $InAs_{1-X}Sb_X$ alloys with composition x=0.12 threshold energy for the hole impact ionization becomes larger than for the case InAs where $\varepsilon_{ihso} \sim E_g = \Delta_O$, and one can observe resonant curves as shown in Section 3.4.

5.2 Noise and Response Speed of Long Wavelength APDs

InAs and InGaAs alloys are of great interest as materials for high-speed infrared photodiodes including APDs working in the spectral range up to 4 μm. Photodiodes operating in this spectral range can be produced from InAs in combination with multi-component solid solutions such as InGaAs and InAsSb.

In Section 2 it was shown that in constructing low-noise, fast response avalanche photodiodes it is necessary to utilize semiconductor materials with large asymmetry of charge carrier ionization coefficients [5]. InAs is a semiconductor with strong asymmetry of electron and hole ionization coefficients. As was demonstrated in Section 5.1 the ionization coefficient of the holes β is over an order of magnitude greater than the electron ionization coefficient α. Multiplication noise is a key parameter determining the performance of avalanche photodiodes and Lukovsky and Emmons [101] found that in InAs the noise increased faster than the signal for multiplication values higher than unity. For InSb, p-n junction noise measurements interpreted using McIntyre's theory agreed with the conclusion that α is much greater than β. That is, the noise increased only slightly faster than the signal. Below we present some results of multiplication noise measurements in InAs and InGaAs p-n junction APDs. The design structure of these APD was described previously in Section 6.1.

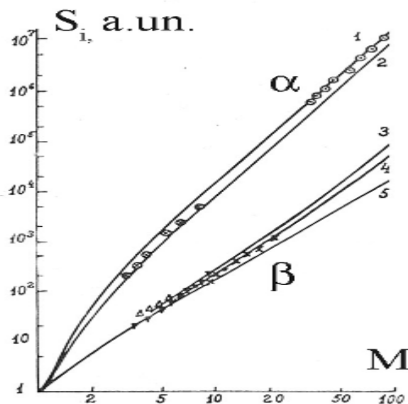


Figure 25. Dependence of spectral noise density on multiplication coefficient for APD structures and for various $k_{eff}=\beta/\alpha$ in the case of electron injection (α) and hole injection (β). *Solid curves* – theoretical calculation after McIntyre [21]. k_{eff}: 1,3-10; 2,4-20; 5-∞, *points* – experimental data

The multiplication noise was studied at f=100 kHz in a 1 kHz bandwidth. The APDs were illuminated by monochromatic light, which was strongly absorbed in the p- and n- side of the junction creating pure electron or hole injection into the multiplication region. The results of noise measurements in InAs and InGaAs APD

structures are shown in Figure 25. When the avalanche process was initiated by electrons, the spectral noise density was proportional $\langle i^2 \rangle \sim M^3$, in the case of holes initiating the avalanche the spectral noise density was proportional $\langle i^2 \rangle \sim M^2$.

Experimental results of the spectral noise densities compared with McIntyre's theoretical curves are presented in Figure 25. The calculated curves are shown for various values of the ionization coefficient ratios (solid lines) and experimental data (dots and triangles). It is clear from this Figure that the best agreement is obtained for $k_{eff} > 20$. This result is in agreement with the data obtained by an investigation of the field dependences of the ionization coefficients (see Section 5.1, β/α=20-50). The experimentally determined value of the noise factor was $F(M)=M^{0.1}$, so the excess noise factor was determined to be F~1.0 .

Pulsed measurements of InAs APDs under GaAs laser diode (λ=0.85 μm) illumination at frequency f= 50 MHz and 10 MHz bandwidth have shown that the output signal in avalanche mode grows with increasing multiplication coefficient linearity.

One of the main merits of InAs-based APDs is the possibility to make an APD with a low reverse bias operating voltage and to reach high values of signal-noise ratio by means of reduced excess noise factor under near monopolar multiplication. The possibility to achieve these parameters only at liquid nitrogen temperatures and with a relatively high level of reverse dark current have prevented the use of InAs-based APDs in optoelectronics. We suppose that new methods of narrow gap material growth (MOVPE and MBE [103, 104]) could make it possible to reduce dark currents and to increase working temperatures up towards room temperature.

6 Methods of Separating Ionization Coefficients Using Quantum Structures

As discussed above in Section 3.5, there are only a few III-V semiconductor compounds based on GaSb, InAs and their alloys (GaInAsSb, GaAlAsSb, InAsSb) where due to "lucky" peculiarities of the band structure (band gap "resonance" $E_g \approx \Delta$, and the location of the subsidiary L-valley in GaSb) one can obtain large ionization coefficient ratios and design high speed low-noise APDs for the spectral range 2-4 μm.

Some physical approaches have been proposed to separate the electron and hole ionization coefficients by using peculiarities of energy diagrams of the heterostructures (band offsets), to realize "impact-engineering" in APDs based on AlGaAs/GaAs and AlGaSb/GaSb [5] superlattice and multiquantum well structures. Staggered avalanche photodiodes based on multi-layer structures with quantum well were proposed in [4,5]. A theoretical consideration of electron heating by strong electric fields in semiconductors with a superlattice was given by Suris *et al.* [105]. It is especially important for the fabrication of APDs in III-V semiconductor structures, where in bulk materials the electron and hole coefficients are similar, $\alpha \approx \beta$ (GaAs, InP, GaAlAs, AlInAs *etc.*). APDs based on these materials are of great interest for lightwave communications in the spectral range 1.3-1.55 μm.

Capasso *et al.* [106] proposed a staircase APD based on a multilayer structure using quantum wells. It was shown experimentally [107] that effective coefficients for electrons and holes in AlGaAs/GaAs multilayer structures can differ strongly (α/β=8, see Figure 22 a,b), although their values in bulk materials are similar. This is connected with the difference in band-offsets at the heterointerface in the conduction and valence bands and which can be obtained in heterostructures such as GaAsSb-GaInAs, In(GaAl)As-GaAsSb and GaInAs-GaAlAsSb lattice-matched to InP.

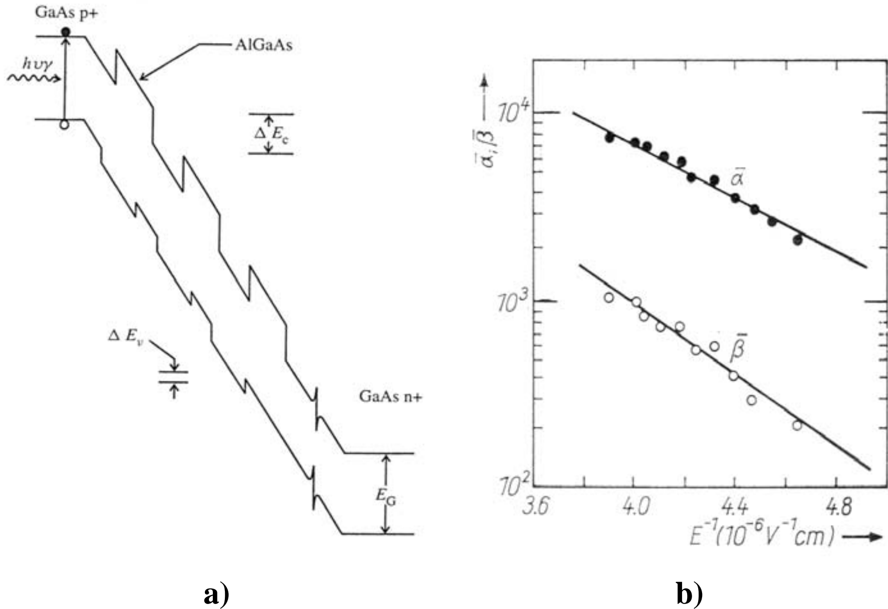


Figure 26. a) Sketch of the energy diagram of a multiquantum well APD with high potential steps after Brennen *et al.* [108]; **b)** Electron and hole ionization coefficient in GaAs-GaAlAs staircase multilayer structure after Capasso *et al.* [106].

The physics and performance of APDs based on superlattices and quantum well structures were recently reviewed in [108]. The theory of impact ionization across a potential step was considered and experimentally studied. Enhancement of the electron impact ionization coefficient across the potential step in the GaInAs/AlInAs and InAlGaAs/InAlAs systems was described. A new design for a high speed APD, the so-called channel avalanche photodiode, was first proposed in [5]. This device can be made by MBE from the channels of n-GaSb and p-AlAsSb. A very high ratio of electron and hole ionization coefficients α/β is predicted for such a device. It is interesting to note also a proposition to increase the β/α ratio in staggered heterostructures where "clusters" with asymmetric band offset are formed at the interface (*e.g.* in AlGaAs, InGaAsP, InGaAs and GaAsSb) [109].

Well-developed MBE technology now allows one to fabricate artificial clusters and grow self-organized quantum dots, *e.g.* InAs/GaAs. In Type II quantum dots one charge carrier is localized inside the dot and the other one in the barrier [110,111]. Optical properties of GaSb/GaAs, InSb/GaAs and AlSb/GaAs quantum dot systems have been studied [112]. But, up to now only low-threshold lasers have been made using quantum dot structures. Maybe there is an opportunity to use these quantum dot structures for creation of APDs. However, this problem requires a special study of impact ionization processes in quantum dot heterostructures.

7 Conclusion

The main objective of this chapter was to draw attention to high-speed low-noise avalanche photodiodes based on the narrow-gap III-V semiconductors and their alloys. Such APDs are very promising for a number of applications in the 2-5 μm spectral range, including; laser-diode spectroscopy, mid-infrared fiber optics, laser range-finding, free-space optical links for high frequency communications *etc.*

Impact ionization processes in III-V semiconductors and their connection with band structure have been considered. Experimental and theoretical studies of ionization coefficients of electrons and holes in several important III-V semiconductors and their dependence on applied electric field have been examined. The findings show that these materials have the advantage of a large electron/hole ionization coefficient ratio. Furthermore, the rather interesting phenomena of anisotropy of electron ionization coefficient in multi-valley semiconductors (GaAs, InP, GaSb) and mono-polarity of multiplication in GaSb have been considered and reviewed. Results of experimental and theoretical studies of impact ionization processes in these materials and their solid solutions (GaInAsSb, GaAlAsSb, InAsSb, InGaAs, GaInSb) have also been presented.

A new physical approach has been proposed for designing avalanche photodiodes with high ionization coefficient ratio using resonant hole ionization in materials where the band gap energy is equal to the value of the spin-orbit splitting, $E_g \sim \Delta_0$ (*i.e.* InAs, GaSb and their alloys). The performance of mid-infrared fast, low-noise avalanche photodiodes for the spectral range 2.0-2.5 μm, including APDs with separate absorption and multiplication regions (SAM APD) based on GaInAsSb/GaAl(As)Sb as well as narrow-gap APDs (λ=3.0-5.0 μm) based on InAs and their alloys with a "resonant" composition have been reported.

Finally, some methods of separating ionization coefficient of electrons and holes in materials where they are equal (GaAs, GaAlAs *etc.*) by using quantum wells and superlattices have been briefly described. The authors hope that this review will be useful for scientists and engineers studying both the fundamentals of impact ionization, as well as in the design and creation of avalanche photodiodes and other devices with internal gain for high frequency applications.

References

1. Stillman GE and Wolfe CM In: Semiconductors and Semimetals **12**; ed. by Willardson RK and Beer AC Academic Press, 1977
2. Law HD, Nakano K Tomasetta LR. J. Quant. Electr. 1979; **QE15**: 549
3. Pearsall T, Nahory RE and Pollack MA. Appl. Phys. Lett. 1976; **28**: 403
4. Korolkov VI, Mikhailova MP, Fiz.Tekh.Poluprov. 1983; **17**: 569
5. Capasso F In Semicond. and Semimetals **22D;** Photodetectors 1985, ed. by Tsang W.T.
6. Dmitriev AP, Mikhailova MP and Yassievich IN. Phys. Stat. 1987; **140**: 9.
7. Kogelnik H In: Springer Series in Optical Sciences, Lasers and Applications, 1981, ed. Guimaraes WON, Lin CT and Mooradian A:235
8. Stilman GE, Robbins VM, Tabatabaie N. IEEE Trans Electr. Dev. 1984; **ED-31**: 1643.
9. Washington MA, Nahory RE, Pollack MA, Beeb ED, Appl. Phys Lett. 1978; **33**
10. Olsen GH, IEEE Electr Dev Lett 1981; **EDL-2**: 217
11. Chin R, Hill CM, Appl Phys Lett, 1982, **40**, 332
12. Dianov EM, Sov Quant Electr. 1980: **7**: 453
13. Alferov ZhI, GorelenokAT, Kolyshkin VI,Kop'ev PS, Tarasov IS, Mdivaki VI *et al.* Zhurnal Tech Phys. Pisma 1978: **4**: 1329
14. Alferov ZhI, Andreev VM, Korol'kov VI, Portnoi EL, Yakovenko AA. Fiz. Tekh Poluprov 1969; **3**: 930
15. Tsarenkov BV, Akperov YaG, Verschak GG, Evstropov VV, Imenkov AN and Yakovlev YuP. Fiz. Tekh Poluprov 1972; **6**: 921
16. Pearsall TP and Pollack MA In: Semicond and Semimetals 1985; **22**, Part D, Photodetectors, ed. by Willardson RK and Beer AC Academic Press.
17. Pearsall TP, Hopson RW J Electr Mater 1978; **7**: 133
18. Chin R, Hill CM, Appl.Phys.Lett. 1982; **40**: 332
19. Kanbe H, Susa N, Nakagone H, Ando H Electr Lett 1980; **16**: 163
20. Capasso F, Mohammed K, Alavi K, Cho AY, Foy PW Appl Phys Lett 1984;**45**: 968.
21. McIntyre RJ. IEEE Trans Electr Dev 1966; **13**:164
22. Tager AS. Sov Phys Solid-State 1965; **6**: 1919
23. Webb PP, McIntyre RJ and Conrade RCA Rev 1974; **35**:234
24. Miyashita T, Manabe T. IEEE J Quant Electr 1982; **QE-18**:1432
25. Dianov EM. Izvestja Vuzov. Radiotechnika 1983; **26**: 27
26. Andreev IA, Baranov AN, Voznitskii MV, Ermakov BA, Mikhailova MP *et al.*. Sov.J.Optical Technol 1991; **58**:339
27. Bechers RJ. Laser Focus World 1973; **71**
28. Barnes NP. SPIE 1986; vol. **663**: 2
29. Kincade K. Laser Focus World 2004; May: 154
30. Mikhailova MP, Smirnova NM, Slobodchikov SV Fiz Tekh Poluprov 1976; **10**: 860
31. Zhingarev MZ , Korol'kov VI, Mikhailova MP and Yassievich IN. Zh Tech Fiz. Pisma (Soviet Tech Lett) 1979; **5**: 862.
32. Zhingarev MZ , Korol'kov VI, Mikhailova MP, Sazonov VV. Zh Tekh Fiz Pisma 1981; **7**: 1487.
33. Hilderbrandt O, Kuebart W, Pilkuhn M. Appl Phys Lett 1980; **37**: 807.
34. Andreev IA, Mikhailova MP, Semenov AN, Slobodchikov SV, Stus' NM and Filaretova GM. Fiz Tekh Poluprov 1984; **18**: 545.
35. Matveev BA, Mikhailova MP, Slobodchikov SV, Smirnova NM, Stus' NM and Talalakin GM. Fiz Tekh Poluprov 1979; **13**: 498.
36. Andreev IA, Mikahilova MP, Mel'nikov SV, Smorchkova YuP, Yakovlev YuP. Fiz Tekh Semicond (Sov Phys Semicond) 1991; **25**: 1429.

37. Dmitriev AP. Fiz Tekh Poluprov 1982; **16**: 978
38. Anderson CL and Crowell CR. Phys Rev 1962: **B5**: 2267
39. Pearsall TP, Capasso F, Nahory RE, Pollack MA and Chelikovsky JR Solid-State Electr. 1978; **21**: 297
41. Shocley W. Czech J Phys 1961; **b11**: 81
42. Davydov BT. Zh Exper Theor Fiz 1936; **7**: 269
43. Wolf PA. Phys Rev 1959; **95**: 1435
44. Baraff GA. Phys Rev 1962; **128**: 2502
45. Gribnikov ZS. Zhurn experim teor fiz. 1978; **24**: 2112
46. Gribnikov ZS. Zhurn experim teor fiz. 1981; **15**: 1572
47. Armenito CA, Growes Sh and Hurwitz CE Appl Phys Lett. 1979; **35**: 333
48. Dmitriev AP, Mikhailova MP, Yassievich IN. phys stat sol(b), 1982; **113**: 125
49. Hess K and Schichijo H. Phys Rev, 1981; **1323**: 4197
50. Vul BM and Shotov AP Solid State Phuysics in Electronics and Telecommunication 1960; **1**: 491.
51. Moll JL and Overstraeten RV. Solid-State Electr. 1962; **6**: 44
52. Gribnikov ZS, Ivastchenko VM and Mitin VV. Phys Stat Sol (b) 1981; **105**: 451
53. Albrecht H and Lerach L Solid-State Electr. 1980; **23**: 357
54. Emmons, RB and Lucovsky L IEEE Trans Elect Dev. 1966; **ED-13**: 297
55. Kaneda T In: Semicond and Semimet. 1985; Academic Press, ed. by Willardson RK and Beer AC, vol.**22 Part D**, Photodetectors
56. Chelikovsky JR, Cohen ML. Phys Rev 1976; **B14**: 556
57. Denis V, Kazleris J, Matulenis A In: Electrons in Semiconductors, 1978, ed. by Pozhela Yu, vol. **1** Mosklas, Vilnius
58. Brennen K and Hess K Phys Rev. 1984; **B29**: 5584
59 Dumke WP. Phys Rev. 1968; **167**: 783
60. Kane EO, J Phys Chem Solids 1954; **1**: 249
61. Baertch RD. J Appl Phys. 1967; **38**: 4267
62. Gavrushko VV, Kosogov OV and Lebedev VD. Fiz Tekh Poluprov 1978; **12**: 2351
63. Gutkin AA and Kumenkov SE. Fiz Tekh Poluprov 1981; **15**: 48
64. Bauer G and Kuchar F Phys Stat Sol (a) 1972; **13**: 169
65. Dick Cl and Anker-Johnsin B, Phys Rev. 1972; **B5**: 526
66. Mikhailova Mp, Nasledov DN, Smirnov VS and Dlobodchikov SV In: Issledovania po poluprovodnikam, 1968, Karta Moldavenyaska, Kishinev: 13.
67. Mikhailova MP, Nasledov DN, Slobodchikov SV. Fiz Tekh Poluprov 1967; **1**: 123
68. Mikhailova MP, Slobodchikov SV, Smirnova NN, Filaretova GM Fiz Tekh Poluprov. 1976; **10**: 978
69. Mikhailova MP, Rogachev AA and Yassievich IN. Fiz Tekh Poluprov 1976; **10**: 1480
70. Takeshima M. J Appl Phys 1972; **43**: 4114.
71. Hang D, Kerkhoff D, Lochman W. Phys Stat Sol(b) 1991; **89**: 357.
72. Zotova NV and Yassievich IN. Fiz Tekh Poluprov 1977; **11**: 1802.
73. Gelmont BL, Sokolova ZN, Yassievich IN Fiz Tekh Poluprov 1982; **16**: 592.
74. Haiecker W. Solid St.El. 1974; **17**: 993
75. Zhingarev MZ, Korol'kov VI, Mikhailova MP, Sazonov VV, and Tretyakov DN Sov.Phys.-Semicond. (English Transl.) 1980; **14**: 801
76. Hildebrand O, Kuebart W, Benz KW, and Pilkuhn MH IEEE J. Quantum Electron. 1981; **QE-17**:284
77. Miura S, Mikawa T, Kuwatsuka H, Yasuoka N, Tanahashi T, and Wada O, Appl. Phys.Lett. 1989; **54**:2422
78. Luquet H, Perotin M, Gouskov L. *et al.*. J.Appl.Phys. 1990; **68**(8): 3861
79. Andreev IA, Il'inskaya ND, Kunitsyna EV, Mikhalova MP, and Yakovlev Yu P. Semiconductors 2003; **37**: 949

80. Andreev IA, Afrailov MA, Baranov AN, S. G. Konnikov,, Mirsagatov MA, Mikhailova MP, Salata, V. B. Umanskii, G. M. Filaretova,, and Yakovlev Yu P. Sov. Tech. Phys. Lett. 1989;**15**: 253
81. Crowell CR and Sze SM Appl.Phys.Lett. 1966; **9**: 242
82. Andreev IA, Afrailov MA, Baranov AN, Mirsagatov MA, Mikhailova MP, and Yakovlev Yu P. Sov. Tech. Phys. Lett. 1987; **13**: 199
83. Sulima OV, Mauk MG, Shellenbarger ZA *et al.*. IEE Proc. Optoelectron.2004; **151(1)**;1
84. Andreev IA, Baranov AN, Zhingarev MZ, Korol'kov VI, Mikhalova MP, and Yakovlev YuP. Sov. Phys. Semicond. 1985; **19**: 987
86. Law HD, Chin R, Nakano K, and Milano R. IEEE J.Quantum Electron. 1981: **QE-17**: 275
87. Andreev IA, Kunitsyna EV, Lantratov VM, L'vova TV, Mikhalova MP, and Yakovlev YuP. Semiconductors 1997; **31**: 556
88. Chin R and Hill CM Appl.Phys.Lett. 1982; **40**: 332
89. Esaki L. In Tunneling Phenomena in Solids (Proc. NATO Advanced Study Institute, Riso, Denmark,1967, ed. by E.Burstein and S.Lundqvist) Plenum Press, NY, 47, 1969.
90. Kurnosova OV,and Yassievich IN Sov.Phys. Solid State 1984; **26**: 1988
91. Shelenbarger ZA, Mauk MG, Sims PE *et al.*.Mat.Res.Soc.Symp.Proc. 1998; **484**: 135
92. Carter BL, Shaw E, Olesberg JT, Chan WK, Hasenberg TC, and Flatte ME Electron.Lett. 2000; **36**: 1301
93. Susa N, Nakagome H, Mikamo O, Ando H, and Kanbe H IEEE J.Quantum Electron. 1980; **QE-16**: 864
94. Ando H, Yamaguchi Y, Nakagome H, Susa N, and Kanbe H, IEEE J.Quantum Electron 1981; **QE-17**: 250
95. Andreev IA, Afrailov MA, Baranov AN, Mar'inskaya NN, Mirsagatov MA, Mikhailova MP, and Yakovlev, YuP. Sov. Tech. Phys. Lett. 1989; **15**(9): 692
96. Sun YM, Wang JM, and Wu MC. Jpn.J.Appl.Phys. 1996; **35**: 5246
97. Dries JC, Gokhale MR, and Forrest SR. Appl. Phys. Lett. 1999; **74**(18): 2581
98. Andreev IA, Afrailov MA, Baranov AN, Konnikov SG, Mirsagatov MA, Mikhailova MP. *et al.*. Sov. Tech. Phys. Lett. 1989; **15**: 253
99. Smith RC, Personick In Topics in Applied Physics ed by H.Kressel, Springer Verlag; **39**: 1980
100. Benoit J, Boulou M, Soulage G, Joullie A, Mani H. J.Opt.Commun. 1988; **9**(2): 55
101. Lukovsky G and Emmons RB. Proc.IEEE 1965; **53**: 180
102. Gutkin AA, Kosogov OV, Kumekov SE Sov.Phys.Semicond. 1980; **14** (6): 1161
103. Kizhaev SS, Mikhailova MP, Molchanov SS, Stoyanov ND and Yakovlev YuP. Techn.Phys.Lett. 1998 ; **24** : 247
104. Lin Ray-Ming, Tang Shiang-Feng, Lee Si-Chen, Kuan C.H. Journal of Crystal Growth 2001; **227-228**: 167
105. Suris RA, Shamkhalova BS. Fiz Tekh Poluprov 1984; **18**: 1178.
106. Capasso F, Pearsall TP, Thornber KK, Nahory RE, Pollack MA *et al.*. J.Appl.Phys. 1982; **53**: 3324
107. Chin R, Holonyak N Jr., Stillman GE, Tang JT, Hess K. Electr Lett 1980; **16**: 467
108. Brennen KF, Haralson J. Superlat and Microstruct 2000; **28**: 77
109. Bulman GE, Robbins VM, Brennen KF, Hess K, Stillman GE, IEEE Electr Dev Lett 1983; **EDL-4**: 181
110. Krestnikov IL, Sakharov AV, Ledentsov NN *et al.*. Proc 6th Int Symp. Nanostructures: Physics and Technology, St.Petersburg, Russia, 1998: 257.
111. Bimberg D, Grudmann M, Ledentsov NN. Quantum Dot Heterostructures, Willey, 1988

112. Ledentsov NN *et al.* Phys Rev 1995; **B52**:14058
113. Zhukov AE, Kovsh AR, Mikhrin *et al.*. Proc. 8th Conf. Nanostructures: Physics and Technology , St.Petersburg, Russia, 2000: 38.
114. Diadiuk V, Groves SH, Hurwitz CE. Appl. Phys. Lett. 1980; **37**:807.
115. Tomasetta LR, Law HD. Laser Focus. 1978; **14**(12):42.
116. Forrest SR, DiDomenico MJr,Smith RG, Stocker HJ. Apll.Phys.Lett. 1980; **36**(7):580.

Part IV

Applications

Infrared Methods for Gas Detection

J.G. Crowder
Physics Department, Heriot-Watt University
Edinburgh, EH14 4AS

S.D. Smith, A. Vass and J. Keddie
Edinburgh Instruments Ltd
Edinburgh, EH54 7DQ

1 Introduction

Many gases that we wish to detect or monitor in our environment are infrared active, exhibiting bands of absorption lines in the infrared part of the spectrum, typically from 2.5 to 15μm. These absorptions provide a means of detecting the presence of a gas and measuring its concentration. The narrow emission line of a coherent laser source can be tuned to a single absorption line of a gas, giving highly sensitivity gas detection and this represents the high-cost, low-volume end of the gas detection market. On the other hand, thermal sources and infrared Light Emitting Diodes (LEDs) are spectrally broad sources and are used in the low-cost, high-volume sector, where thermal sources are well established and infrared LEDs are showing promise. The major part of this chapter is devoted to the latter, high-volume end of the market.

2 Gas Absorption Spectra

Infrared active gas molecules absorb infrared radiation due to changes in the vibrational and rotational states of the molecule. The resulting infrared spectra typically lie between 2.5 and 15μm because the absorption is based on the strong, fundamental transition of a vibrational mode of the molecule and these modes have a frequency, and associated wavelength, in the infrared part of the spectrum [1]. Fine structure is added by simultaneous rotational transitions in the molecule, resulting in a band of absorption lines centred around the wavelength corresponding to the vibrational transition in the molecule. Each absorption line in the band typically has a pressure-broadened width of 0.1cm^{-1} at N.T.P[1]. and the lines are typically spaced by 1 to 2cm^{-1}. Such spectra are illustrated in Figures 1, 2 and 3. The bands are commonly about 2% wide and a filter of this width is indicated in the Figures.

[1] (Normal Temperature and Pressure, defined as 20ºC and a pressure of one atmosphere).

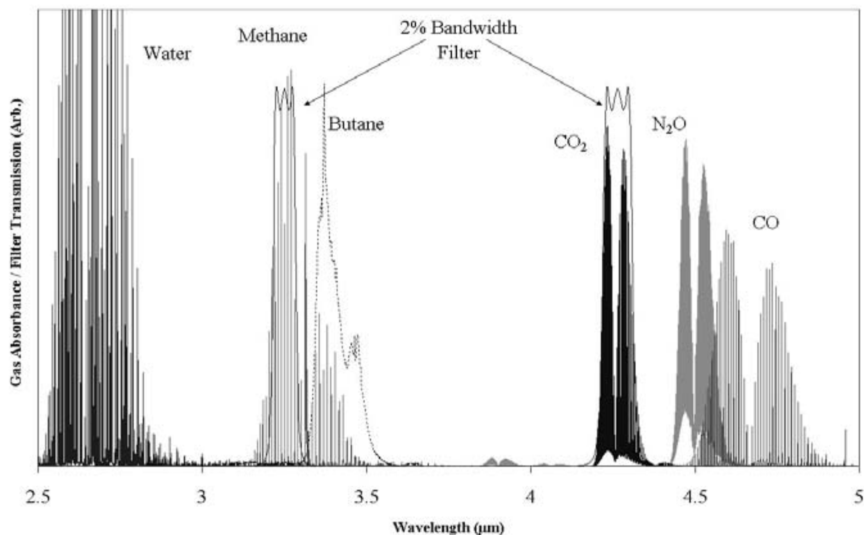


Figure 1. Absorption bands of common gases from 2.5 μm to the glass envelope cut-off at 4.7 μm. The figure includes data used with the kind permission of Infrared Analysis Inc [2], US EPA [3] and US NIST [4].

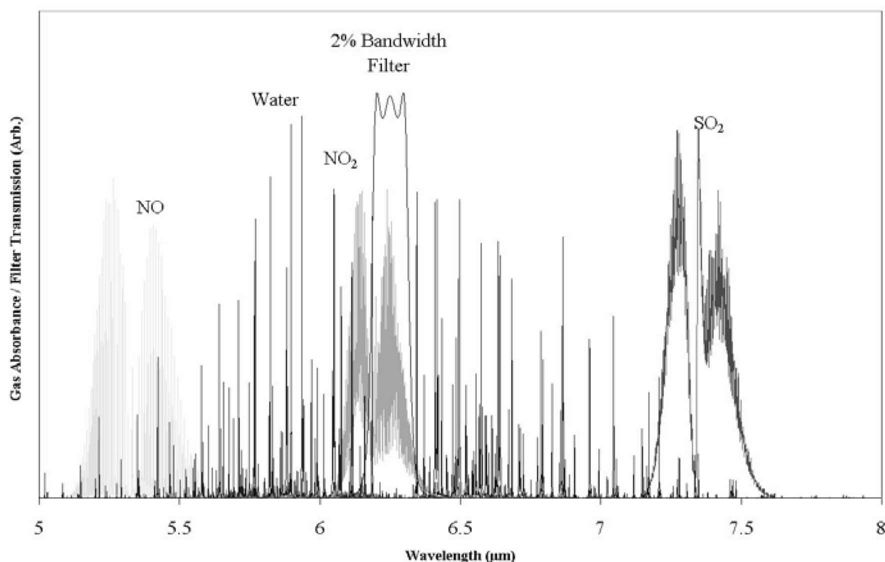


Figure 2. Absorption bands of common gases in the water band from 5 to 8 μm. The figure includes data used with the kind permission of Infrared Analysis Inc [2], US EPA [3] and US NIST [4].

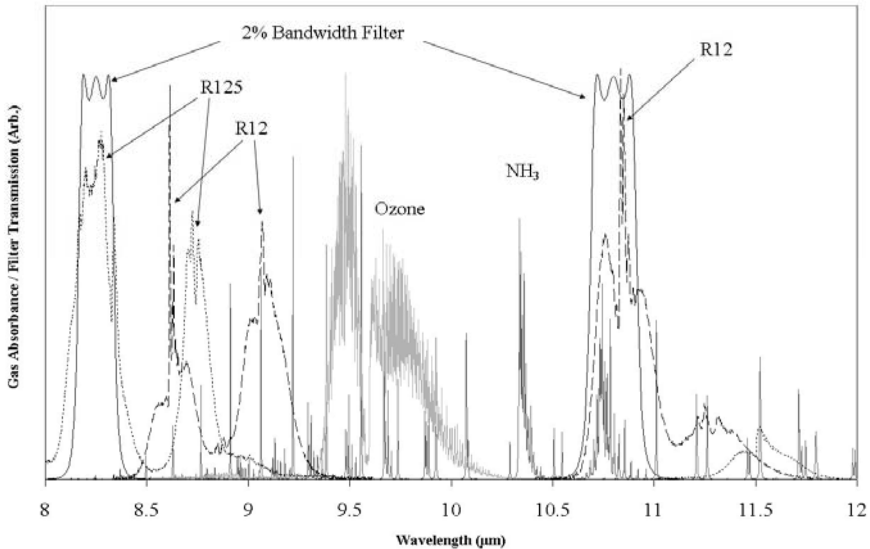


Figure 3. Absorption bands of common gases in the 8 to 12 μm atmospheric window. The figure includes data used with the kind permission of Infrared Analysis Inc [2], US EPA [3] and US NIST [4].

The various gas absorption bands are distributed across the infrared spectrum and this provides the means to distinguish one gas from another. Whether this is straightforward or not depends on which gases may be present in the sample under investigation. It should be noted that water vapour, which is present in most situations, unfortunately absorbs strongly in several parts of the infrared spectrum. It has a complex rotational structure because of its non-linear shape and exhibits many lines in its broad absorption bands. Furthermore, there are weaker infrared absorptions, typically by an order of magnitude, from vibrational overtones and combination bands. These can be useful for detecting gases in a more convenient part of the infrared spectral region but they can also make it harder to distinguish one gas from another.

3 Methods of Gas Detection

In competition with spectroscopic techniques for gas detection are a number of methods [5], of which we summarise some important ones below:

- Electrochemical – essentially, surface effects change the electrical properties. These have poor selectivity but a high sensitivity and are low cost
- Catalytic detectors for combustible gases, commonly known as pellistors
- Magnetic detectors for oxygen

- Colorimetric detectors, *e.g.* paper impregnated with lead nitrate for detecting H_2S by the black lead sulphide reaction product
- Chemical detection by fluorescence from a reaction, *e.g.* for NO_x and O_3

The electrochemical and catalytic methods can be inexpensive and very sensitive but sometimes suffer from the disadvantages of poisoning, shorter lifetimes, failing to danger and having low selectivity. For example, pellistors can be poisoned by common solvents and, since they operate by burning any gas present, will give a zero response if the concentration of the combustible gas is 100%. By contrast, infrared methods can be constructed to fail to safety, are immune to poisoning and can be highly selective. On the debit side, they tend to be more expensive than electrochemical sensors.

The reasons for the extra expense of infrared measurement reside mainly in the current cost of the infrared detector/filter package, the transfer optics and, in some cases, the cost of the sources. Here, we make quantitative comparisons of thermal sources and detectors, currently responsible for the vast majority of industrial infrared sensors, with semiconductor diode based sources and detectors. At present, the cost and performance of the semiconductor devices do not give them a clear advantage for most infrared applications. Conventional pyroelectric and thermopile detectors have already been developed to be low-cost, high volume items. The pyroelectric detector is used for passive intruder alarms and is usually fashioned as ‘opposed pairs’ to increase sensitivity to differential irradiation from faceted mirrors or lenses. Thermopile detectors have found volume application in infrared thermography. To compete effectively, semiconductor devices would need to attain similar high volumes and approach the low cost of the established thermal detectors. Thus, there is much promise in research activity in infrared diodes for both emission and detection, but convincing commercial advantages over thermal technology remain to be demonstrated to achieve widespread application to gas detection.

4 Infrared Sources for Gas Detection

A source need only be good enough, rather than ideal, and a number of preferred characteristics can be set down, against which to judge a particular source. Certainly, different applications will favour particular aspects of a source but a general guide can be drawn up as follows:

- High spectral radiance, *i.e.* >200 μW emitted power into a minimum 4% bandwidth, situated between 2 μm and 12 μm in a device a millimetre or so in size.
- Small emitting area (<20mm^2).
- Fast modulation (>1 Hz at 50% modulation depth, and preferably >50 Hz).
- Long term stability (<0.5% drift per year absolute or differential between two closely spaced wavelengths).
- Temperature stability of radiance (*e.g,*<200 ppm/ºC, either as an absolute value, for single channel gas sensors, or as a differential value between two closely spaced wavelengths for two-channel gas sensors).

- Expected life in normal operation (>5 years).
- Electrical drive power <10 W and preferably <0.5 W, drive voltage between 1V and 10V and drive current <1 A and preferably <100mA.
- Low manufacturing cost (<€10).
- High spectral efficiency (*i.e.* as little radiation as possible lost to wavelengths outside the spectral region of interest) – an important consideration for thermal sources but which is intrinsically a feature of LEDs.

4.1 Thermal Sources

Historically, the most flexible and well-known infrared source for the near infrared was the Nernst Glower. This was a bar of ceramic material heated by an interior electrical conductor, without any window, and operating at approximately 1000 ºC. This was very suitable for the slit geometry of spectrometers but over-specified for simple gas detectors. Circa 1980, in researching suitable infrared gas sensor systems for use in diving bells, the use of ordinary glass envelope tungsten lamps, with a drive voltage modulated at 8Hz, was introduced by one of the authors (JGC). This gave a 60% modulation of the infrared output and provided a convenient modulated source with no moving parts. When combined with infrared interference filters dating back to, for example, double half wave filters [6], we have a popular current method of providing appropriate wavelength-selected radiation for infrared gas detection. However, these lamp sources are limited to less than 5μm ($2000cm^{-1}$) by the cut-off of the glass envelope. The wavelength distribution of the radiation obeys the Planck distribution law (Figure 4), which peaks near the visible for the tungsten lamp, resulting in only a small proportion of the provided power being used. For IR gas detection, however, the extra heating available around the body of the sensors is often useful to prevent unwanted condensation.

Access to wavelengths longer than the glass cut-off is now available in small modulated sources, which have recently extended the useful range of this technology [7-10]. These are large area (>$4mm^2$), relatively low temperature (600 to 800°C), electrically heated, thin film devices. The substrate is often prepared with MEMS (Micro Electro-mechanical Systems) technology and the thin film may be silicon, metal, typically nichrome, or a diamond-like carbon (DLC). The emissivity can be tailored using, for example, a surface textured by a plasma or a photonic bandgap approach.

An example is referred to as a 'gray body' in Figure 4. The power spectral density is limited by the source temperature, area, background temperature and emissivity. A filter bandwidth of $50cm^{-1}$ is typically used for an infrared CO_2 gas sensor and the amount of energy then available, in a convenient geometry, is about 1mW, emitted into a hemisphere. In practice, we find that such sources, combined with a thermal detector having a D* of 10^9 $cmHz^{½}W^{-1}$, give more than adequate sensitivity for many gas sensing applications.

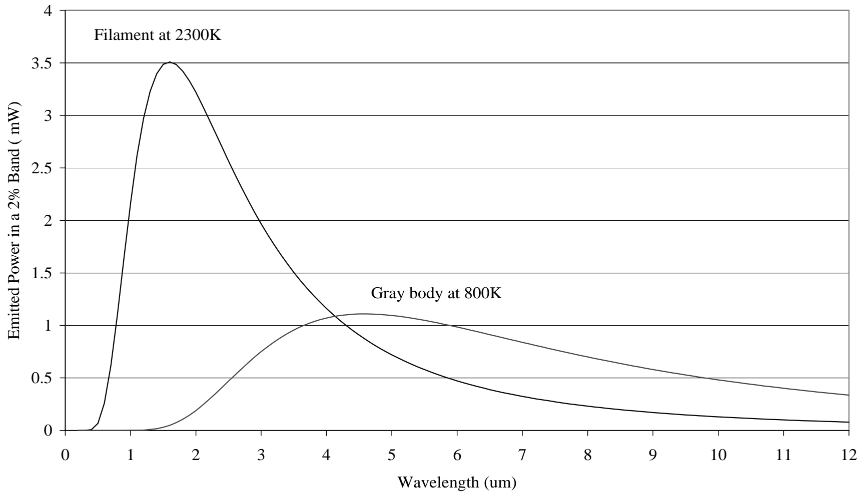


Figure 4. Two thermal sources modulated at 8Hz, a tungsten type filament at 2300 K in a vacuum with an area of $1mm^2$ and a gray body source at 800K with an area of 4 mm^2, both radiating into a hemisphere, in a 2% bandwidth. Note that Wien's Displacement Law does not apply directly to these curves since a 2% bandwidth is specified.

The temperature dependence of the emission from a filament is very small. Because the electrical power input to the lamp is mostly radiated, rather than conducted away, it requires a very small percentage change in this radiated power to adjust to a change in the equilibrium with the surroundings caused by a change in ambient temperature. Practically, from measurements, we infer a value of no more than $0.02\% K^{-1}$ for the temperature coefficient of radiation in the 2% bandwidth spectral region used for a gas sensor at 4.3 μm wavelength.

4.2 Semiconductor Sources

In the wavelength region of interest, from 2.5 to 15 μm, cost precludes the use of a laser or an optical parametric oscillator (OPO) in most applications. However, the advent of commercial quantum cascade (QC) lasers [11] does present possibilities for practical systems in the mid-infrared. Lasers, by their nature, do allow techniques not possible with LEDs, particularly for high-sensitivity gas detection using Cavity Ring-Down spectroscopy (CRDS) and Intra-cavity Laser Absorption Spectroscopy (ICLAS).

The most likely semiconductor sources for high-volume gas detection applications at the moment are the mid-infrared LEDs. Considerable research work and development effort have gone into extending room temperature LED (and laser) operation to longer wavelengths (from 3 to 12 μm), where thermally activated carriers reduce the radiative efficiency [12]. Excitation methods employed to this end include carrier injection, carrier extraction and optical pumping, and there is an extensive literature on these approaches. Carrier injection is the method employed in

conventional electrically stimulated LEDs but complexity in the structure of the LED has been added to improve performance. For example, barriers to confine electrons have been added to improve quantum efficiency and "band structure engineering", using quantum wells and superlattices, has been applied also to achieve this [13-28].

Carrier extraction, or negative luminescence, devices maintain a much reduced, non-equilibrium carrier concentration under reverse bias such that the device appears colder than the background [29-34]. The negative luminescence efficiency can be around 90%, which corresponds to a reduction in the blackbody emission from the device by an order of magnitude over a spectral region several microns wide in the mid-infrared [35]. For a room temperature background, the blackbody emission is modest, rising from a spectral radiance of 2.3 $\mu Wmm^{-2}sr^{-1}\mu m^{-1}$ at 5 μm wavelength to a value of 9.3 at 10μm. However, this is a usable output and has the merit of increasing with wavelength.

Optical pumping uses a near-infrared LED or diode laser to excite carriers in the mid-infrared device. This approach has been used out to 9μm wavelength and can produce a source radiance comparable with or exceeding that of electrically excited LEDs [36,37].

Other techniques applied during fabrication include the incorporation of cone concentrators to optically immerse the detector and increase the radiated output [32,38]. Also, the spectral bandwidth of the diode emission can be around 25% of the peak wavelength, which is several times larger than desirable to accommodate a 2% gas absorption bandwidth and a reference wavelength region. To reduce this emission bandwidth and improve the directionality of the output, resonant cavity optics have been added to LEDs. This approach has improved the LED characteristics but not yet as much as is theoretically possible [39,40].

A disadvantage of semiconductor sources is that, in general, their emission is strongly temperature dependent, because a semiconductor whose energy gap is close to kT at room temperature is intrinsically sensitive to temperature. For example, we have measured a temperature coefficient of about 1% K^{-1} for a non-equilibrium, room temperature, InSb LED at 6 μm. This is far larger than the 0.02% K^{-1} figure for a filament lamp. In practice, this requires either temperature compensation or the LED needs to be maintained at a constant temperature. A second disadvantage is that, without the advantage of high-volume production, they are significantly more costly than thermal sources. Third, a total output power of around 1 mW at a wavelength of 4 μm, falling to tens of μW at 10 μm, is smaller than the power emitted by thermal sources. However, an advantage is that semiconductor sources can be modulated at high frequencies, although to exploit this naturally requires a comparably fast detector.

5 Infrared Detectors for Gas Detection

The thermal detectors used for gas detection are generally the pyroelectric detector and the thermopile. Both have a specific detectivity (D*) around 10^8 cm $Hz^{½}$ W^{-1}. The pyroelectric detector is an a.c. device and so requires a modulated signal but this would be the normal mode of operation anyway. It has a high voltage responsivity, which decreases above a few Hertz, becoming inversely proportional to frequency.

For this reason, it is generally used below 100Hz. However, its NEP (Noise Equivalent Power) remains low up to about 1 kHz and the pyroelectric detector can, with care in the signal amplification, be used at such higher frequencies. It is also microphonic, which is a disadvantage though one that can be ameliorated by electronic signal processing.

The thermopile is also a low bandwidth device. Both its responsivity and noise performance fall with increasing frequency above a few Hertz and it is generally used only at low frequencies. When used with a modulated thermal source, the long thermal time constant of the source similarly restricts the modulation frequency and so thermal sources and detectors are a natural pairing for low frequency operation. Bolometers (temperature dependent resistors) have comparable performance but are being developed mainly as arrays for thermographic applications.

Photovoltaic and photoconductive detectors of an appropriate material can cover the entire mid-infrared region of interest, with high detectivity, when cooled but the cooling is an unwanted additional expense and a practical inconvenience. A good detectivity can be maintained in room temperature photodiode detectors out to about 4 μm [41-44].

At longer wavelengths, a novel design of photodiode, pioneered by QinetiQ [45,48], has been researched over a number of years and is a non-equilibrium device of similar design to the negative luminescence source described earlier. It comprises a four-layer structure that creates a non-equilibrium distribution of carriers under reverse bias through carrier extraction. In principle, this should provide a D* in excess of 10^{10} cmHz^{½}W^{-1} in a room temperature device. In practice, the application of the bias generates a large 1/f noise component that dominates up to 1 MHz, too high a frequency for serious consideration in a gas detection system [47]. However, even under zero bias, the device provides a D* in excess of 10^{9} cmHz^{½}W^{-1} with minimal cooling, for example by a Peltier cooler [48]. These devices have been investigated for InSb and CMT (Cadmium Mercury Telluride) systems, giving coverage from 4 to 12 μm [49]. Complementary source and detector pairs of this type can operate at very high modulation frequencies (~MHz) but their performance has yet to lead to commercialisation.

5.1 Optical Immersion

Semiconductor LED and optically pumped sources at mid-infrared wavelengths suffer greatly from total internal reflection at the interface of their planar emitting surface with air. The fraction of internal radiation that escapes from the device is only $1/n^2$ or $1/2n^2$, depending on whether the emission region is optically thick or not. For InSb, the refractive index is 4 and so only about 6% of the radiation can escape, at best.

The situation can be retrieved by fixing the emitting device to an immersion lens, where the refractive index of the lens material must be at least equal to the index of the emitter. Germanium is suitable for the mid-infrared and readily available. It is important that the LED and lens be in close optical contact and this presents practical difficulties in the mid-infrared where there is no glue of similar refractive

index. The LED and lens must then be brought into close proximity (typically $\lambda_{air}/40$) over the whole emitting area of the source and without causing damage.

The effect of a hemispherical lens (with a diameter n times the size of the emitter) is to allow all the radiation to escape into the lens and subsequently out of the lens provided that the surface is anti-reflection coated. A second effect is that the apparent size of the source is increased by a factor of n^2. Note therefore that the overall brightness (radiance) of the source is unchanged but the source is larger and more powerful. A hyperspherical lens, as shown in Figure 5, causes the apparent size of the source to increase by n^4 and the emitted radiation to be confined to a cone with semi-angle $\sin^{-1}(1/n)$. Thus, the source radiance is again unchanged but the narrower emission cone may be useful practically. Unlike the hemispherical case, the apparent source position is also changed.

A conventional lens, as described here, is bulky and a diffractive optical element (DOE) may offer a more attractive solution. A suitably prepared DOE could be used as the substrate for the emitting device in some instances, giving a wafer-scale production method.

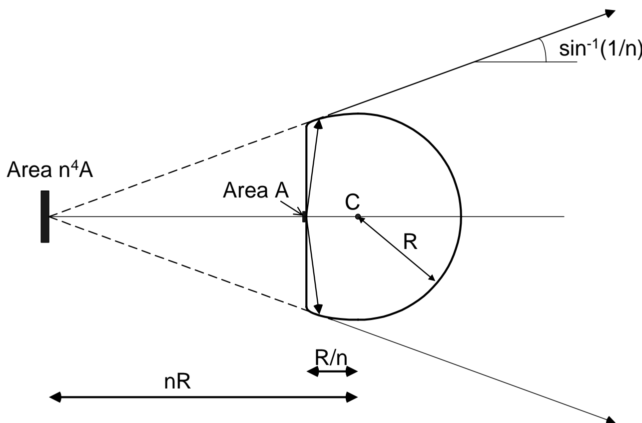


Figure 5. Hyperspherical immersion of a detector of area A (from Crowder *et al.* [50] with permission of IOP Publishing Ltd)

Photodiode detectors can also gain from immersion techniques. The D* value can be increased by the addition of a hemispherical immersion lens because the apparent size of the detector is increased by n^2 while the detector noise remains unchanged. Since D* is given by $A^{½}$/NEP, where A is the detector area and NEP is its Noise Equivalent Power, D* will increase by n.

If a detector of a particular apparent size is required, the introduction of an immersion lens of this type allows the detector to be smaller in area by n^2 and this increases the electrical resistance of the detector. For infrared detectors at or near room temperature, this increases their generally low resistance by a significant factor, which can be extremely useful for matching the detector to an amplifier. When the detector resistance is small (< 50 ohms), the front-end voltage noise of the amplifier will be larger than the detector noise and will degrade the signal to noise ratio of the measurement.

A hyperspherical lens can also be used with a detector. In this case, the D* value is further increased compared with the hemispherical lens but the field of view of the detector is reduced. The reduced field of view may be of no concern while the detector's field of view is obscured anyway by its packaging but ultimately it will restrict the light-gathering capability of the detector. A laboratory-based, low-power gas detection system with high gas sensitivity using such immersed components is described later, with a quantitative consideration of the optical performance [50].

6 Design of Optical and Gas Sampling Systems

A single-channel instrument measures the amount of infrared radiation from the source that reaches the detector having passed through the optical system, the gas sample volume and the optical filters that define the spectral region of interest. This measurement is affected not only by the amount of gas present but also, for example, by variations in the source strength, detector sensitivity and any deterioration in the transmission of optical surfaces exposed to the atmosphere being monitored. To minimise these additional effects, it is usual to employ a reference channel that uses a wavelength not absorbed by the gases that may be present. The ratio of sample channel to reference channel then provides a far better measure of the absorbing gas present. This approach is not necessary if a calibration measurement can be made immediately prior to the gas measurement though, of course, this too could rightly be construed as a reference measurement.

The two-channel system can be set up as two channels in time or two channels in space. In a common implementation of the former, a rotating filter/chopper wheel interposes the optical channel filters sequentially in the light path. This ensures that essentially the same optical path is used for both channels, which gives good stability but at the expense of incorporating a moving part in the system. In a two channels in space system, there are no moving parts but the light travels from source to detector along two different paths and any relative change between the paths, other than caused by the gas, will give rise to a false reading. The two paths will, for that reason, be kept overlapped and common for as large a part of the path as possible.

The filters employed are usually optical interference filters but, in cases where the absorption bands of the gases overlap, a sealed gas cell containing a sample of the gas to be detected can be added to a channel, usually mounted in a rotating filter wheel. When such a cell is used to modulate the light beam, the spectral modulation is highly specific to the gas being measured. This correlation spectroscopy method requires that the gas absorption lines be sufficiently well separated that the gas cell is not simply opaque within the gas absorption band. The regular spacing of transmission peaks in a Fabry-Perot etalon might be thought a suitable alternative but, unfortunately, gas absorption lines steadily deviate from an even spacing and so using an etalon as a comb filter is of limited use.

The optical system can also be imaging or non-imaging. If the source is non-uniform or if the system employs two filter/detectors then mechanical distortion, due for example to temperature gradients, can cause a signal change, *i.e.* a drift in the measured gas concentration. In a non-imaging system, a light pipe is often

employed, which scrambles any source non-uniformity yet still allows a good collection efficiency. In both imaging and non-imaging optical systems, the efficiency of the collection optics is governed by the concept of 'étendue', also known as geometric extent, throughput, acceptance or Lagrange invariant [51]. It is a measure of the light gathering capability of an optical system and also governs how the beam diameter must increase if the angular size of the beam is reduced, and vice versa. The etendue in an axially symmetric system is given by

$$U = n^2 a^2 \sin^2\theta \tag{1}$$

where a is the radius of the limiting circular aperture that encloses the light beam, θ is the semi-angle of the cone enclosing the light beam and n is the refractive index of the medium as shown schematically in Figure 6. The etendue is important because it is invariant through an optical system provided that there are no obstructions in the light beam and provided that absorption and scattering are negligible.

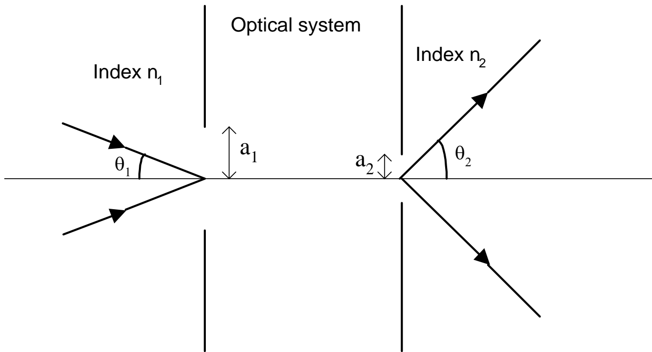


Figure 6. Etendue of an optical system

If the radiance (brightness) of the light beam at the aperture is L then, since the area of the aperture is πa^2, the optical power flux through the system will be

$$L\,\pi^2 a^2 n^2 \sin^2\theta \tag{2}$$

or

$$L\,\pi^2 U \tag{3}$$

This power flux will not change and so both the radiance and etendue are constant through the system. If a source with large etendue is coupled to an optical system with low etendue then light will be lost. To couple the most radiation into a system, a source should have as large a radiance as possible and should fill both the area and acceptance angle of the optical system. Similar considerations apply when coupling light from the optical system to the detector(s). Since the maximum value for $\sin\theta$ is unity, there is a limit to how small an area the light can be concentrated into. If the detector is smaller than this then light will be lost. In practice, detectors will not have a full hemisphere for their field of view and so the light can only be

concentrated onto a somewhat larger area. For an axially symmetric system, we can simply use the constancy of u, where

$$u = n\, a\, \sin\theta \tag{4}$$

to determine the interplay of aperture and divergence.

6.1 A Non-imaging Gas Sensor with Thermal Source and Detector

A natural design for a gas sensor is an imaging system incorporating a collimating lens or mirror at the source and a condensing optic near the detector, with an optical filter placed in front of the detector. (The properties of optical interference filters allow the use of convergent rays [52]). In practice, an arguably simpler and cheaper optical arrangement is the use of a gold coated tubular light guide, typically 100mm in length, with no further optics between the source and detector. This non-imaging design has proved to be effective in, for example, the Edinburgh Instruments Ltd Gascard, a schematic of which is shown in Figure 7. Although this simple design is relatively inefficient in radiation gathering at the light guide exit, it has proven to be stable over the long term and against temperature variation. We discuss its properties here as an example of its type.

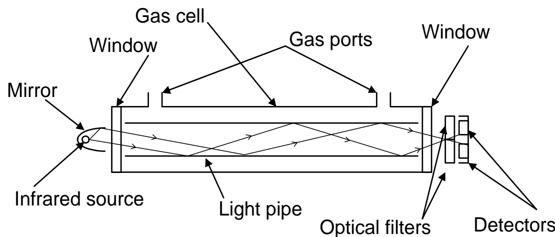


Figure 7. Gas sensor using a light guide

This sensor uses a dual detector to provide a two-channel system with the spectral region of each channel being defined by an optical interference filter directly in front of each detector. The length of the light guide can be chosen up to a maximum of 100mm to accommodate full scale readings of, for example, 100% to 1000ppm of carbon dioxide. The source is fitted with a concave reflector, not to collimate the light, but to generate an f/1 beam within the light guide. This is to provide sufficient reflections at the guide wall to produce, at the exit of the light guide, a smooth intensity profile with a broad central maximum,.

The discussion of etendue showed that the concentration of light in an optical system has a limit and this is directly relevant to the optics here in showing by how much the absence of collection optics reduces the detected light below what is theoretically achievable. The exit aperture of the light pipe has a 7mm diameter and the angular extent of the beam has a semi-angle of about 26° corresponding to the designed f/1 optics. If the detector is a flat single-sided absorber of 1.5mm diameter,

a collection efficiency of 4.6% is the best that can be achieved, being the ratio of areas. It will be somewhat less if the plane of the detectors is distant from the exit of the guide. If the light leaving the guide were collected by a concentrator [51] then the collection efficiency would be a maximum when the field of view of the detector was at its theoretical maximum, *viz.* the whole hemisphere. In that case, the beam would be reduced to a diameter of 3.1 mm and, given the size of the detector, the ratio of areas would give a collected fraction of 23%, an increase by a factor of about 5. However, the detector does not have such a wide field of view and, moreover, the collection angle of 26° is as large as should be used with the optical interference filters. Beyond this value, the characteristics of the optical interference filter are degraded [52].

Thus, as covered above, collecting nearly 4.6% of the light at the exit of the light pipe is as good as can be done with a detector field view of semi-angle 26° given the beam divergence introduced in the light guide. There are, of course, two adjacent detectors and also the beam intensity is weaker away from the axis of the light guide and so the actual arrangement is not as far from optimum as might at first appear. As might be expected, there is a trade-off between optical complexity and optical efficiency. In practice, in the Gascard case, there is an estimated 1mW of light, within the spectral bandwidth of each filter, that enters the guide and the light flux reaching each detector is between 10 and 50 μW.

Table 1. Performance of Edinburgh Instruments Gascard

Model	Gas	Accuracy[1] % of range	Stability % of range, over 12 months	Repeatability @ zero	Repeatability @ span
Gascard II (0-3000ppm)	CO_2	± 2%	± 2%	± 0.3%	± 1.5%
Gascard II (0-100%)	CO_2	± 2%	± 2%	± 0.3%	± 1.5%
Gascard II (0-2000ppm)	C_6H_{14}	± 2%	± 2%		
Gascard II (0-2000ppm)	Hydro-carbons	± 2%	± 2%		

Response Time: T_{90} = 10 s (= 30 s for diffusion model)
Operating Temperature: 0-45ºC
Warm-up Time: 1 min (initial), 30 min (full specification)
Humidity: Measurements are unaffected by 0-99% relative humidity, non-condensing.

[1]stated accuracy includes a calibration gas tolerance of ±1%

To relate to diode sources, the operating wavelength for a carbon dioxide sensor is normally the 4.3 μm band and the spectral radiance of LEDs in this region currently offers no advantage over thermal sources. The advantage of the thermal source is then clearly the very low cost of a filament lamp.

6.2 A Long-path, Imaging Gas Sensor with Semiconductor Source and Detector

In contrast to the preceding section, we describe here a gas sensor using infrared semiconductor devices for the ppm-level detection of nitrogen dioxide at a wavelength of 6.2 μm. This system uses room temperature, immersed, semiconductor diode devices for both source and detector, which provides a very low power requirement while retaining good signal to noise at the detector. The system was set up as a laboratory demonstrator and benefits from the simpler optics of a single measurement channel. The high sensitivity is obtained from the use of a White cell, which is a multiple-pass system. Here, the White cell has an overall length of about 300 mm and the 4m path in the gas mixture contained within the cell is provided by 20 passes. A schematic, showing just four passes, is given in Figure 8.

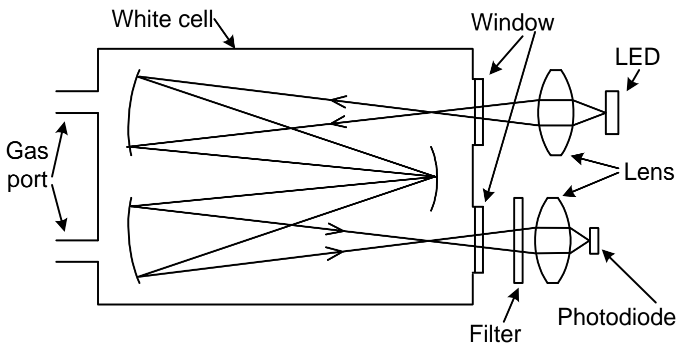


Figure 8. Long path length gas sensor with semiconductor source and detector (from Crowder *et al.* with permission of IOP Publishing Ltd. [50])

A difficulty with using a White cell is the small acceptance angle of the optics The internal optics of the cell used have an f/number in the horizontal plane of f/6.4 and, in the vertical plane, f/4.6, with an aperture that is 10 mm high and 9 mm wide. When the emission angle of the source radiation is reduced optically to this acceptance angle, the conservation of etendue forces the light beam to a larger radius. However, the White cell aperture is considerably larger than the semiconductor source and so all the source radiation can be coupled into the cell.

The source is an InSb-based LED with a diameter of 82 μm and it incorporates a near-hyperspherical immersion lens that gives an apparent source size of 1.1 mm diameter (area = 1 mm^2). The lens causes the radiation to be emitted in a cone of semi-angle 15°, which is then easily coupled into the White cell with a convex, calcium fluoride lens, which, from etendue considerations, increases the source radiation to a diameter of 5.5 mm. At the exit of the White cell, a lens images the exit aperture onto the detector through an optical interference filter with a centre wavelength of 6.2 μm, a 2.2% bandwidth and a 60% peak transmission. The photodiode detector has the same dimensions as the source, is of similar construction and is similarly fitted with a near-hyperspherical immersion lens. The light from the White cell can therefore be collected efficiently. With a time constant of 10 s, the presence of nitrogen dioxide was measured to 2 ppm (limited by system noise), which is a sensitivity appropriate, for example, to NOx detection in vehicle exhausts. The emission spectrum of the immersed LED source extends from 4.5μm to 7μm, with a maximum near 6μm, and gives a spectrally integrated output of 1.6 μW at a drive current of 4 mA. This is a modest output power and could be increased by higher drive current (a factor of 5) and increasing the source size (subject to the White cell etendue in this system). However, as above, a useful sensitivity to NO_2 was obtained with a power dissipation in the LED of only 0.25 mW.

The immersed detector has a D* of 4.5 x 10^9 $cmHz^{1/2}W^{-1}$ with a field of view of 30°. Importantly, the small physical size of the detector (82 μm diameter) gives it a resistance of 35 ohms. The signal from such a source resistance can be amplified without being overwhelmed by the front-end voltage noise of an amplifier by using a small number of amplifiers in parallel. However, a detector with a physical size of 1.1mm diameter would have had a resistance of less than 1 ohm and would require less convenient impedance matching, *e.g.* by transformer or resonant coupling.

In this set of measurements, the LED and detector were both at room temperature. In a practical instrument, the temperature dependence of the LED and detector characteristics would need to be addressed. Temperature stabilisation using a thermoelectric device would be simple and also attractive since a modest cooling would improve the detector's detectivity, raise the detector's resistance and, in general, increase the LED output. Indeed, overall, this system represents a realistic starting point for a practical instrument for gas monitoring at low gas concentrations.

7 Laser Techniques

This wide area will be described only briefly. Systems employing lasers or optical parametric oscillators for gas LIDAR are well established commercially for applications where their cost is not a barrier. Currently, of particular interest are the techniques of Cavity Ring-Down Spectroscopy (CRDS) and Intra-Cavity Laser Absorption Spectroscopy (ICLAS). Both of these approaches give very high sensitivities because of the large effective path length of light within the optical cavity (~100 km). Trace gas detection (parts per billion) is possible, with measurable absorption coefficients as low as 10^{-10} cm^{-1}. The progress in developing room temperature infrared semiconductor lasers, *e.g.* vertical cavity, surface emitting

(VCSEL) devices and quantum cascade (QC) lasers, has been a key factor in making these systems more attractive.

In cavity ring-down spectroscopy, a short laser pulse enters the optical cavity, which contains the gas to be measured. On each pass of the cell, some light leaves the cell and its amplitude is measured by a fast detector. The rate of decay of the signal amplitude is quicker when an absorbing gas is present and this provides the method of measurement, with the gas being removed from the cell to get a system calibration. Variants on the method include the use of CW lasers and fast optical switching [53].

Intracavity laser absorption spectroscopy requires a laser with a homogeneously broadened gain bandwidth greater than the linewidth of a gas absorption line [54] and a laser cavity long enough that several axial modes lie within the gas linewidth. When the laser pulse begins, the laser modes compete and the developing laser generation spectrum reveals the presence of gas absorption. The effective path length in the gas is given by the elapsed time of the pulse times the speed of light [55]. At present, both systems remain generally laboratory-based because of relative complexity and cost.

8 Conclusions

It might reasonably be expected that semiconductor devices would inevitably displace thermal devices in all aspects of gas sensing. However, thermal sources and detectors are inexpensive and suffice for a very wide range of gas sensing applications. The advent of thin film sources provided the necessary extended wavelength coverage beyond 4.5 μm compared with glass-envelope filament lamps, and the emitted power is generally higher for thermal sources.

Because of the complexity of fabricating semiconductor devices with the materials used, there is a significant economic barrier to them entering the marketplace. They do offer the advantages of being robust and having a long lifetime and, when used as both source and detector, much higher modulation frequencies are available. The temperature dependence of semiconductor devices is higher than thermal devices, which requires temperature stabilisation, or compensation. Incorporating a modest cooling with the temperature stabilisation in semiconductor devices would be advantageous but at the expense of electrical power and some complexity. It is, perhaps, where reliability is paramount that semiconductor systems are likely to be introduced.

References

1. "Fundamentals of Molecular Spectroscopy", Banwell C N (McGraw-Hill, 1983)
2. Infrared Analysis, Inc. (Anaheim, CA,USA) see www.infraredanalysisinc.com
3. US Environmental Protection Agency. See www.epa.gov
4. US National Institute of Standards and Technology. See www.nist.gov
5. "Techniques and Mechanisms in Gas Sensing", Ed. Moseley P T, Norris J O W and Williams D E, Publ. IOP (1991)
6. Smith S D, "Design of multilayer filters by considering two effective interfaces", J. Opt. Soc. Am. **48** 43 (1958)

7. Hawkeye Technologies (Connecticut, USA), see www.hawkeyetechnologies.com
8. Leister Microsystems (Sarnen, Switzerland), see www.leister.com/english/leister_english.html
9. Cal Sensors Inc. (California, USA), see www.calsensors.com
10. Ion Optics (Mass., USA), see www.ion-optics.com
11. Alpes Lasers (Neuchatel, Switzerland), see www.alpeslasers.ch
12. Smith S D, Hardaway H R & Crowder J G, "Recent developments in the applications of mid-infrared lasers, LEDs, and other solid state sources to gas detection", Proc. SPIE 4651 157-172, May 2002.
13. M.J.Kane, G.Braithwaite, M.T.Emeny, D.Lee, T.Martin, D.R.Wright, "Bulk and surface recombination in InAs/AlAsSb 3.45μm light emitting diodes", Appl.Phys.Lett. **76** (8) 2000, p.943-945
14. T.Ashley, C.T.Elliott, N.T.Gordon, R.S.Hall, A.D.Johnson, G.J.Pryce, "Room temperature narrow gap semiconductor diodes and sources and detectors in the 5-10μm wavelength region.". J.Crystal Growth **156,** 1996, p.1100-1103
15. A. Popov, V.V.Sherstnev, Y.P.Yakovlev, A.N.Baranov, "Powerful mid-infrared light emitting diodes for pollution monitoring", Electron. Lett. **33** (1) 1997, pp 86-88
16. A.Krier, V.V.Sherstnev, "Powerful interface light emitting diodes for methane detection" J.Phys.D:Appl.Phys. **33**, 2000, p.101-106
17. A.Krier, V.V.Sherstnev, H.H.Gao, "A novel LED module for the detection of H_2S" J.Phys.D:Appl.Phys. **33**, 2000, p.1656-16661
18. V.V.Sherstnev, A.M.Monahov, A.Krier, G.Hill, "Superluminescence in InAsSb circular-ring-mode light emitting diodes for CO gas detection" Appl.Phys.Lett. **77** (24) 2000, p.3908-3910
19. A.Krier, H.H.Gao, V.V.Sherstnev, Y.Yakovlev, "High power 4.6μm light emitting diodes for CO detection" J.Phys.D: Apll.Phys. **32**, 1999, p.3117-3121
20. X.Y.Gong, H.Kan, T.Makino, T.Iida, K.Watanbe, Y.Z.Gao, M.Aoyama, N.L.Rowell, T.Yamaguchi, "Room-temperature mid-infrared light-emitting diodes from LPE InAs/InAsSb/InAsSbP heterostrucutres" Jpn.J.Appl.Phys. **39**, 2000, p.5039-5043
21. X.Y.Gong *et al*, "Light emitting diodes fabricated from liquid phase epitaxy InAs/InAsSbP/InAsSbP and InAs/InAsSb multi-layers" Cryst.Res.Technol. **35** (5) 2000, p.549-555.
22. X.Li, J.Heber, M.Pullin, D.Gevaux, C.C.Phillips, "MBE growth of mid-infrared antimonide LEDs with strained electron barriers, Journal.Crystal Growth, **227**, 2001, p.600-604
23. P.J.P.Tang, H.R.Hardaway, J.Heber, C.C.Phillips, M.J.Pullin, R.A.Stradling, W.T.Yuen, L.Hart, "Efficient 300K light emitting diodes at 5 and 8μm from InAs/InAsSb single quantum wells" Appl.Phys.Lett. **72** (26) 1998, p.3473-3475
24. M.J.Pullin, H.R.Hardaway, J.D.Heber, C.C.Phillips, W.T.Yuen, Ra.A.Stradling, P.Moeck, "Room-temperature InAsSb strained layer superlattice light emitting diodes at λ=4.2μm with AlSb barriers for improver carrier confinement", Appl.Phys.Lett. **74** (16) 1999, p.2384-2386
25. B.Gietens, S.Nemeth, C.Van Hoof, P.Van Daele, G.Borghs, "Growth and characterisation of InAsSb/InAs strained multiple quantum well light emitting diodes grown on InAs substrates" IEE Porc.-Optoelecton. **144** (5) 1997, p.295-298
26. R.M.Biefeld, S.R.Kurtz, A.A.Allerman, "The metal-organic chemical vapour deposition growth and properties of InAsSb mid-infrared (3-6μm) lasers and LEDs" **3** (3) 1997, p.739-748
27. "Demonstration of the cascade process in InAs/GaInSb/AlSb mid-infrared light emitting devices" Appl.Phys.Lett. **72** (12) 1998, p.1495-1497
28. Gevaux DG, Green AM, Phillips CC, Vurgaftman I, Bewley WW, Felix CL, Meyer JR, Lee H, Martinelli RU, "3.3 μm 'W' quantum well light emitting diode", IEE P-Optoelectron **150** 351-355 (2003)

29. Lindle JR, Bewley WW, Vurgaftman I, Kim CS, Meyer JR, Johnson JL, Thomas ML, Piquette EC, Tennant WE, "Negative luminescence from large-area HgCdTe photodiode arrays with 4.8-6.0 μm cutoff wavelengths", IEEE J Quantum Elect **41** 227-233 (2005)
30. Lindle JR, Bewley WW, Vurgaftman I, Meyer JR, Varesi JB, Johnson SM, "Negative luminescence from MWIR HgCdTe/Si devices", IEE P-Optoelectron **150** 365-370 (2003)
31. L. J. Olafsen *et al*, "Negative luminescence from type II InAs/GaSb superlattice photodiodes", Appl.Phys.Lett **74** (18), 1999, pp2681-2683
32. Nash GR, Gordon NT, Hall DJ, Ashby MK, Little JC, Masterton G, Hails JE, Gless J, Haworth L, Emeny MT, Ashley T, "Infrared negative luminescent devices and higher operating temperature detectors", Physica E-Low-dimensional systems & nanostructures **20** 540-547 (2004)
33. Remennyi MA, Matveev BA, Zotova NV, Karandashev SA, Stus NM, Talalakin GN, "InGaAsSb negative luminescent devices with built-in cavities emitting at 3.9 μm", Phsica E **20** 548-552 (2004)
34. T. Ashley, N. T. Gordon, G. R. Nash, C. L. Jones, C. D. Mazey and R. A. Catchpole, "Long wavelength HgCdTe negative luminescent devices", Appl. Phy. Lett. **79**, no.8, pp1136-1138, 2001
35. Haigh MK, Nash GR, Gordon NT, Edwards J, Graham A, Giess J, Hails JE, Houlton M, "Long-wavelength HgCdTe on silicon negative luminescent devices", Appl. Phys. Lett. **86** 011910 (2005)
36. M Aidaraliev, N V Zotova, S A Karandashev, B A Matveev, M A Remennyi, N M Stus and G N Talalakin, "Optically Pumped "Immersion-Lens" Infrared Light Emitting Diodes Based on Narrow-Gap III-V Semiconductors", Semiconductors, vol 36, No. 7, pp828-831, 2002.
37. Zotova NV, Karandashev SA, Matveev BA, Remennyi MA, Stus' NM, Tarakanova NG, "Luminescence of multilayer structures based on InAsSb at lambda=6-9 μm", Semiconductors **39** 214-217 (2005)
38. Nash GR, Ashley T, Gordon NT, Jones CL, Maxey CD, Catchpole RA, "Micromachined optical concentrators for IR negative luminescent devices", J. Mod. Optics **49** 811-820 (2002)
39. E.Hadji, J.Bleuse, N.Magnea, J.L.Pautrat, "3.2μm infrared resonant cavity light emitting diodes" Appl.Phys.Lett. **67** (18) 1995, p.2591-2593
40. Green AM, Gevaux DG, Roberts C, Phillips CC, "Resonant-cavity-enhanced photodetectors and LEDs in the mid-infrared", Physica E **20** 531-535 (2004)
41. Green AM, Gevaux DG, Roberts C, Stavrinou PN, Phillips CC, "lambda approximate to 3 μm InAs resonant-cavity-enhanced photodetector", Semicond. Sci. Tech. **18** 964-967 (2003)
42. Matveev BA, Aydaraliev M, Zotova NV, Karandashev SA, Il'inskaya ND, Remennyi MA, Stus' NM, Talalakin GN, "Flip-chip bonded InAsSbP and InGaAs LEDs and detectors for the 3μm spectral region", IEE P-Optoelectron **150** 356-359 (2003)
43. Chakrabarti P, Krier A, Morgan AF, "Analysis and simulation of a mid-infrared P+-InAs0.55Sb0.15P0.30/n(0)-InAs0.89Sb0.11/N+-InAs0.55Sb0.15P0.30 double heterojunction photodetector grown by LPE", IEEE T Electron Dev **50** 2049-2058 (2003)
44. Chakrabarti P, Krier A, Morgan AF, "Double-heterojunction photodetector for mid-infrared applications: theoretical model and experimental results", Opt. Eng. **42** 2614-2623 (2003)
45. T Ashley, C T Elliott and A T Harker, Infrared Phys. **26** 303 (1986)
46. M K Haigh, G R Nash, N T Gordon, J Edwards, A Graham, J Giess, J E Hails and M Houlton, Appl Phys Lett **86** 011910 (2005)

47. Ashby MK, Gordon NT, Elliott CT, Jones CL, Maxey CD, Hipwood L, Catchpole R, "Investigations into the source of 1/f noise in HgxCd1-xTe diodes", J. Electron. Mat. **33** 757-760 (2004)
48. Crowder JG, Elliott CT, Hardaway HR, "High performance, large area, uncooled detectors for mid-infrared wavelengths", Electron. Lett. **37** 116-118 (2001)
49. Hall DJ, Buckle L, Gordon NT, Giess J, Hails JE, Cairns JW, Lawrence RM, Graham A, Hall RS, Maltby C, Ashley T, "High-performance long-wavelength HgCdTe infrared detectors grown on silicon substrates", Appl. Phys. Lett. **85** 2113-2115 (2004)
50. Crowder J G, Hardaway H R & Elliott C T. ”Mid-Infrared Gas Detection using Optically-Immersed, Room Temperature, Semiconductor Devices”, Meas. Sci. Technol. 13 (2002) 882-884
51. "High Collection Nonimaging Optics", W T Welford and R Winston (Academic Press, 1989)
52. Pidgeon C R and Smith S D, "Resolving Power of Multilayer Filters in non-parallel light", J. Opt. Soc. Am. **54** 1459 (1964)
53. Y. B. He and B. J. Orr, “Cavity ringdown spectroscopy: New approaches and outcomes”, J. Chin. Chem. Soc. **48** (3) pp 591-601 2001.
54. L. A. Pakhomycheva, E. A. Svirichenkov, A. F. Suchkov, L. V. Titova, and S. S. Churilov, JETP, Lett **12**, 4, 1970.
55. A. Garnache, A. A. Kachanov, F. Stoeckel, and R. Planel “High sensitivity intracavity laser absorption spectroscopy with vertical-external-cavity surface emitting semiconductor lasers”, Optics Letters, **24** 826-828, 1999

Mid-infrared Biomedical Applications

I.K. Ilev and R.W. Waynant
US Food and Drug Administration
Division of Physics,
Office of Science and Engineering Laboratories,
Center for Devices and Radiological Health, HFZ-130,
12725 Twinbrook Parkway,
Rockville, M D 20857

1 Introduction

The mid-infrared (mid-IR) spectral region is of great importance in biophotonics research and instrumentation because of its specific sensitivity to so many of the building blocks of life. The most identifiable molecules have specific absorption and radiation features in this range that depend upon their exact chemical composition [1-3]. The unique specificity of the biological molecule's mid-IR spectra opens an extensive potential for development of minimally invasive, effective and safe medical diagnostic techniques based on such highly sensitive biophotonics approaches as mid-IR spectroscopy, biosensors and imaging [4-15]. Moreover, the mid-IR spectral range is attractive for highly precise surgical procedures and medical ablation of both soft and hard tissue because of the following main reasons: 1) it has absorption-dominant properties and its high absorption in water is associated with minimal penetration depth within tissue; 2) the mid-IR tissue ablation effects are superficial with submicron ablation rates and with minimal submicron damage layers around the exposed area; and 3) the mid-IR radiation lacks the carcinogenetic potential that has generated concern at ultraviolet laser tissue interaction [16-43]. However, the mid-IR region presents significant technological limitations to the functioning of the basic optical components of a biophotonics delivery and imaging system. These limitations can be associated with the lack of effective broadband tunable laser sources, low-loss fiber delivery systems and high-sensitive imaging techniques [44-50].

In this chapter, we present the current status as well as future trends in the most significant and extensively used mid-IR biomedical applications. The first part of the chapter summarizes the fundamental principles and features of mid-IR biophotonics applications. It includes description of a basic mid-IR biophotonics delivery system as well as key optical elements of this system such as: typically used mid-IR biomedical lasers and incoherent light sources including their critical parameters and characteristics, mid-IR biomedical delivery fibers, and some modern all-hollow-waveguide mid-IR laser delivery systems. The second part of the chapter is focused on specific mid-IR biophotonics applications such as precise laser surgery and tissue ablation, bioimaging and biosensors, and mid-IR biomedical spectroscopy.

2 Mid-IR Biophotonics

The field of mid-IR biophotonics is a substantial part of the extensively developed modern photonics science and technologies. This field is based on specific fundamental features of the mid-IR spectral range including its water absorption dominant properties, molecular spectral specificity and minimal penetration depth within tissue. In this *Section*, we summarize the basic principles of mid-IR biophotonics applications that include specific features and components of a mid-IR biophotonics delivery system such as coherent laser and incoherent light sources, and mid-IR biomedical delivery fiber systems.

2.1 Biophotonics

Biophotonics is a modern cross-disciplinary scientific field (illustrated in Figure 1) involving developments in areas of laser physics, electro-optics, optical fibers, detectors, sensors, and imaging on one hand, and cell and molecular biology, medical science and new biomaterials on the other. This is a science of generating and using light (photons) for imaging, sensing and manipulating biological materials. In biology, biophotonics is utilized to probe for molecular mechanisms, functions and structures. In medicine, it is used to study light-tissue interaction mechanisms at the macro (large-scale) and micro/nano organism level to detect, diagnose and treat diseases.

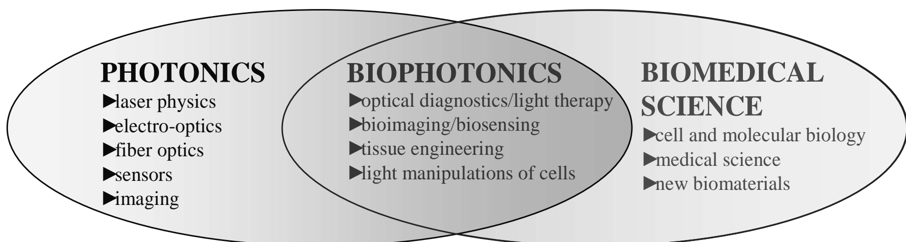


Figure 1. Biophotonics as a cross-disciplinary scientific field between photonics and biomedical science

The biophotonics field is a rapidly emerging medical technology. Minimally invasive biophotonics imaging and biosensor techniques are potential alternatives to conventional medical methods for diagnostics and monitoring of diseases, drug discovery, proteomics, and environmental detection of biological agents. These techniques offer an effective, fast and painless way for sensing and monitoring of various biomedical quantities. Devices utilizing minimally invasive biosensor technology are rapidly finding their way into the mainstream for early disease diagnosis and improved patient acceptance and comfort. As part of the whole area of biophotonics science and technology, mid-IR biophotonics basic principles and applied systems possess specific advanced features and critical parameters which will be considered below.

2.2 Fundamentals of Mid-IR Biophotonics Applications

In the mid-IR spectral range the optical properties of biological tissue, and, therefore, the associated mid-IR biophotonics applications, are governed by the following fundamental factors.

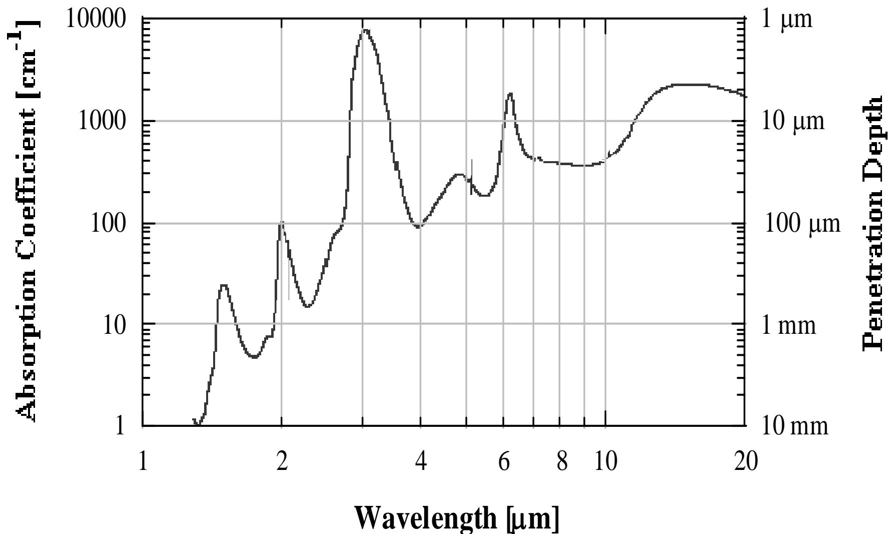


Figure 2. Absorption spectrum of water and penetration depth in the 1-20 mm IR spectral range, after D.Wieliczka, S.Weng and M. Querry, "Wedge shaped cell for highly absorbent liquids: infrared optical constants of water," Applied Optics, v. 28, pp.1714-1719, 1989

1. Water absorption dominates tissue properties

The mid-IR spectral range has absorption-dominant properties and the primary absorption component in biological tissue is water. Figure 2 illustrates a typical water absorption spectrum in the IR range between 1 and 20 µm [51,52]. Water absorption increases through the mid-IR to a maximum at approximately 3 µm, and with some additional fundamental OH absorption bands around 1.9, 4.8 and 6.2 µm.

2. Minimum penetration depth in tissue

Since water is the major constituent (~70-80 %) in virtually every type of soft tissue, it has been demonstrated both theoretically and experimentally that the high absorption of mid-IR laser radiation by water in tissue, particularly in the highly absorbing OH spectral bands (see Figure 2), will lead to a minimum penetration depth. Typical mid-IR penetration depths are in the range 1 µm – 1 mm as shown in Figure 2 (see the right vertical axis). This fundamental mid-IR feature opens a great potential for various laser systems using this spectral range to be exploited for highly effective tissue ablation and precise surgical procedures with minimal submicron tissue damage effects.

3. Molecular spectral specificity

The mid-IR is an attractive area for medical research and instrumentation because of its specific sensitivity to various substantial biomolecular components. This is the range where the most identifiable molecules have specific absorption and radiation features that depend upon their chemical composition.

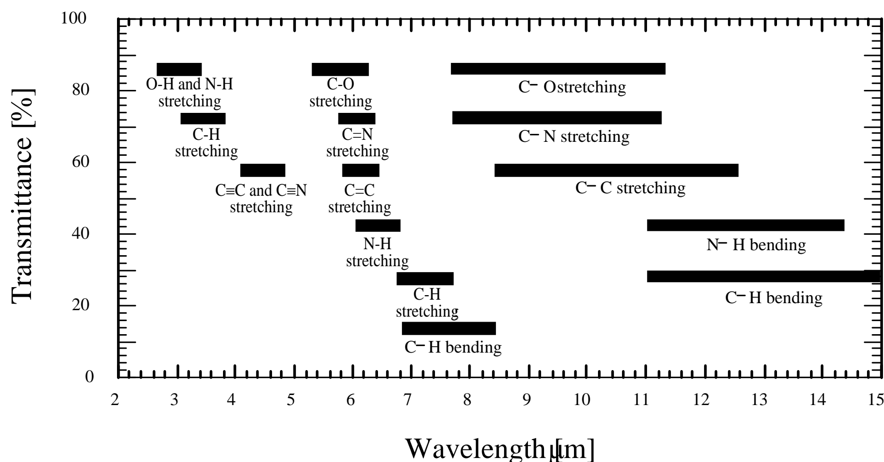


Figure 3. Simple correlations of various group molecular vibrations. [From Phil. Trans. R. Soc. A 359, pp. 635 - 644 (2001)].

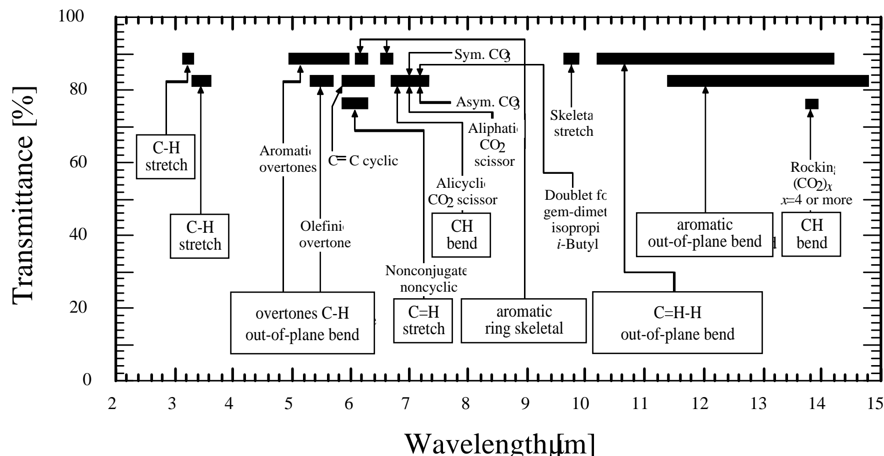


Figure 4. Characteristic hydrocarbon IR absorption. [From Phil. Trans. R. Soc. A 359, pp. 635 - 644 (2001)].

Photons in the IR spectral region have much less energy than photons in the visible and ultraviolet. They can excite vibrations in molecules and the IR interactions are characterized by transitions from one vibrational energy level to another. Figures 3

and 4 show the region of the mid-IR where specific molecules interact and absorb radiation depending on their chemical composition [1-3]. The IR spectrum can be split into four regions: 1) 4000-2500 cm^{-1} (2.5-4 μm) the absorption of single bonds to hydrogen, the high frequency is due to low mass of the hydrogen atom; 2) 2500-2000 cm^{-1} (4-5 μm) the absorption of triple bonds requires relatively high frequencies; 3) 2000-1500 cm^{-1} (5-6.6 μm) the absorption of double bonds, it is an important region because the C=O and C=C absorption occurs; and 4) 1500-400 cm^{-1} (6.6-25 μm) this region is know as "fingerprint region" because it is unique to each compound. By using characteristic absorption frequencies it is easy to identify the presence of functional groups in unknown compounds which is a significant advantage of the mid-IR biomedical field.

2.3 Mid-IR Biophotonics Delivery Systems

A basic tool of each photonics system used in biomedicine (including mid-IR biophotonics designs) is a biophotonics delivery system which involves various biomedical lasers and flexible fiber-optic delivery links for controllable transmission of the laser emission to the tissue samples to be manipulated. With regard to fiber-based biomedical laser delivery systems, recent research efforts have been concentrated on development of multifunctional and highly effective laser delivery systems working in the broad spectral range covering the ultraviolet, visible and infrared. The new developments which relate to each element of a commonly used biomedical fiber-optic delivery system will be considered below.

2.3.1 Basic Mid-IR Biophotonics Delivery System

A principal design of a conventional mid-IR biophotonics delivery system is illustrated in Figure 5. This is a fiber-optic based design that includes the following main components: a mid-IR biomedical laser with specific spectral, temporal and power characteristics; a laser-to-fiber coupling device; a flexible laser delivery fiber; and a handpiece for tissue imaging and sensing. Specific characteristics and critical parameters of each of these elements from the mid-IR biophotonics delivery system will be summarized in the next *Section*s.

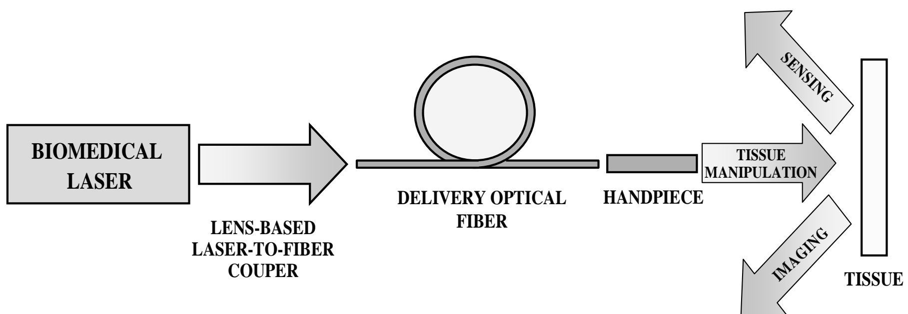


Figure 5. A conventional mid-IR biophotonics delivery system

2.3.2 Mid-IR Biomedical Lasers

In comparison with the visible and near-IR spectral ranges where the manufacturing of various laser sources is a pretty well-developed field that provides a significant number of alternatives, the mid-IR range includes a much shorter list of available biomedical lasers. Table 1 summarizes some fundamental characteristics and basic applications of the most extensively used mid-IR biomedical lasers. It includes both gas (*i.e.* CO_2 and CO) and solid-state (Ho:YAG, Er:YAG, semiconductor, and nonlinear-conversion-based) lasers. These lasers can be classified using different criteria; however, the most important of their characteristics from a biomedical application prospective are wavelength, output power/energy and CW/pulse width.

Among the mid-IR biomedical lasers presented in Table 1, the most widely studied and exploited in the field of mid-IR biophotonics are the following laser sources.

1. CO_2 laser

The CO_2 laser is the first used and the most recognized mid-IR biomedical laser named often as the “surgical laser”. This laser emits continuous wave (CW) or pulsed mid-IR emission at 10.6 μm that can be delivered precisely to a local tissue area using flexible fiber-optic systems (see *Section* 2.3.4). Because of its absorption-dominant properties and high water absorption associated with minimal penetration depth within tissue that provides its cutting and vaporizing potential, the CO_2 mid-IR laser has various biomedical applications including microscale surgical procedures, minimally invasive tumor removal, tissue ablation and reshaping. The unique features of this laser in biomedical applications ensure substantial advantages such as minimally invasive/no-touch technique, pain and swelling reduction, control of laser-tissue infection, minimal surgical bleeding, sterilization of the impact tissue site, and precise tissue manipulation.

2. Solid-state mid-IR biomedical lasers

This is a relatively large group of mid-IR biomedical lasers that consists of various types of lasers such as: widely used in biomedicine Er:YAG (2.94 μm) and Ho:YAG (2.1 μm) lasers; semiconductor and quantum cascade lasers; and laser sources based on nonlinear frequency conversion including either optical parametric oscillator (OPO) or difference frequency generation (DFG). These lasers provide mid-IR emission at both fixed and tunable wavelengths within a broad range in the interval 2-30 μm. They can also produce either CW or pulsed laser emission with various pulse widths. Solid-state mid-IR lasers are used traditionally in well established biomedical procedures such as precise tissue ablation, microscale surgical procedures, ophthalmic surgery, dentistry, as well as in IR biomedical spectroscopy, biosensors and breath analysis.

3. Free-electron laser (FEL)

Free-electron lasers involve major facilities. They ensure the free combination of wavelength, pulse width and energy. Some FELs produce a rapid sequence of short (pico-, nano-, and micro-second), broadband tunable (2-12 μm), intensive mid-IR laser emission that is used for various biomedical applications including ophthalmic microsurgery, tissue ablation, skin resurfacing, and dentistry.

Table 1. Basic mid-IR biomedical laser sources including some of their fundamental features and biomedical applications

Mid-IR biomedical laser	Wavelength [μm]	CW/ Pulsed	Biomedical applications	References
Ho:YAG (holmium yttrium-aluminium-garnet)	2.1	Pulsed Normal (free-running) mode, 200 μs	Kidney stone fragmentation Treatment of enlarged prostates General surgical procedures Coronary angioplasty	34,36,53
Er:YAG (erbium YAG) and Er:YSGGG (2.79 μm)	2.94	Pulsed Normal mode, 200 μs Q-switched, 50 ns	Tissue ablation Skin resurfacing Corneal surgery General surgical procedures Orthopedics Dentistry	16,21,24, 27, 28,33,37, 38
FEL (free-electron laser)	Fixed or tunable (2–12 μm)	Pulsed pico-, nano- and micro-second pulse width	Ophthalmic microsurgery Tissue ablation Dentistry Skin resurfacing General surgical procedures	20,23,32
Tunable quantum cascade lasers	Tunable (3–24 μm)	Pulsed	Infrared biomedical spectroscopy Infrared biosensors Breath and gas analysis Early detection of colon cancer and ulcers	50,54
Semiconductor lasers (lead-salt, antimonide laser diodes)	Fixed or tunable (3–30 μm)	Pulsed	Infrared biomedical spectroscopy Infrared biosensors Breath and gas analysis	50,54
Tunable OPO/FFG lasers (optical parametric oscillator/difference frequency generation)	Fixed or tunable (1–10 μm)	Pulsed	Infrared biomedical spectroscopy Infrared biosensors Breath and gas analysis	50,54,55
Fiber-optic lasers	Fixed or tunable (2–6.5 μm)	Pulsed	Infrared biomedical spectroscopy Infrared biosensors Breath and gas analysis	25,27,31
CO_2 (carbon dioxide) and CO (carbon monoxide, tunable 5-8 μm)	10.6	CW Pulsed	Microscale surgery (laser scalpel) Soft and hard tissue ablation Tumour removal Skin resurfacing Dentistry	21,24,28, 38,39

2.3.3 Mid-IR Incoherent Light Sources

Although for some biomedical applications mid-IR incoherent light sources are considered as an alternative approach, the field of mid-IR incoherent light sources is

not so well-developed as that of the mid-IR coherent (laser) biomedical sources. Practical biomedical applications so far include the use of various IR emitting lamps (for instance, tungsten-halogen as an IR coagulator [26]) or ceramic light sources.

2.3.4 Mid-IR Bio-medical Delivery Fibers

As a basic element in bio-photonics delivery systems, biomedical optical fibers (OFs) are now commonly used in various modern bio-photonics systems such as minimally invasive techniques, laser therapy, *in situ* laser diagnostics and optical imaging [44-49]. They possess a number of unique features such as low attenuation losses, micro-scale transverse waveguide dimensions, long interaction lengths, flexibility, immunity to external interferences and low cost. These advanced features make biomedical OFs an ideal medium for effective laser delivery and sensor design. With regard to fiber-based laser delivery systems used in biomedicine, the recent research efforts are concentrated on development of multifunctional and highly effective laser delivery systems working in a broad spectral range covering the ultraviolet, visible and infrared.

Mid-IR laser delivery OFs including biomedical OFs have been intensively investigated both theoretically and experimentally for more than three decades because of their promising applications in industrial and biomedical laser delivery systems. Research efforts are focused on producing low-loss IR fibers that provide an efficient broadband delivery of powerful mid-IR laser radiation including Er:YAG, free-electron and CO_2 lasers. However, mid-IR delivery OFs have some specific features and limitations that can be summarized as follows.

1. In comparison with the attenuation losses of standard visible and near-IR fibers (normally these are far below 1 dB/km), typical attenuation losses of IR fibers are a few orders of magnitude larger (about 10^2-10^4 dB/km).

2. The high attenuation losses impose a limitation on the maximum IR fiber length that can be used and in practice this limits construction to basically short-fiber-length laser delivery systems (typically about 1-10 m).

3. Mainly due to technological problems, the conventionally used solid-core and hollow IR fibers have large core diameters (usually 100-1000 μm) and strongly multimode waveguide core structures.

There are basically two types of mid-IR biomedical OFs that have been studied and applied in practical laser delivery systems so far, solid-core OFs and hollow waveguides.

The solid-core mid-IR OFs may be divided additionally into two broad categories: 1) glass OFs including heavy metal fluoride (ZBLAN), germanate (GeO_2-PbO) and chalcogenide (As_2S_3 and AsGeTeSe) mid-IR OFs and 2) crystalline OFs including single crystal (sapphire) and polycrystalline (halide, AgBrCl) mid-IR OFs. The solid-core mid-IR fibers provide relatively low attenuation and bending losses below 1 dB/m. These OFs have been used for delivery of basic mid-IR biomedical lasers, for example, fluoride glass, sapphire or germanate OFs for Er:YAG lasers, chalcogenide glass OFs for CO lasers, and silver halide OFs for CO_2 lasers. However, solid-core mid-IR OFs have certain critical parameters basically related to their mechanical properties, for instance, they are more fragile than silica fibers. In addition, some mid-IR medical OFs have unfavourable properties. The use

of chalcogenide OFs for biomedical purposes presents a safety issue related to possible toxic effects. Fluoride OFs are hydroscopic with possible surface crystallization effects due to moisture, whereas silver halide OFs are photosensitive to visible and ultraviolet radiation.

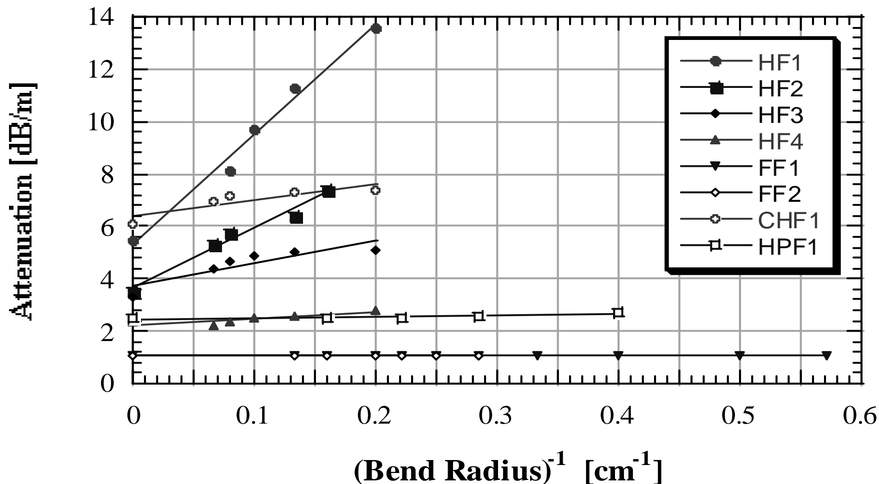


Figure 6. Measured attenuation losses *vs* bending losses for conventionally used mid-IR biomedical laser delivery fibers. HF1-HF4, hollow waveguides, FF1 and FF2, fluoride fibers, CHF1, chalcogenide fiber, HPF1, high-power germanium-doped fiber

The mid-IR hollow waveguides include three basic categories: 1) hollow glass waveguides which are prepared using fused-silica or Teflon tubing internally coated with metal (Ag or AgI) or dielectric layers, 2) hollow sapphire (refractive index <1) waveguides, and 3) hollow photonic bandgap waveguides based on the use of high-reflectance photonic bandgap mirror surfaces that consist of alternating concentric layers of optical materials with different indexes of refraction. The air core of hollow waveguides in combination with the high-reflectance inner surfaces gives significant advantages over solid-core OFs in terms of high damage threshold and highly efficient delivery of powerful biomedical lasers in the mid-IR. A critical parameter in using hollow waveguide laser delivery systems is the dependence of the waveguide attenuation losses on the bending radius. Typical bending attenuation losses measured for both solid-core and hollow fiber samples from mid-IR biomedical OFs most widely utilized currently are shown in Figure 6 [49]. These samples are available from a few fiber vendors as well as from several university laboratories around the world. Solid-core OF samples include two fluoride fibers (FF1 and FF2) from various producers, a chalcogenide fiber (CGF1) and a high-power germanium-doped fiber (HPF1). Hollow fibers include four samples that are made of either Teflon (HF1) or fused silica (HF2-HF4) tubing internally coated with dielectric layers. In general, the solid-core fibers show significantly lower attenuation losses while the larger bending losses of hollow waveguides can be

explained by the increasing number of reflections when the bending radius decreases, which leads to a relatively large growth of total attenuation.

2.3.5 All-Hollow-Waveguide Mid-IR Laser Delivery System

For matching and launching the mid-IR laser emission into delivery OFs, various focusing optical elements are conventionally used. However, in some cases, especially when powerful lasers are utilized, the focusing elements introduce some limitations connected with their own losses and aberrations, alignment problems, and damage effects. Recently, an alternative all-hollow-waveguide technique for direct (without any focusing elements) launching of laser emission into OFs was demonstrated [56-58].

A principal optical arrangement of the demonstrated all-optical-waveguide mid-IR laser delivery system is shown in Figure 7. It is a simple lens-free technique for direct laser-to-waveguide coupling that includes the following basic components: a mid-IR laser source, a laser-to-waveguide coupler and an IR delivery waveguide. As a mid-IR laser source, we use a compact free-running Er:YAG laser (λ=2.94 μm) that generates single pulses with a 150-μs pulse width and energy in the range of 5-500 mJ. In our experiment, we utilize an uncoated glass tapered hollow funnel as a laser-to-fiber

A principal optical arrangement of the demonstrated all-optical-waveguide mid-IR laser delivery system is shown in Figure 7. It is a simple lens-free technique for direct laser-to-waveguide coupling that includes the following basic components: a mid-IR laser source, a laser-to-waveguide coupler and an IR delivery waveguide. As a mid-IR laser source, we use a compact free-running Er:YAG laser (λ=2.94 μm) that generates single pulses with a 150-μs pulse width and energy in the range of 5-500 mJ. In our experiment, we utilize an uncoated glass tapered hollow funnel as a laser-to-fiber coupler instead of a conventional lens-based coupler that introduces significant limitations in terms of additional losses, aberrations, damage effects and alignment problems especially in the case of powerful invisible laser pumping. The operating principle of the optical funnel is based on the grazing incidence effect. It ensures a direct launching of mid-IR laser emission into the optical delivery waveguide providing a high laser-to-fiber coupling efficiency. Moreover, because of its mode filtering properties, the optical funnel serves as a mode converter that transforms the highly multimode intensity distribution of the input laser emission (Figure 7a) into a high-quality smooth profile at the funnel output (Figure 7b). This profile can provide an effective laser-to-fiber coupling with higher damage thresholds and more homogeneous coupling that is essential for the tissue ablation processes. The delivery mid-IR fiber is a key optical element in the tissue ablation system. Here, a mid-IR-transparent hollow waveguide with a 600-μm-core diameter and a 0.6-m length is used. It is a fused-silica tubing internally coated with high reflectance dielectric layers. The hollow waveguide shows low attenuation losses in the range of 1-1.5 dB/m at λ=2.94 μm and high damage threshold for powerful laser transmission. In order to realize a regime of on-the-spot mid-IR laser delivery to a precise tissue area, we use a smart, tissue-activated solid-core fiber tip with a specially angle-shaped profile (Figure 7c). When the fiber tip is not touching the tissue area, the laser emission is backreflected at the angled tip due to total-internal-reflection. However, when the tip is placed on an absorbing tissue, it becomes

"transparent" (or tissue-activated) for laser emission because of the frustrated-total-internal-reflectance and the energy then is coupled into the absorber. Thus, an effective and safe method for mid-IR laser radiation delivery to a microscale tissue area is ensured. This system can be used for precise on-the-spot laser tissue ablation, micro-scale surgical procedures, laser therapy and hair removal.

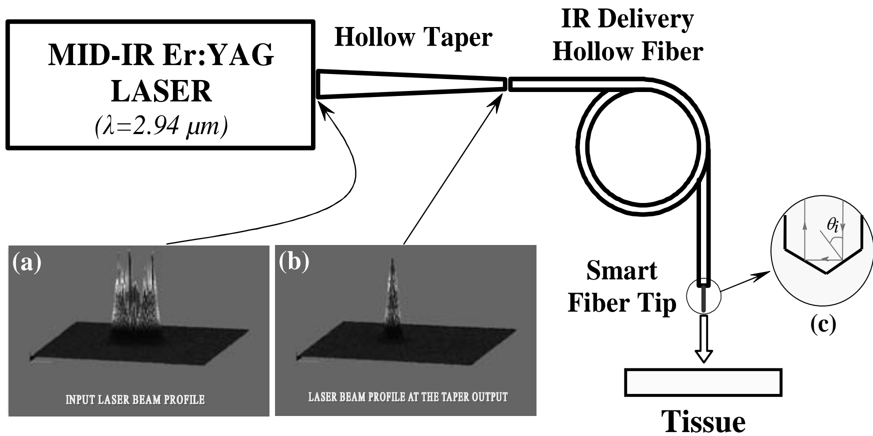


Figure 7. Experimental setup of the all-optical-waveguide delivery system for precise mid-IR laser ablation of specific brain regions in a goldfish model to simulate Parkinson's disease

3 Mid-IR Bio-photonics Applications

There are a significant number of different available biomedical applications based on the mid-IR bio-photonics delivery systems described in Section 2. In the present *Section*, we will categorise these applications into three major groups: precise laser surgery and tissue ablation, bio-imaging and biosensors, and mid-IR biomedical spectroscopy.

3.1 Mid-IR Laser Surgery and Tissue Ablation

In recent years, mid-IR bio-photonics delivery systems have become an increasingly attractive tool for highly precise and effective medical ablation of both soft and hard tissue as well as alternatives to mechanical drills and scalpels in areas of surgery and medicine [16-41]. As we described in Section 2.2, the main reasons leading to this interest are related to the water absorption dominant tissue properties and as a result, the minimum penetration depth of mid-IR laser emission in tissue.

Moreover, pulsed mid-IR tissue ablation effects cause minimal damage around the exposed area and mid-IR radiation lacks the carcinogenetic potential that has generated concern for ultraviolet laser tissue interactions. Figure 8 shows optical penetration depths in soft tissue of typical mid-IR biomedical lasers described in *Section* 2.3.1.

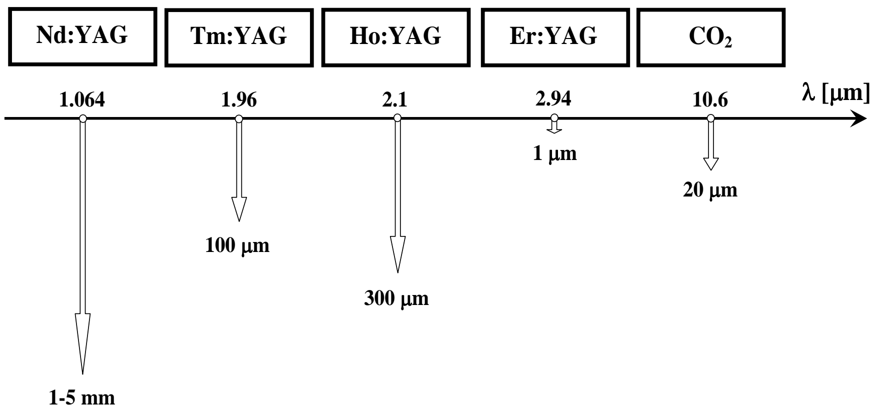


Figure 8. Penetration depth in soft tissue of typical mid-IR biomedical lasers and Nd:YAG laser for comparison

The Er:YAG laser has a minimum penetration depth of about 1 μm because its wavelength falls into the water absorption band with the absolute maximum absorption. The CO_2 laser has a relatively small penetration depth and it is one of the most frequently used mid-IR tissue ablative lasers. The typical penetration depth data are given in Figure 8 compared with the near-IR Nd:YAG laser that has a few orders of magnitude higher penetration depth in tissue.

There are four basic photoablation mechanisms that are conventionally considered in mid-IR laser tissue ablation: 1) photochemical processes including bio-stimulation and photodynamic therapy that describe laser stimulation effects on biochemical and molecular interactions, and the therapeutic use of lasers to induce tissue reactions for treatment of pathological conditions; 2) photothermal processes defined in clinical use as photoablation or tissue removal by vaporization and superheating of tissue fluids; 3) photomechanical processes leading to breaking apart of tissue structures and tissue removal with laser-induced shock-wave generation; and 4) photoelectrical processes including tissue removal by the formation of electrically charged ions and particles.

In the ideal case, photoablation should remove tissue with minimum thermal damage and dissections. This can be achieved using mid-IR lasers having optimum combination between the laser wavelength, pulse width and fluence. However, photoablative and surgical requirements are often contradictory. For example, a high ablation rate is required in dental ablation and surgery, while it must be minimal to reshape the human cornea. In other cases, cutting through vascularized tissue (*e.g.* brain surgery) requires some amount of surface coagulation (thermal damage), while for wound healing there is no need of thermal damage. Lasers with higher pulse repetition rates are used for reducing the surface roughness while, at the same time these types of lasers increase the overall thermal load on the tissue target.

The most commonly used mid-IR lasers for precise tissue photoablation and surgery are presented in Table 1 and Figure 8. Basically, these lasers have been used in various photoablation and surgical procedures including ophthalmic microsurgery, corneal ablation, soft and hard tissue microablation, tumor removal, skin resurfacing, and hair removal. Some of the most significant advantages of these

medical procedures are: minimally invasive /no-touch technique, minimal pain and swelling reported after surgery, sterilization of the laser cut, reducing the risk of blood-born disease transmission, less postoperative infection, less damage to adjacent normal tissue, minimal surgical bleeding, better access to difficult to access tissue locations (*i.e.* oral cavity), and precise tissue manipulation.

3.2 Mid-IR Bioimaging

Mid-IR bioimaging is an advanced biophotonics technique based on thermal tissue imaging and monitoring that is capable of providing non-contact, *in vivo* diagnostic information. Being a passive technique (*i.e.* without external sources of radiation), infrared bioimaging is a non-invasive, intrinsically harmless and real time method

Figure 9. Typical thermal images obtained using mid-IR laser irradiation (see Figure 7) of chicken skin after 10 laser pulses when the OF tip is above the skin surface (**a**) and after a single laser shot when the smart tip is on the surface (**b**)

for monitoring laser beam propagation in tissue [13,14]. For example, Figure 9 shows typical thermal images obtained using mid-IR Er-YAG laser irradiation (see the experimental setup shown in Figure 7) of chicken skin after 10 laser pulses when the smart, tissue-activated OF tip is above the skin surface (Figure 9a) and after a

single laser shot when the OF tip is on the skin surface (Figure 9b). As can be seen, if the smart OF tip doesn't touch the absorbing tissue area there is no tissue illumination even after 10 laser pulses. However, when the OF tip is in contact with the tissue, just a single laser shot is enough for the maximum laser energy to be transferred to the tissue area. Moreover, according to the thermal images obtained, the IR laser emission is delivered to a precise tissue area. Typically the irradiated area observed is localized in the range between the OF tip core diameter and 1 mm. In this study, a thermal image processor system (Bales Scientific) that includes a high resolution thermal camera with a 0.05°C minimum detectable temperature change, a spatial resolution of 600 optical lines at 30° field of view, a focal range from 10 cm to infinity, and a spectral response over the ranges 3-5 and 8-12 μm is used [56].

In recent years research efforts in the field of mid-IR thermal imaging have been focused on the development of mid-IR thermal cameras and techniques possessing high spatial (micron and sub-micron range) and temperature resolution. Another attractive area in this field is the development of special mid-IR-waveguide based bundles for transmission of infrared thermal images.

3.3 Mid-IR Spectroscopy and Biosensors in Biomedicine

Mid-infrared spectroscopy in medicine is currently in its infancy and like many other applications it is limited by the development of all of the previously discussed components, *i.e.* sources, detectors, endoscopic delivery and transmission to and inside the human body. Furthermore, it is also limited by a lack of understanding and the reluctance to use mid-infrared technology by the medical profession. Despite these problems, we still believe there are wide-ranging applications in medicine where mid-infrared technology can contribute less invasive, less expensive diagnostics and therapy than is currently available. Possible medical uses, some of which have already received cursory research, will most likely involve spectroscopy of solids, liquids and gaseous waste products to look for indications of chemicals related to disease. Other uses of spectroscopy are to sense hazards in food, water and air from either natural emissions, pollutants or terrorist injected bio or nuclear products. All these are of potential bio-medical interest and are potentially determined by biomedical sensing or imaging. In this *Section*, we concentrate on the more developed interest in trace gas detection of gaseous indicators in breath and the non-invasive sensing of glucose in blood.

3.3.1 Non-invasive Blood Glucose Monitoring

We will first discuss the long sought capability to detect glucose non-invasively in the blood and to replace the relatively expensive and unpleasant meters used by many millions of people with diabetes that require a small drop of blood to determine the glucose level in blood. Typical fasting levels of glucose in normal patients are 80 - 140 milligrams per deciliter (mg/dl). Fasting levels above this indicate diabetes and the need for medicine or insulin to reduce the glucose level to normal. Currently almost all invasive meters use calibrated strips which react to the glucose in blood to generate a measurement. These strips cost on the order of

$0.50/strip, are not reusable and become more costly as the number of tests/day increases. Meters cost on the order of $50 each, but are often discounted because of the revenue derived from the strips. A non-invasive meter would probably have to be priced somewhere between $500 - $2000 to compete with current meters depending on the expected life of the meter and the value of not requiring the finger stick. In addition, it has been estimated that the required accuracy should be somewhere near +/-10 mg/dl or better to be useful. As yet no non-invasive device has demonstrated accuracy better than about a standard deviation of 27 mg/dl. [4]. Because of the high cost of mid-infrared detectors, many workers have worked close to the near infrared or in the 1-2 μm or 1-5 μm region where it is difficult to eliminate competing components in blood. Work in the 7-11 μm region has less competing spectra, but here detector cost and performance can be a problem. None the less, several studies suggest that it is possible to make sufficiently accurate measurements [5,6]. If these theoretical predictions are correct, then the proper system must still be designed that is within the cost range and competitive with the current meters and strips. A break-through in source, detector or system technology would be helpful in this respect.

3.3.2 Breath Analysis for Medical Applications

Mid-infrared trace gas detection, *i.e.* using mid-infrared spectroscopy, may have the potential to compete as a system with tandem mass spectroscopy for analysis of the 500 or so compounds emitted in breath [7]. The metabolic pathways that produce these compounds are still not all known. Some of them relate to diseases such as diabetes, cirrhosis and renal disease. Others relate to drugs, diet adherence and radiation exposure and others result from exposure to occupational hazards and smoking.

Collection of breath samples is important for quantization of breath compounds since the proportion of alveolar air to dead space air (the amount of air in the mouth and trachea where no compounds originate) will vary from breath to breath. The standard is to sense CO_2 and collect only the latter half of the breath that is expired after the first CO_2 is sensed. Many of these compounds require the ability to sense parts per billion or parts per trillion of the fraction of CO_2. Detection of these parts by optical means adds the ability to sense in real time thus potentially measuring expiration rates.

3.3.2.1 Diagnostic Use of Breath Components

Medical science has known for some time that various breath odours could indicate particular diseases. Quantifying breath is non-invasive and is less costly than blood analysis. Several uses in disease detection are mentioned below:

1. Diabetes - in Section 3.3.1 we mentioned a possible use of mid-infrared spectroscopy to measure glucose in blood. Already many millions of diabetics use the current invasive meters to measure their blood glucose. The readings that invasive or non-invasive meters give is only an indication of their current glucose level and gives them no information on what to do to move toward lower readings and better health. Combined, however, with readings from breath analysis, the readings of glucose and breath acetone together can suggest treatment [8]. For

example, a) when glucose is high and breath acetone is normal (10-50 nmol/L) the diagnosis is overeating and the treatment is less food; b) when glucose is high and breath acetone is high (>50 nmol/L) the diagnosis is insulin insufficiency and treatment is to give more insulin; c) when glucose is normal and breath acetone is high (>50nmol /L) the diagnosis is carbohydrate insufficiency and the treatment is to continue the diet if weight loss is the goal or reapportion diet by adding more carbohydrate; d) when both glucose and breath acetone are normal the diagnosis is good metabolic control and treatment is to continue current method. Breath acetone can be a more sensitive indicator of poor control of diabetes than blood glucose, but millions of diabetics today have no such supporting device to assist them in managing their diabetes.

2. Weight reduction - as indicated in c) of the example above, normal glucose and high breath acetone indicate that an obese subject is adhering to a diet. Breath acetone doubles by the second day of a low carbohydrate diet and by the third or fourth day has increased five fold. A breath acetone of 500 nmol/L is associated with a weight reduction of about one-half pound per week. When breath acetone has stabilized, one can easily determine when the subject has eaten a high carbohydrate meal; for instance, eating a doughnut will halve the breath acetone in 3 h. Exercise increases the concentration of acetone and is a good measure of the effectiveness of a subject's exercise program.

3. Renal disease - dimethyl- and trimethylamine have been found in high concentrations in the breath of uremic patients. After renal dialysis these amines return to normal concentrations.

4. Hepatic disease - dimethyl sulfide, mercaptans and fatty acids are found in the breath of cirrhotic patients [9]. Dimethyl sulfide has been found to be significantly increased in cirrhotic patients as compared with normal subjects. Ammonia is normally present in human breath, but is increased in hepatic disease

3.3.2.2 Real-time Recording of Ethane via Mid-infrared Spectroscopy

Ethane (C_2H_6) has been measured in real time to an accuracy of less than 100 parts per trillion in human breath using mid infrared laser spectroscopy [10]. Ethane is the most stable of the volatile hydrocarbons and is a direct measure of oxidative stress from reactive oxygen species. Ethane is a marker of lipid peroxidation, radiation damage and the effects of smoking. Some of the first real time measures have been done by the group led by Murtz [11,12]. To make these measurements the breath collected must be carefully pre-concentrated and injected into a resonant cavity absorption cell through which a CO laser is passed many times. The data are rapidly recorded, processed and statistically analyzed and are able to measure to 500 ppt in less than 800 ms.

3.3.3 Summary of Mid-IR Spectroscopy and Biosensors

The potential of mid-IR spectroscopy and biosensors for medical detection of diseases and for the detection of toxic bio-products in food and food processing have not been evaluated because the work so far has only been exploratory. Real time measurements of breath using proper breath sampling techniques must become routinely used. Identification of more of the biomarker molecules must be traced

back to their chemical origin within the human body. Once high sensitivity instrumentation is developed, a large number of studies will be needed to determine the relevant biomarkers of disease and to see if important signs indicating major diseases (*e.g.* cancer) can be determined at early stages with this non-invasive technology.

4 Conclusion

Uses of mid-IR biophotonics technology are poised to make important contributions to medicine. Mid-IR biophotonics delivery systems including biomedical lasers and delivery waveguides have become more mature and great progress in mid-IR spectroscopy has been made. The next decade should realize some important mid-IR medical devices exploring some of the fundamental mid-IR biophotonics principles and features summarized in this chapter.

References

1. B. Smith, *Fundamentals of Fourier Transform Infrared Spectroscopy* (CRC Press, Boca Raton, 1995).
2. C. Faust, *Modern Chemical Techniques* (RSC, London, 1992).
3. R. Waynant, I. Ilev and I. Gannot, "Mid-infrared laser applications in medicine and biology", Philosophical Transactions of The Royal Society London, v. A 359, pp. 635-644, 2001.
4. C. Malchoff, J. Landau, K. Shouki, and J. Buchert, "A novel noninvasive blood glucose monitor", Diabetes Care, v. 25, pp, 2268-2275, 2002.
5. Y. Kim, S. Hahn, G. Yoon, "Determination of glucose in whole blood samples by mid-infrared spectroscopy", Applied Optics, v. 42, pp. 745-749, 2003.
6. P. Zheng, C. Kramer, C. Barnes, J. Braig and B. Sterling, "Noninvasive glucose determination by oscillating thermal gradient spectrometry", Diabetes Technology and Therapeutics, v. 2, pp. 17-25, 2000.
7. D. Manolis, "The diagnostic potential of breath analysis", Clinical Chemistry, v. 29, pp. 5-15, 1983.
8. O. Crafford, et. al. "Acetone in breath and blood", Trans. Am. Clin. Climatol. Assoc. v. 88, pp. 128-139, 1977.
9. F. Kaji, M. Hisamura, N. Saito and M. Murao, "Evaluation of volatile sulfide compounds in the expired alveolar gas in patients with liver cirrhosis," Clin. Chem. Acta, v. 85, pp. 279-284, 1978.
10. T. Risby, and S. Schnert,"Clinical applications of breath biomarkers of oxidative stress status", Free Radic. Bio Med., v. 27, pp. 1182-1192, 1999.
11. G. Basum, H. Dahnke, D. Halmer, P. Hering and M. Murtz, "Online recording of ethane traces in human breath via infrared laser spectroscopy", J. Appl. Physiol., v. 95, pp. 2583-2590, 2003.
12. S. Stry, P. Hering and M. Murtz, "Portable difference-frequency laser-based cavity leak-out spectrometer for trace gas analysis" Applied Physics B., v. 75, pp. 297-303, 2002.
13. H. Fabian, P. Lasch, M. Boese and W. Haensch "Mid-IR microspectroscopic imaging of breast tumor tissue *Sections*", Biopolymers, v. 67, pp. 354-357, 2002.
14. R. Mendelsohn, H. Chen, M. Rerek and D. Moore, "Infrared microspectroscopic imaging maps the spatial distribution of exogenous molecules in skin", Journal of

Biomedical Optics, v. 8, pp. 185-190, 2003.
15. J. Viator, B. Choi, G. Peavy, S. Kimel and J. Nelson, "Spectra from 2.5-15 microm of tissue phantom materials, optical clearing agents and ex vivo human skin: implications for depth profiling of human skin", Phys Med Biol., v. 48, pp. N15-24, 2003.
16. M. Pierce, S. Jackson, M. Dickinson, T. King, and P. Stoan, "Laser-tissue interaction with a continuous wave 3-mm fiber laser: Preliminary studies with soft tissue", Lasers in Surgery and Medicine, v. 26, pp. 491-495, 2000.
17. E. Ross, F Sajben, J.Hsia, D.Barnetta, C. Miller, J. McKinlay,Nonablative "Selective Dermal Heating Skin Remodeling with a Mid-infrared Laser and Contact Cooling Combination", Lasers in Surgery and Medicine, v.26, pp. 186-195, 2000.
18. K. Nahen, A. Vogel, "Investigations on acoustic On Line Monitoring of IR Laser Ablation of Burned Skin",Lasers in Surgery and Medicine, V.25, pp. 69-78 1999.
19. M.Ostertag, J. McKinley, L. Reinisch, D. Harris, N. Tolk, "Laser Ablation as a Function of the Primary Absorber in Dentin", Lasers in Surgery and Medicine, V.21, pp.384-394, 1997,
20. J. Ovelmen-Levitt, K. Straub, S, Hauger,E. Szarmes, J. Madey, R. Pearlstein, B. Nashold Jr, Brain "Ablation in the Rat Cerebral Cortex Using a Tunable-Free Electron Laser", Lasers in surgery and Medicine, v. 33, pp. 33-61, 2003.
21. A. Rosa, A. Anupama, V. Sarma, C. Le, R. Jones, D. Fried, "Peripheral Thermal and Mechanical Damage to Dentin With Microsecond and Sub-microsecond 9.6 µ.m, 2.79 µ.m, and 0.355 µ.m Laser Pulses" Lasers in Surgery and Medicine, v. 35, pp. 214-238, 2004
22. M. Staninec, J. Xie, C. Le, D. Freid, "Influence of an Optically Thick Water Layer on the Bond Strength of Composite Resin to Dental enamel After IR Laser Ablation", Lasers in Surgery and Medicine, v. 33, pp. 264-269, 2003.
23. N. Spector, L. Reinisch, J. Spector, D. Ellis, "Free-Electron Laser and Heat-Conducting Templates: A Study of Reducing Cutaneous Lateral Thermal-Damage",Lasers in Surgery and Medicine, v. 30, pp. 117-122, 2002
24. D. Freid, J. Ragadio, A. Champion, "Residual Heat Deposition in Dental Enamel During IR Laser Ablation at 2.79, 294, 9.6 and 10.6 µm", Lasers in Surgery and Medicine, v. 29, pp. 221-229, 2001
25. M. Pierce, S. Jackson, M. Dickinson, T. King, "Laser Tissue Interaction With a High-Power 2-µm Fiber Laser: Preliminary Studies With Soft Tissue", Lasers in Surgery and Medicine, v. 25, pp. 407-413, 1999
26. H. Kubota, A. Furuse, M. Takeshita, Y. Kotsuka, S. Takamoto, "Atrial Ablation With an IRK-51 Infrared Coagulator", Ann Thorac Surg, v. 66, pp. 95-100, 1998
27. M. Pierce, S. Jackson, M. Dickinson, T. King, P. Sloan, "Laser Tissue Interaction With a Continuous Wave 3-µm Fiber Laser: Preliminary Studies With Soft Tissue" Lasers in Surgery and Medicine, v. 26, pp. 491-495, 2000
28. G. Peavy, L. Reinsch, J. Payne, V. Venugopalan, "Comparison of Cortical Bone Ablations by Using Infrared Laser Wavelengths 2.9 to 9.2 µm", Lasers in Surgery and Medicine, v. 26, pp. 421-434, 1999
29. K. Tanhuanpaa, K. Cheng, K. Antonen, J. Virtanen, P. Someharju, "Characteristics Pyene Phospholipid/y-Cyclodextrin Complex", Biophysical Journal, v. 81, pp.1501-1510, 2001
30. S. Sato, M. Ogura, M. Ishihara, S. Kawauchi, T. Arai, T.Matsui, A. Kurita, M. Obara, M. Kikuchi, H.Ashida, "Nanosecond, High-Intensity Pulsed Laser Ablation of Mycardium Tissue at the Ultraviolet, Visible and Near-Infrared Wavele: In Vitro Study, Lasers in Surgery and Medicine, "v. 29, pp. 464-473, 2001
31. M. Pierce, S,Jackson, M. Dickinson, T. King. P. Sloan, "Laser Tissue InteractionWith a Continuous Wave 3-µm Fiber Laser: Preliminary Studies With Soft Tissue", Lasers in Surgery and Medicine, v.26, pp. 491-495, 2000
32. M. Heya, S. Sano, N. Takagi, Y. Fukami,K. K. Awazu, "Wavelength and Average

Power Density Dependency of the Surface Modification of Root Dentin Using an MIR-FEL", Lasers in Surgery and Medicine, v.32, pp. 349-358, 2003

33. H. Hoffman, W. Telffair, "Photospallation: A New Theory and Mechanism for Mid-Infrared Corneal Ablations", Journal of Refractive Surgery, v. 16, 2000
34. O. Topaz, M. Mclvor, G. Stone, M. Krucoff, , E. Perin, A. A. Fosschi, J. Sutton, R. Nair, E. deMarchena, and the Holmium:YAG Laser Multicenter Investigators, "Acute Results, Complications, and Effect of Lesion Characteristics on Outcome With the Solid State, Pulsed-Wave, Mid-Infrared Laser Angioplasty System: Final Multicenter Registry Report", Lasers in Surgery and Medicine, v. 22, pp. 228-239, 1998
35. D. Harris, M. Yessik, "therapeutic Ratio Quantifies Laser Antisepsis Ablation of Porphyromonas gingivalis With Dental Lasers", Lasers in Surgery and Medicine, v> 35, pp. 206-213, 2004
36. W. Cecchetti, F. Zattoni, F. Nigro, A.Tasca, "Plasma Bubble Formation Induced by Holmium Laser: An in Vitro Study", Urology v. 63, p3,2004
37. C. Channey, Y. Yang, N. Fried, "Hybrid Germanium/ Silica Optical Fibers for Endoscopic Delivery of Erbium: YAG Laser Radiation" Lasers in Surgery and Medicine, v. 34, pp. 5-11, 2004
38. A. Dela Rosa, A. Sarma, C. Le, R. Jones, D. Fried, "Peripheral Thermal and Mechanical Damage to Dental With Microsecond 9.6 μm, 2.79 μm, and 0.355μm, Laser Pulses", v. 35, pp. 214-228 2004
39. A. Dayan, A. Goren, I. Gannot, "Theoretical and Experimental Investigation of the Thermal Effects Within Body Cavities During Transendoscopical CO2 Laser-Based Surgery", Lasers in Surgery and Medicine, v. 35, pp. 18-27, 2004
40. D. Simhon, T. Brosh, M. Halpern, A. Ravid, T. Vasilyev, N. Kariv, A. Katzir, Z., Nevo, "Closure of Skin incisions in Rabbits by Laser Soldering: I: Wound Healing vPattern", Lasers in Surgery and Medicine, v. 35, pp. 1-11, 2004
41. L. Santana-Blank, H. Reyes, E. Rodriguez-Santana, "Microdensitometry of T2-Weighted Magnetic Resonance (MR) Images From Patients With Advanced Neoplasias in a Phase 1 Clinical Trial of an Infrared Pulsed Laser Device (IPLD)", Lasers in Surgery and Medicine, v> 34, pp. 398-406, 2004
42. V. Zharoy, S. Ferguson, J. Eidt, P. Howard, "Infrared Imaging of Subcutaneous Veins", Lasers in Surgery and Medicine, V. 34, pp.56-61, 2004.
43. I. Ishikawa, K. Sasaki, A. Aoki, H. Watanabe, "Effects of Er:YAG Laser on Periodontal Therapy", Journal of the International Academy of Periodontology, v. 5/1, pp. 23-28, 2002.
44. J. Sanghera and I. Aggarwal, eds., *Infrared Fiber Optics*, (CRC Press, Boca Raton, 1998).
45. J. Harrington, "A review of IR transmitting hollow waveguides", Fiber and Integrated Optics, v. 19, pp. 211-217, 2000.
46. J. Harrington, C. Rabii, and D. Gibson, "Transmission properties of hollow glass waveguides for the delivery of CO2 surgical laser power", IEEE J. Select. Topics Quantum Electron. **5**, 948-953 (1999).
47. N. Groitoru, J. Dror, and I. Gannot, "Characterization of hollow fibers for the transmission of infrared radiation", Appl. Opt. **29**, 1805-1809 (1990).
48. R. Driver, G. Leskowitz, and L. Curtiss, "The characterization of infrared transmitting optical fibers", Mat. Res. Soc. Symp. Proc. **172**, 169-175 (1990).
49. I. Ilev, R. Waynant and M. Bonaguidi, "Attenuation measurement of infrared optical fibers using a hollow-taper-based coupling method", Applied Optics, v. 39, pp. 3192-3196, 2000.
50. F. Titel, D. Richter and A. Fried, *Mid-Infrared Laser Applications in Spectroscopy*, in Solid-State Mid-Infrared Lasers Sources, I. Sorokina, K. Vodopyanov, Eds. (Springer-Verlag Berlin, 2003).
51. G. Hale and M. Querry, "Optical constants of water in the 200-nm to 200-micrometer

wavelength region**",** Applied Optics, v. 12, pp. 555-563, 1973.

52. D. Wieliczka, S. Weng and M. Querry**, "**Wedge shaped cell for highly absorbent liquids: infrared optical constants of water**,** Applied Optics, v. 28, pp. 1714-1719, 1989.
53. O. Topaz, "Holmium laser angioplasty", Semin Intervent Cardiol., v. 1, pp. 149-161, 1996.
54. R. Curl and F. Tittel, "Tunable infrared laser spectroscopy", Annu. Rep. Prog. Chem., Sec, C, v. 98, pp. 219-272, 2002.
55. D. Lancaster, A. Fried, B. Wert, B. Henry and F. Tittel, "Difference-Frequency-Based Tunable Absorption Spectrometer for Detection of Atmospheric Formaldehyde", Applied Optics, v. 39, pp. 4436-4443, 2000.
56. I. Ilev, R. Waynant, K. Byrnes and J. Anders, "On-off laser delivery into a precise tissue area using smart tissue-activated fiber probes", IEEE Journal of Selected Topics of Quantum Electronics, v. 9, pp. 331-336, 2003.
57. I. Ilev and R. Waynant, "Uncoated Hollow taper as a simple optical funnel for laser delivery", Review of Scientific Instruments, v. 70, pp. 3840-3843, 1999.
58. I. Ilev and R. Waynant, "Grazing-incidence-based hollow taper for infrared laser-to-fiber coupling", Applied Physics Letters, v. 74, pp. 2921-2923, 1999.

Development of Infrared Countermeasure Technology and Systems

D.H. Titterton
DSTL, Sensors Dept, Farnborough, Hants. UK

1 Introduction

The requirement for the development and perfection of infrared countermeasures is a direct consequence of extensive proliferation of infrared (IR) sensors that may be used on the battlefield, or in guerrilla warfare. These sensors and systems have been shown to be very effective and have provided additional military capability to regular, as well as irregular, forces. This fact is often exemplified by the effectiveness of heat-seeking missiles, which have been responsible for the majority of aircraft losses since their introduction into service during the 1960s (Figure 1). Some statistics suggest that heat-seeking missiles have been responsible for more than 80% of all combat aircraft losses over the last 40 years [1].

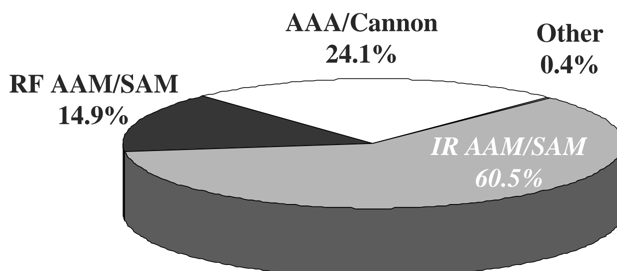


Figure 1. A summary of the documented aircraft losses over a 35-year period to the 1990s

The rise of world-wide terrorism and the proliferation of these simple weapons has only served to reinforce the need for more robust and effective countermeasures. Moreover, the requirement for effective countermeasure systems has promoted the research into the science underpinning the understanding of the behaviour of IR photons in the atmosphere, their interaction with materials and the efficient generation of IR photons. The effectiveness of countermeasure techniques has been boosted by many technological developments, but in particular by the design and demonstration of compact laser sources emitting in the IR part of the spectrum.

An effective countermeasure system may be a reactive system, which responds to the detected, or known, presence of a threat and in this case would undertake some action such as the emission of IR photons to disrupt the performance of the threat system or weapon. The countermeasure system may also be used in a pre-emptive (or pro-active) mode so that the threat weapon system is unable to acquire a

suitable target, although the target may have been detected by the threat system. In general, countermeasure systems are considered to be soft-kill weapons as interference take place with some system or sub-system in the target, causing it to malfunction and miss its intended victim, rather than the use of bullets, high explosive and/or fragments from a warhead to destroy a target.

The principle underpinning any successful IR countermeasure system is the application of any technique that degrades the notional performance and normal mode of operation of a sensors or system. A number of effective approaches have been demonstrated and these may be considered as falling into one, or more, of the following broad categories:

- Denial – in this case the sensor may be denied information about the thermal scene by the deliberate use of an obscurant or smoke.
- Deception – the use of deception techniques is often associated with the injection of false information into a sensor, such as a seeker, so that a target-tracking algorithm is confused. Modest power IR sources are required for this application, which are modulated to provide a pulsed waveform An alternative approach concerns having a "reactive surface" so that the emission from an object matches the background like a chameleon.
- Dazzle – this defeat mechanism is similar to denial, but this is normally associated with the use of an active system to create a very bright object that cause scatter of the incident photons in the optical train of a sensor. This technique can be invoked with modest power lasers or other bright sources.
- Damage – this form of defeat mechanism requires an active countermeasure device designed to deposit photon energy rapidly into a vulnerable component and offers the potential of a very robust defeat mechanism, This damage route is normally associated with attack of optical systems with IR pulse lasers.
- Destruction – this form of countermeasure is associated with the American "Star Wars" project where lasers were proposed for melting components or structures of platforms such as missiles or aircraft. This type of directed energy weapon requires high-performance and very powerful laser devices.

This chapter will consider each of the five defeat mechanisms outlined above in some detail, particularly with regards to fundamental principles of the operation of the sensor and relevant countermeasure techniques. Active countermeasure techniques will be considered in this chapter, covering both so-called on-board systems as well as expendables. There will be a particular emphasis on laser-based systems *i.e.* on-board systems, but mention will be made of incoherent sources and the development of effective systems for platform protection. Consideration will be given to the fundamentals of propagation, development of modern IR countermeasure systems, their sources and future requirements.

2 Historical Development

Optical countermeasures have been used for several millennia to provide a defensive capability. The initial approaches concerned the use of focussed sunlight to either destroy objects or dazzle an opponent's soldiers. More recent applications have concentrated on the use of techniques to disrupt the operation of a sensor, or to prevent observation of a scene or normal operation of a target tracking system causing a weapon to miss its intended target. Clearly, the development of the laser has provided a stimulus for a number of soft-kill weapon systems used to enhance platform survivability and anti-air missile applications.

Optical countermeasure techniques are believed to have started several millennia ago at the battle of Syracuse, in 212 BC. During this battle Archimedes appears to have been the pioneer in the application of this technique by using an array of plane mirrors to focus sunlight on to a distant target, in an attempt to burn the sails of the ships of the attacking Roman fleet. Each mirror was "silvered" on both sides and had a hole in the middle to act as a sighting device, so when the sun was imaged on the hole the sun was directed at the object seen in the hole. "Optimum" use of the position of sun has also featured, or played a decisive role, in a number of battles, such as the battles at Hastings (UK 1066) and Agincourt (France 1415).

Smoke screens have been exploited for centuries as a visible-band obscurant, offering a simple technique for disrupting a sensor's function, for example concealing manoeuvres and deployment of weapons: their use continues today. The original approach involved the use of combustible materials that would generate copious amounts of small particles, which would scatter the visible light giving complete extinction of the light beams. Development of sensors operating in other optical wavebands, including the IR, has required the generation of far more sophisticated obscurants; the same physical processes of extinguishing the light beams apply to denying a sensor information. However, the modern "smokes" provide obscuration over many parts of the spectrum, including the IR.

The 1960s saw the emergence of heat-seeking missiles into operational service [2]. This was the start of a technical push for techniques to defeat the "optical" seeker system in these weapons which are still difficult to intercept with hard-kill weapons, owing to their small size and short engagement timeline.

Many of the active countermeasure projects were initiated in the 1960s in response to the introduction of the heat-seeking missile into operational service. These systems were required to defeat this relatively simple, but effective, IR-homing missile technology. On-board systems, known as jammers as well as expendables were developed to confuse the target-tracking system in the missile's seeker, *i.e.* deception of the missile's sensor. Other pragmatic techniques were also devised including the extension of the jet pipes in fighter aircraft to mitigate the effects of an interception by a heat-seeking missile. This has been a classic "cat and mouse" activity between the measures and the countermeasure, which continues to this day, as shown in Figure 2. An important operational consideration is the flexibility and adaptability of the new technology, so that the older weapon systems can be defeated, as well as the new ones.

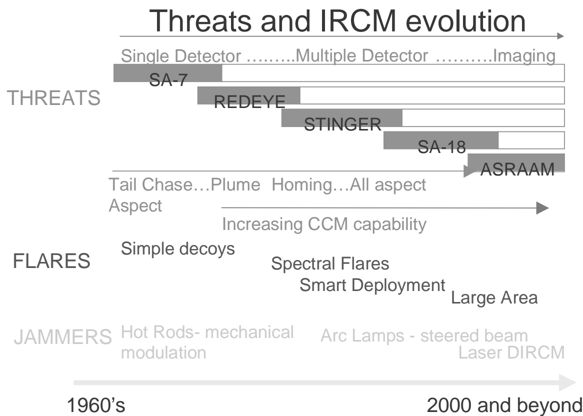


Figure 2. A summary of threat development

Very high-power lasers were first demonstrated in the 1970s, only a decade after the first demonstration of a laser by Theodore Maiman [3]. Early in the decade the US Air Force demonstrated a 100–kW IR device; later in the decade powers approaching 0.5 MW was recorded from mid–IR lasers [4]. Systems were developed to demonstrate the potential performance of directed energy weapons for destruction of targets, such as the US Navy Army chemical laser (NACL) being integrated with an accurate pointing system to shoot down a Tow anti-tank missile. Another spectacular demonstration was the destruction of a drone, known as Vandal, flying at supersonic speed. These were early examples of the feasibility of a damage-class laser system. Other significant experimental demonstrations around that time involved the US Air Force's Airborne Laser Laboratory (ALL), which operated with a gas-dynamic carbon dioxide laser; this system defeated a Sidewinder missile in flight. During 1977 work started on the chemical oxygen iodine laser at the USAF's Phillips Laboratory at Kirtland Airforce Base New Mexico, now used in the Air-Borne Laser system.

Activity in the former Soviet Union included the design and demonstration of similar laser technology and also the development of explosively-pumped iodine lasers. These devices have demonstrated output energy of the order of a megajoule. The 1980s saw the start of the American "Star Wars" project, which did much to end the "Cold War"; this project aimed to advance high technology for the defence of the US and its allies, and saw an emphasis on directed-energy techniques, including laser-based devices. Examples included the commissioning of the mid–IR advanced chemical laser (MIRACL) at the HELSTAF site at White Sands Missile Range New Mexico, which helped to consolidate progress with directed energy weapon (DEW) technology. This decade also saw progress with solid-state and liquid-laser technology for anti-sensor applications, known as laser sensor damage weapons (LSDW). This is the so-called in-band damage technique, in contrast to the out-of-band approach used for structural damage defeat systems requiring very powerful "Star Wars" laser-based devices.

Directed infrared countermeasure (DIRCM) systems were developed in the 1990s, using more conventional IR-source technology, in response to the continuing menace presented by IR-homing missiles [5]. Developments involved use of arc-

lamp technology and focussing optics to concentrate the photons into a narrow beam. The Nemesis DIRCM system became operational in this decade. This period saw large advances with compact solid-state IR lasers for use in DIRCM systems. The development of periodically-poled lithium niobate, and its adaptation for use in compact mid–IR systems, was a significant technological advance, together with semi-conductor lasers for pumping the gain medium; semiconductor devices were designed and demonstrated for direct generation in the IR, for a range of applications including countermeasures.

Another notable development in the last 20 years has been the application of IR lasers for atmospheric sensing. A typical application is differential absorption lidar, usually known as DIAL. These are discussed above in the chapter on sensing.

3 Propagation and Atmospheric Windows

IR sensors and systems that operate on or close to the surface of the Earth function in a gaseous blanket that surrounds this planet, which gives a transition between the solid or liquid of the surface of the Earth and the vacuum of outer space. The atmosphere consists of a number of gases and suspended particles, the temperature and pressure of the atmosphere varying with latitude and altitude, as well as exhibiting seasonal, diurnal and local variations. The dynamical variation of the atmosphere is instrumental in weather patterns; moreover, these dynamics have important consequences on the behaviour of IR sensors and countermeasure systems, which may hinder the performance of these systems unless careful design and compensation techniques are applied. Conversely, favourable weather conditions can result in enhanced propagation, but reliance on luck does not lead to a robust countermeasure technique. Hence, favourable propagation conditions may be seen as an enhancement in nominal performance, but should be ignored in effectiveness calculations.

The lowest layer of the atmosphere is known as the troposphere, which extends from sea level to about 11 km (7 miles), and generally the temperature decreases with increasing altitude, at the rate of 6.5 K per kilometre (1 ^{o}F per 300 ft); this is known as the lapse rate. Variation in this rate leads to some interesting refractive effects, such as the formation of mirages. The pressure variation ranges from one atmosphere at sea level to a few tenths of an atmosphere at the top of this layer.

The fundamental physical processes that attenuate the propagating electro-magnetic radiation in the Earth's atmosphere are [6]:

- Resonant absorption by the rotational and vibrational motion of diatomic and triatomic molecules in the atmosphere.
- Rayleigh scattering by molecules in the atmosphere much smaller than the incident wavelength of the energy.
- Mie scattering by the particles comparable in size with the incident wavelength.

The gases in the troposphere that have the greatest impact on the transmission of IR radiation are water vapour, carbon dioxide and ozone. A range of aerosols and particles of varying sizes may be present, such as fog, cloud, smoke and dust. Consequently, the transmission of IR radiation in the atmosphere is affected in the

same manner, as is that of visible light: by absorption and scattering by molecules and by the presence of aerosols and particles. Of the two processes, absorption proves to be the more dominant at IR wavelengths, whereas scattering is the more dominant for visible light.

Figure 3 shows a plot of the percentage transmission against wavelength measured over a horizontal path length of about one nautical mile (1.9 km) at sea level for an atmosphere typical of north-west Europe. The plot shows transmission resulting from the combined effect of molecular and aerosol scattering and absorption; but selective absorption by water vapour, carbon dioxide and ozone molecules are the dominant processes. This figure clearly exhibits several regions of high transmittance known as *atmospheric windows* separated by regions of high absorption. There are two principal windows that may be used by IR sensors, such as thermal imaging cameras; they are known as the 3–5 μm and 8–14 μm atmospheric windows. However, the former one is often considered to be from 2 to 5 μm for other sensors, such as heat-seeking missiles and has been exploited by these systems.

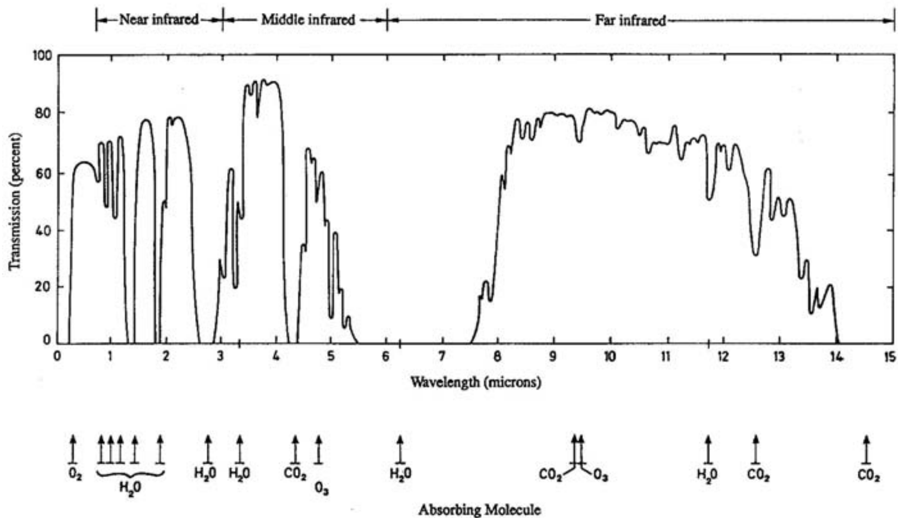


Figure 3. Transmission of a typical atmosphere in NW Europe for a one nautical-mile horizontal path at sea level

3.1 Humidity

Increasing humidity reduces the transmission within each window, owing to increased Mie scattering and continuum absorption. However, the 8–14 μm atmospheric window is more susceptible to increased humidity, and variations in atmospheric humidity are generally greater in summer than in winter in temperate conditions. Therefore, atmospheric transmission within the windows will be less in summer than in the desert or winter, particularly for the 8–14 μm window. From investigation of the contribution of the molecular absorption for typical summer and

winter conditions, and the aerosol absorption and aerosol scattering for clear and hazy conditions for the two atmospheric windows the following conclusions may be reached:

- Transmission in the 8–14 μm window is dominated by molecular absorption.
- The 3–5 μm atmospheric window is influenced more by aerosol scattering in summer than in winter.
- For short ranges (less than 10 km) the 8–14 μm window appears to have the better atmospheric transmission for most circumstances.
- Only in conditions of high humidity and extreme clarity in visual terms does the 3–5 μm atmospheric window have a superior transmission performance, particularly for ranges in excess of 10 km.

3.2 Haze, Fog, Cloud and Rain

The transmission of IR radiation is superior to that of visible wavelengths, particularly if conditions are hazy or conventional smoke is present. This is because the mean particle size of about 0.5 μm is a relatively small fraction of the IR wavelength (say 10 μm) and therefore any scattering by haze or smoke particles will be negligible, particularly if the scattering particles are very much smaller than the IR wavelength (Rayleigh scattering). Scattering by fog and clouds, on the other hand, is intense, because the size of the droplets present extends from 5 to 15 μm, *i.e.* comparable with the wavelength (Mie scattering processes). The presence of a high concentration of aerosols has a an impact on image quality leading to a blurring of the image, which limits the ability of a thermal system to detect, acquire and then track, or identify, a target.

Rain affects the performance of an IR sensor in two ways: it tends to reduce the thermal contrast of the scene by coating everything with a thin layer of water, and it affects transmission in both atmospheric windows almost equally, particularly in heavy rain. Attenuation in these circumstances may amount to as much as 20 per cent per kilometre, giving rise to only 33% transmission over a 5-km path.

3.3 Turbulence

The phenomenon of optical turbulence occurs as a result of fluctuations in the refractive index of the atmosphere within the air mass, along the optical path. Turbulence is responsible for a variety of effects, such as scintillation, *i.e.* the temporal intensity fluctuations of a beam. In the case of laser-beam propagation, owing to the coherent nature of the photons in the beam, the effects of turbulence result in beam wander and beam broadening. In the case of imaging systems, the effects of turbulence result in motion of the image (image dancing), as well as blurring of the image.

Variations in the absorption of energy by the atmosphere, such as solar radiation as well as other effects, leads to the formation of small zones or cells with slight changes in density. The motion of these zones leads to atmospheric turbulence and also, owing to the changes in density, small fluctuations in the index of refraction of these cells in the atmosphere. Despite the fact that the magnitude of the individual

fluctuations in the index refraction is very small, the cumulative effect on the spatial and temporal characteristics of the photon beam as it travels along a path through the atmosphere may be very significant. The turbulence-induced fluctuations in the refractive index produce a distortion in the wavefront of the beam as it passes through the various turbulent zones, and it is this distorted wavefront that continues to propagate and is, itself distorted further. The critical parameter used for characterising the effects of turbulence in the atmosphere on a propagating beam is the refractive index structure parameter C_n^2.

The effects of turbulence [7] are usually most severe close to the ground, particularly on hot sunny days. The effects tend to decrease with altitude[1]; moreover the effects of scintillation are approximately inversely proportional to wavelength. There is a strong link in the characteristics of the turbulence to the thermal structure parameter of the atmosphere (C_T^2) and the difference in the temperature of the ground and the surrounding air [8]. Consequently, there is a diurnal variation in the value of C_n^2 throughout the entire day, with minima close to dawn and dusk when there is a close match between the temperature of the ground and the air. This is shown is shown in Figure 4. Adaptive-optical techniques [9] have been developed to compensate for turbulence-induced effects.

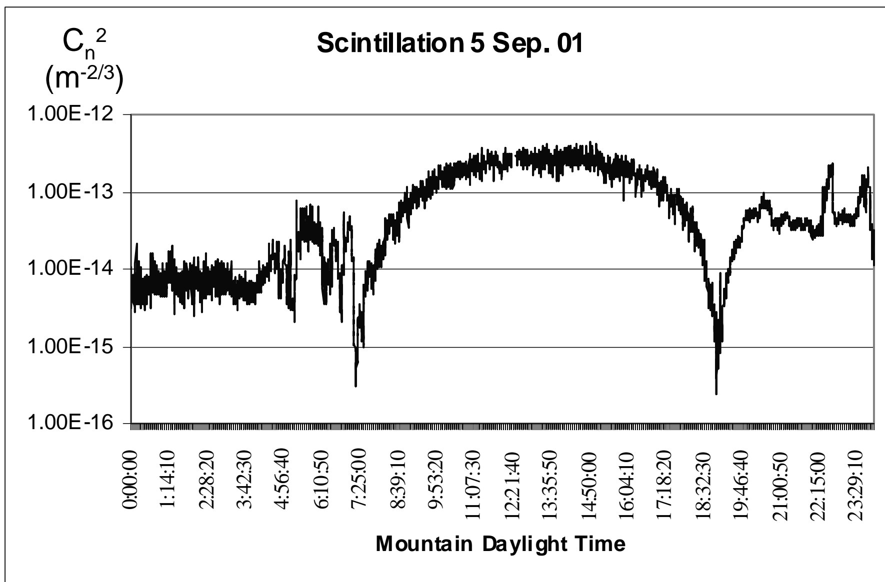


Figure 4. Diurnal variation of the refractive index structure parameter C_n^2

[1] Height (h) $h^{-4/3}$ rule.

Beam spreading effects, as a result of atmospheric turbulence effects, are shown in Figure 5. On the right of the figure is a compensated beam, resulting from the effective compensation from the application of adaptive-optical techniques.

Figure 5. Atmospheric turbulence effects (*left*) and the corrected beam (*right*)

3.4 Extinction

This is the reduction in the intensity of a beam of radiation as it propagates through a medium, owing to the absorption of the photons by the species in the medium and scattering of photons out of the beam by aerosols and particles present in the medium. The application of this phenomenon is exploited in the denial defeat mechanism, with the use of smokes and obscurants.

3.5 Thermal Blooming

This phenomenon usually arises when high-power propagation occurs in the lower atmosphere, which results in changes to the spatial characteristics of the propagating beam. Absorption of energy from photons in the beam by the molecules and aerosols causes local heating of the atmosphere and a consequent change in the local refractive index of the atmosphere. Generally, for a Gaussian beam, more energy is absorbed along the centre of the propagation path, reducing the refractive index. This change in the refractive index creates a lens in the atmosphere that may be either positive or negative. Hence, the IR energy may be focussed, or more usually de-focussed leading to reduced power density at the target area.

Thermal blooming is a problem for high-power directed energy weapon systems, particularly if there is no lateral motion of the beam nor a cross-wind and consequently, is a potential limiting feature. Therefore, great care is required in selecting the wavelength of the high-intensity laser beam. If the target has some lateral component to its motion, so that the laser beam has to pan across a sector, then there is less likelihood of thermal blooming reducing the fluence on the target. The effects of thermal blooming also reduce with increasing altitude.

3.6 Ionisation

When high-intensity laser beams are brought to a focus then the work function of any species may be exceeded by the absorption of the photons, liberating electrons

and creating a plasma. The presence of particles and aerosols may lower the threshold power density for plasma formation; additionally the ablation of solid material from the surface of a target may provide the mechanism for the creation of a plasma. In general, very intense beams of IR radiation are required to ionise clean or pure air.

The formation of a plasma along the propagation path of a laser beam is detrimental to the intensity of the propagating beam, as the beam becomes de-coupled from the target's surface and generally the plasma cloud moves away from the target's surface. However, if the plasma can be constrained to the immediate vicinity of the target's surface, then enhanced coupling may result, but this can be very difficult to maintain in a dynamic environment. Similarly, this phenomenon may be used as a protection measure for blocking intense beams by seeding the formation of a plasma from material ablated from a target's surface.

3.7 Wakes and Plumes

The efflux from a jet engine has hot exhaust gases being mixed rapidly with the cooler ambient atmosphere. This leads to a region of highly turbulent gas stretching out behind an aircraft. This zone behaves as a region with a very large refractive index structure parameter (C_n^2), leading to intensive beam wander and spreading, as well as break-up of the spatial profile of the beam. The impact on a countermeasure system is that the operation of a laser beam from such a device, through an efflux, is to be avoided, *i.e.* it is a blind arc in the system's target engagement zone.

3.8 Aero-optical Effects

This phenomenon arises from the variation in pressure in the airflow around a structure. This variation in pressure leads to a variation in the refractive index of the atmosphere and a bending of an IR beam as it passes through the zone of varying refractive index.

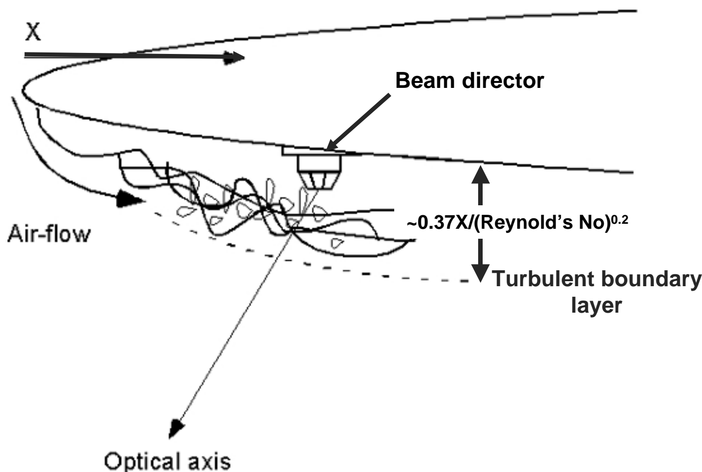


Figure 6. Aero-optical effect on laser propagation

Compensation techniques include estimation of the "optical wedge" using probe beams and applying a suitable aiming correction to the pointing direction of the laser beam. The aero-optical effect is shown schematically in Figure 6.

4 Defeat Mechanisms

A range of techniques [10] and associated technologies may be used to create a defeat mechanism and these are influenced by the type of effect that is required for disruption of a sensor or system. In the sections below there will be some discussion of the normal operation of a sensor and how the defeat mechanism may be applied to achieve its objective, *i.e.* deception or destruction of a sensor or system.

Active countermeasure techniques may operate in a pre-emptive mode or react to an alert concerning the presence of a perceived threat for defeating sensors or seekers. They may be off-board devices, such as smokes and obscurants or flares; alternatively the system may be an integral part of a platform such as an IR jammer; these countermeasure devices are known as "on-board" systems. The damage-class devices may be operated as an independent defensive system or as an adjunct to another defensive asset, such as an air-defence system.

4.1 Denial

The use of denial countermeasure techniques is normally deployed against surveillance and reconnaissance systems such as thermal-imaging devices. A thermal-imaging camera has an array of small detector elements that are placed in the focal plane of an optical system. Each detector is a picture element, or pixel, which records the temperature of the thermal scene within its instantaneous field of view. A range of techniques exists for the sensing and detecting the intensity of photons in a scene by a thermal camera [11]. The image-processing algorithm assembles all of the output signals from each of the pixels in the detector array and creates an image based on the thermal contrast recorded over the complete field of view observed by the entire array of detectors in the focal plane of the sensor. Hence a thermal picture is created using a gray scale and it is refreshed at the frame rate of the camera system.

The use of the denial approach is to confuse the thermal sensor observing a scene, by the reduction of the thermal contrast across the scene so that hot and cold objects appear identical and hence cannot be distinguished. A relatively simple approach is to dispense materials that will extinguish the IR photons emitted by objects in the thermal scene, through absorption, scattering or reflection processes. The fundamental physical processes of these effects have been discussed above in the section on propagation.

Other techniques such as dazzling of a sensor may also be used, but this is considered further below as a separate defeat mechanism. This particular section will concentrate on expendable stores that are fired from a launcher or issued from a dispenser to create a screen that has the specific purpose of extinguishing the photons attempting to propagate through the region where the countermeasure has been applied.

4.1.1 Smokes and Obscurants

Smokes and obscurants rely on optical scattering and absorption techniques to extinguish the light travelling from an object to a sensor, as part of their sensor defeat mechanism. They are used to break lines of sight and deny the enemy the use of their surveillance, target-acquisition or weapon-aiming sensors. Smoke offers the benefit of a low-cost countermeasure: these techniques are a force multiplier. Moreover, it is feasible to deliver a multi-spectral countermeasure; effective simultaneous visible and near IR, mid-wave IR, far IR and millimetre wave (mmW) obscuration has been demonstrated. Importantly, the utility of an obscurant is dependent on prevailing weather conditions and depends on wind speed, wind direction and humidity. Consequently, these factors must be borne in mind before the deployment of an effective smoke screen. Obscuration of objects in the visible part of the spectrum are shown in the upper right-hand image of Figure 7 and the same object, viewed in the IR, is shown in lower part of Figure 7. It should be noted that the bridge can be "seen" by a thermal camera through the visible-band obscurant (lower left image), but is then obscured to this sensor when the IR-band obscurant is applied (bottom right-hand image).

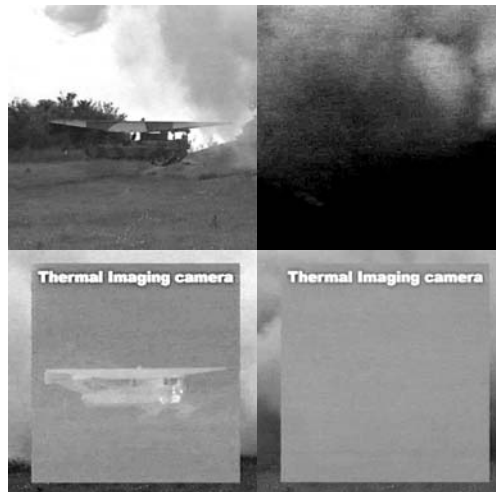


Figure 7. Image obscuration. [Courtesy of QinetiQ]

Typical obscurant applications include.

- Concealment of troop or vehicle advances.
- Protection of vulnerable operations such as bridge laying or minefield breach.
- Protection of land and naval vehicles.

It is possible to deliver a smoke munition either through artillery fire or as a vehicle-launched grenade. A number of different mechanisms are used to generate the obscurant, which are considered in more detail below. Applications include vehicle protection grenades, either stand-alone or as part of a defensive aids system (DAS).

4.1.2 Pyrotechnic Smoke

Pyrotechnic smoke is generated typically through the deployment and initiation of a phosphorus-based munition. The munition can be delivered by artillery or through a vehicle-protection grenade. This can achieve effective obscuration in the visible and near-IR waveband; an advanced phosphorus munition has also achieved simultaneous mid-wave IR, far IR obscuration.

4.1.3 Rapid-bloom Obscurant

Rapid-bloom obscurants (RBO) aim to form an obscurant cloud in less than 1 s. The RBO concept is to disseminate scattering particles such as metal (brass) flakes rapidly by a high explosive burster charge at the centre of the grenade. This can achieve effective obscuration in the visible and near IR, mid-wave IR, far IR wavebands. Recent advances include the incorporation of carbon fibre into the grenade to achieve effective mmW obscuration and a novel scheme has been demonstrated to package the carbon fibre and metal flake payloads into a compact grenade.

4.1.4 Large-area Smoke Screening System

A large-area smoke screening system generates smoke mechanically. Typically, fog oil (a highly refined mineral oil) is injected into the exhaust of a gas turbine. The oil is heated and as it exhausts it expands it cools and condensation forms on the oil droplets. This technique generates an extinguishing large-area visible and near-IR smoke cloud, by injecting graphite powder. Other dopants are available to extend the absorption band to longer wavelengths.

This class of countermeasure system is sensitive to meteorological conditions, and consequently these factors need to be considered before using this type of countermeasure.

4.2 Deception

The aim of a deception-class countermeasure system is to inject false information into a sensor system to create confusion. This countermeasure approach may be illustrated with the injection of false signals into the tracking sensor of a heat-seeking missile, *i.e.* its homing head. A summary of the development of heat-seeking missile homing-head technology is shown in Figure 8.

Many optical seekers have a single detector that is used to sense the position of its victim in its field of view, as shown in Figure 8. A seeker with a single detector has a spinning reticle, or mechanical modulator, in the focal plane of the optical system that collects energy from the thermal scene and focuses it on to the detector [12]. The reticle consists of a series of opaque and transparent segments that create the modulation. Hence, if the reticle has a spin frequency of 100 Hz and 10 transparent and opaque spoke pairs in its structure, the a carrier frequency is generated of 1 kHz, thus the origins of the term spin by carrier jam waveform.

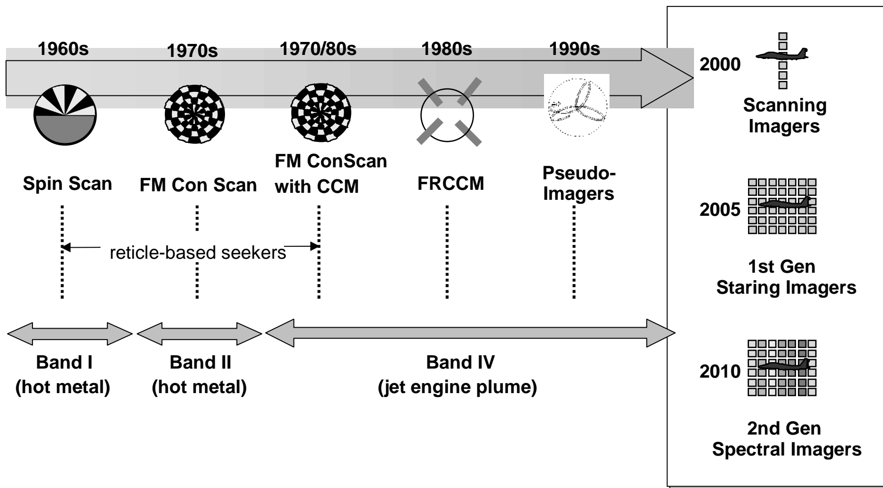


Figure 8. Heat-seeker developments

The reticle creates a series of time-referenced pulses resulting from the operation described above, that are interpreted by the target-tracking system to provide the position of the designated target in polar axes with respect to the optical axis of the seeker system. Another important function of the reticle is to act as a spatial filter to eliminate extended targets, such as clouds. A reticle and its modulation are shown in Figure 9 for a spin-scan system on the left and a more complex, but superior performance system, which undertakes a conical scan of the field of view of the seeker, on the right.

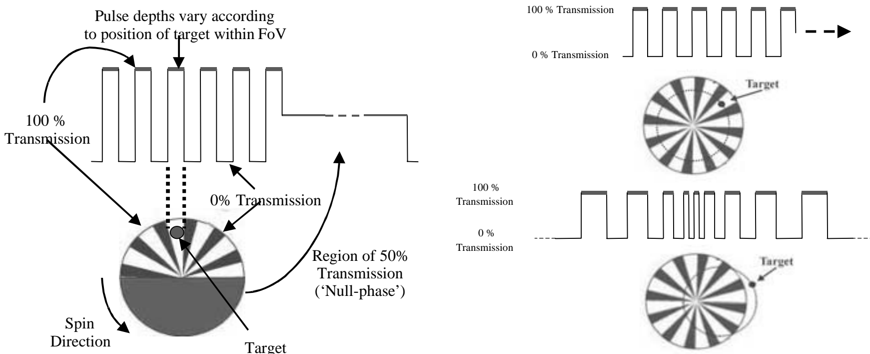


Figure 9. Spin-scan and conically-scanned seeker modulation

The deployment of man-portable air-defence systems (Manpads) and other mobile IR-guided surface-to-air missile systems has meant that aircraft can be attacked at any time by IR-guided weapons. Thus, the option for avoiding the missile

engagement zones by flying over or around these sites, as is used very successfully with systems on fixed sites, is a tactic with limited assured success, particularly if an aircraft is flying from a forward operating base. As a consequence, there is a military operational need for such countermeasure systems.

The success and simplicity of the heat-seeking missile has demanded countermeasure techniques that are simple but effective for defeating these potent weapons. Rapidly effective protection methods are required, as the engagement timelines are very short, typically in the range of a few seconds, particularly for so-called short-range shots of 1 km or so. It is these short engagement time lines that preclude the use of guided weapons. Additionally, these counter-heat-seeking missiles would occupy a number of weapon stations, which would be unacceptable. Moreover, the asymmetric nature of this particular threat means that the use of a hard-kill device is not going to be a cost-effective use of such assets that would require high-performance "guided-weapon" technology.

Use of electro-magnetic energy offers a solution to many of the constraints posed by hard-kill methods. A range of methods has been devised, which fall into two broad classes: on-board techniques embracing jammers and off-board techniques using pyrotechnic flares. In principle the two techniques have the same aim, providing a false target for the threat seeker to "chase". In the former case the jammer sends a series of modulated-IR pulses to the seeker to confuse the target-tracking algorithm, which interprets the pulsed output signals from the detector. In the latter case, a high-intensity IR source is dispensed from the aircraft under attack for the seeker to follow and intercept at a safe distance from the aircraft under attack. Two fundamental approaches have been developed to provide platform protection, based on expendable stores and jammer systems.

4.2.1 Expendables (Flares)

The pyrotechnic store is made of a chemical composition, such as magnesium teflon viton, which burns readily in air and aims to provide a very bright source of IR energy, typically burning at a temperature in the region of 1500–1800 K. Consequently, a point source of energy is created with a large thermal signal in the part of the electromagnetic spectrum used by the homing heads in the heat-seeking missiles. Provided the flare is ignited close to the platform it is protecting and its intensity is greater than the thermal emission from that aircraft, there is a high probability that the seeker will follow (or track) the flare, rather than the aircraft. Hence the seeker will be seduced by the flare and as the flare moves away from the aircraft, the heat-seeking missile will follow the path of the flare and hence miss its intended victim. A further requirement of the flare system is that it should burn for a sufficient period that when it burns out, the "dispensing" aircraft is out of the field of view of the seeker. The launch requirement of a flare being dispensed form an aircraft is shown in Figure 10.

Modern IR-missile systems may use techniques to discriminate between the spectral signature of a simple or conventional flare and an aircraft. As a consequence, spectral flares have been developed that have a "balanced output" in the various spectral regions of the mid–IR, so that the thermal signature of the flare appears similar to that of an aircraft. Other recent developments have included the design and demonstration of flares that spread their emitting fragments over a large

area, in order to screen a target from the advanced seekers, such as those using pseudo-imaging techniques. Special materials have been developed to ensure a covert appearance in the visible band. Other developments have seen the use of pyrophoric materials, which offer a very compact and efficient IR source.

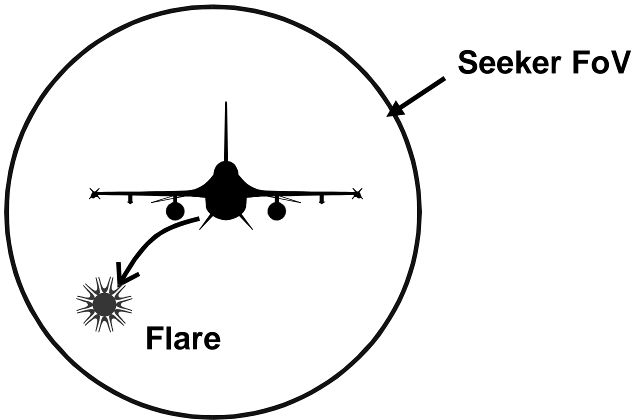


Figure 10. Flare launch geometry

Sequences of flares can be emitted to ensure confusion of the target-tracking system in the approaching missile, or salvoes of missiles; a spectacular example is shown in Figure 11. However, it should be noted that the normal mode of operation is three or four flares in a sequential pattern, shown on the left of Figure 11.

Figure 11. Flare countermeasures [Courtesy of AFRL] reproduced with kind permission of AgustaWestland ©

Off-board systems offer a simple but effective-IR countermeasure system. They are particularly effective for countering a range of IR missiles; in the case of the modern IR-imaging threats they are likely to be particularly effective when an aircraft is

attacked from the rear sector. The major drawback concerns their limited magazine capacity, rules of engagement that limit their deployment in some circumstances and the rapid separation between the flare and the aircraft during engagements in the forward sector.

4.2.2 On-board Techniques (Jammers)

On-board systems are commonly called “jammers”. The basis of this defeat mechanism is a source of thermal energy that is modulated to create a series of pulses of energy such that this waveform of pulses modifies the modulation created by the reticle in the seeker of the heat-seeking missile. Like the flare, the brightness (or radiant intensity) of the thermal source needs to get the attention of the homing head in the heat-seeking missile. This is achieved by ensuring the radiant intensity of the jammer exceeds the radiant intensity of the platform it is protecting by a significant factor, usually by at least 10. This is known as the jam-to-signal ratio.

Most modern seekers can be defeated by using a pulsed modulation known as a spin by carrier waveform that has a spin envelope closely matched to the spin frequency of the reticle and a carrier frequency induced by the pairs of transparent and opaque blades of the reticle. An idealised jamming effect is shown in Figure 12; in reality, the result is more complex owing to the motion of the seeker.

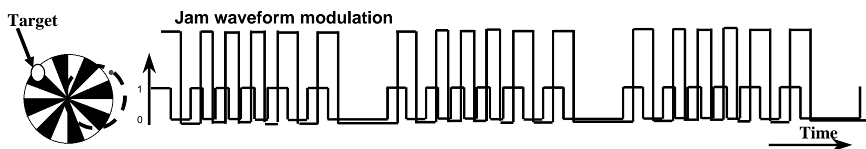


Figure 12. Jamming effects

When the J:S is of the order of 10 then the jammer tends to induce a series of confusing “commands” into the tracking system resulting in perturbations in the missile trajectory, often referred to as guidance-gain degradation. For this defeat mechanism to succeed the jammer must illuminate the target throughout its approach to its victim. The consequence is the complex spiral-like motion of the missile about its notional trajectory to its target, leading to a large miss distance. A more robust defeat mechanism is known as optical break lock (OBL); in this case the deception commands cause the seeker to believe its victim is moving rapidly out of its field of view. As a result the target-tracking system generates a large angular demand on the seeker and the sightline to its victim is lost as the seeker makes a rapid angular deflection in response to the demands from the target-tracking system. Moreover, the missile makes a violent manoeuvre away from its victim. Very large J:S is required to ensure the multiple scatter and reflection of jammer photons in the seeker is needed to achieve this level of deception, typically greater than 100:1. The smoke trail of a missile that has rapidly changed direction as a result of an optical break lock condition is shown in the image in Figure 13.

The two fundamental techniques of flares and jammers used for IR countermeasures (IRCM) apply very different approaches for achieving protection of a platform through inducing a deception defeat mechanism in a heat-seeking

sensor. However, these approaches can be combined to provide a robust technique for defeat of advanced seekers shown on the right of Figure 8. In this case the laser may be used to “mask” the ejection of a flare, so that when the laser is turned off the target-tracking system is confused by the rapidly changing thermal scene.

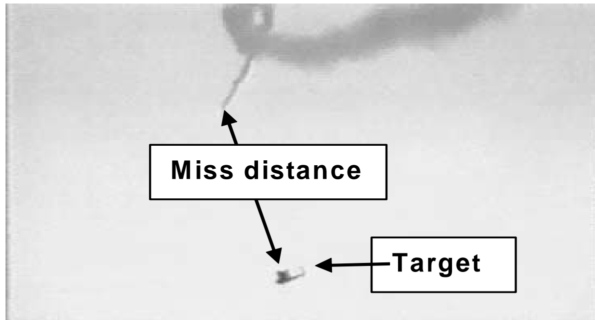


Figure 13. Rapid optical break lock trajectory [Courtesy Northrop Grumman]

An alternative technique that has been investigated is a modern-day form of camouflage used by a chameleon. In this case the asset under threat is able to vary its thermal signature so that it is indistinguishable from its background and hence not able to be detected. This may be achieved with a series of variable emissivity panels; however, this technique is very challenging when the background is complex and rapidly changing, or is a “clear blue sky”. It is a part of the countermeasure technique known as camouflage, concealment and deception.

Laser-based DIRCM devices are considered further in the section on the evolution laser-based systems as they are probably the most advanced of this class of on-board countermeasure systems.

The laser requirements for on-board systems are summarised below.

- The wavelength must be co-incident with the pass band of the seeker.
- There are no special requirements on spectral purity and avoiding chirping of the wavelength, providing it does not incur any detrimental propagation effects.
- The pulse-repetition frequency must be compatible with the modulations required to form the jam waveform.
- The modulation depth is required to be of the order of 1000:1 in order to avoid enhancing the thermal signature of the platform being protected.
- The average-power requirements are modest, of the order of a few watts.
- The beam divergence needs to be matched to the pointing characteristics of the beam director.

Environmental factors, such as optical turbulence, have an impact on the performance of jammer systems, limiting the tracking accuracy. Turbulence distorts the image that the target-tracking system is operating on, as well as modifying the spatial and temporal characteristics of laser-based systems. Turbulence has a negligible impact on the modulation of the waveform, but can decrease the J:S.

4.3 Dazzle

Dazzle-countermeasure techniques may be applied to thermal sensors used for surveillance, reconnaissance, or target tracking. These functions are often undertaken with very small "target" signals in a noisy background, so highly sensitive detection systems are necessary and hence can be vulnerable to intense optical energy directed into their optical apertures. The dazzling phenomenon only persists whilst the bright optical source is present; normally a sensor system will revert to its normal operational performance characteristics once the intense thermal source is extinguished, as the photon flux is usually below the damage threshold of the vulnerable elements in the optical system.

The use of the dazzle-countermeasure technique may be considered as a variant of the denial method, but the important difference is that the effect is within the optical train of the sensor, rather than in the actual scene being observed. This technique is usually associated with a very bright source of photons being directed into the optical train of a sensor, but it is possible to dazzle a system with expendable stores such as pyrotechnic flares and similar hot sources.

A dazzle-countermeasure system directs a very bright source of energy at the objective lens or optical aperture of a sensing system. It is essential that the optical train of the sensor system being dazzled is transparent to the wavelength being used, as described in some detail below in the section on "in-band" damage. The principle of this defeat mechanism is to create intense optical scatter from each optical surface within the optical train of the system, so that "stray light" enters the detection array and the thermal picture is obscured. The development of efficient IR-laser technology has created an efficient means for dazzling thermal sensors. This is because of their inherently large radiant intensity, readily creating intense optical scatter from minor imperfection of any optical component within the optical train of the system and any transparent object between it and the laser-beam director, or laser-pointing system. Moreover, the positioning of the optical elements in the system and the supporting structures for these optical components can create ghost images and striking diffraction effects. These combined effects can result in extensive confusion in the thermal scene being observed, and in the case of a target-tracking system loss of the tracked object. Some characteristic effects are shown in Figure 14.

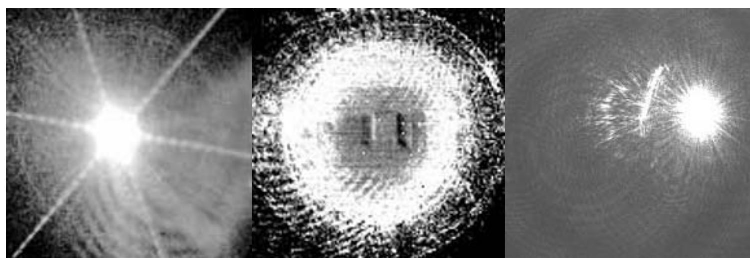


Figure 14. Optical dazzle effects

A subsidiary phenomenon that helps in this defeat mechanism is an effect known as blooming, where the intense beam is focussed on to the detector array. The detectors

within the focussed spot fill their wells in the read-out circuits and there is overspill of electrons into the adjacent wells associated with the corresponding adjacent detectors. Hence, the image of the incident beam appears to be much larger than it actually is and denies information about a larger zone of the thermal scene.

A further aspect of this approach is the thermal heating and subsequent re-radiation from optical components into the optical train that leads to a reduction of thermal contrast in the scene being observed. This effect occurs when the optical wavelength is not well matched to an optical element, or elements, in a refractive optical train; it has some similarities with kinetic heating of IR domes on the front of high-speed missiles.

The fundamental requirements of the laser emission for this countermeasure approach are:

- Wavelength matched to the optical pass band of the sensor, but there are not special requirements on spectral purity and avoiding chirping of the wavelength.
- Temporal characteristics matched to the frame rate of the sensor, so that it is not necessary to use a continuous-wave laser emission unless the frame rate of the sensor is extremely high. In practice, it is normal for the pulse-repetition rate to be somewhat larger than the frame rate of the sensor. Furthermore, the pulse duration should be much larger than the integration time of the sensor, so that the photons enter the sensor during the period it is observing a scene.
- Spatial characteristics are not normally demanding, as the intention is to direct a bright beam at the target area, but not to create permanent damage to the sensor. Relaxing the beam-divergence parameters means that the laser beam is easier to maintain on a target, although this can increase the laser average power requirements.
- Average laser-power requirements are in the region of about 5–10 W and a beam divergence of around a milliradian for longer-range engagements.

Pyrotechnic flares may also be used for this task. The local environmental factors have an impact on laser-based systems similar to those described above for deception systems.

4.4 Damage

This approach to optical countermeasures aims to destroy a sensor, a system or a structure. The defeat mechanism for this type of system is through rapid deposition and, thus, absorption of optical energy, which destroys the integrity of its target material. It is an attractive countermeasure approach as it can have a rapid engagement of a target, enabling the so-called short-range, pop-up target to be engaged and defeated. A typical engagement sequence is shown in Figure 15. As discussed below, there are a couple of approaches to an effective defeat mechanism, known as in-band and out-of-band damage and these are shown in Figure 16. As can be inferred from this figure, these techniques are somewhat different in approach, and consequently, have different requirements for the laser sources used in these devices in order to achieve the desired effects. The critical issues with this type of system concern laser-target interaction effects, propagation and aberrations induced by the atmosphere for endo-atmospheric engagements.

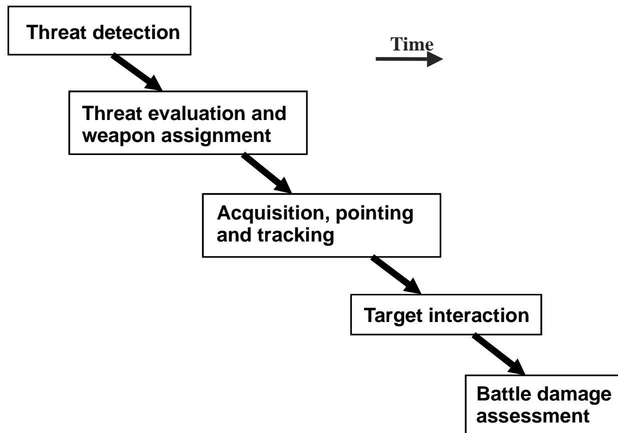


Figure 15. Engagement sequence

The use of the damage defeat mechanism may be considered the next escalation or "step up" in the use of countermeasure approaches, beyond dazzling a target sensor, for defeating targets, to generate a more robust defeat mechanism. The "in-band" technique is usually associated with damaging a component of a sensor, whereas the out-of-band technique aims to defeat a complete sub-system.

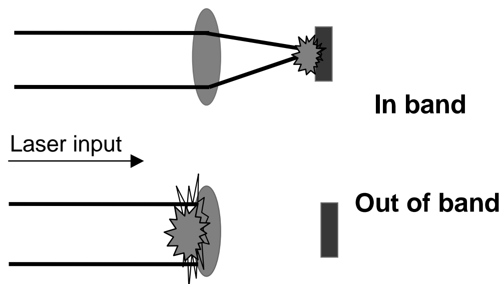


Figure 16. Laser damage routes

4.4.1 In-band Damage Route

The principle of the in-band damage approach is to use the optical gain of the sensor being attacked to focus the incident energy on to a vulnerable component in a focal

plane of the optical sensor[2]. This enables a modest average-power laser to be used to create permanent damage to a sensor, which occurs when the laser-induced damage threshold [13] is exceeded on the vulnerable component. Clearly, the incident energy needs to have a wavelength that is not attenuated by the optical train of the sensor, as well as coinciding with an atmospheric propagation window. This is the basis of a laser sensor-damage-weapon system, mentioned earlier.

An attraction of this approach is that a very short interaction time is required for destroying a sensor, and in theory, could be achieved from a single pulse of energy. In practice, a burst of pulses is used over a very short period in order to increase the hit chance and ameliorating the stochastic processes occurring along the propagation path in the atmosphere, such as turbulence-induced beam-pointing effects. As a result a relatively simple beam-director device can be used, as the interaction time with the target is relatively short and hence a low-cost countermeasures system is possible.

The attraction of laser technology for this type of application is that it offers a robust defeat mechanism and the fact that the countermeasure system can make optimum use of the advantages of laser-weapon technology:

- Potentially a very deep magazine (from the exceedingly high number of "shots" that may be generated)
- Speed-of-light interaction
- Line-of-sight engagement, so it is not necessary to track the motion of a moving target accurately over an extended period
- Very short engagement time lines, so targets that unmask at very short ranges can be engaged
- Multiple-target engagement capability
- Surgical strike with negligible collateral damage

The major drawbacks of laser-based countermeasure systems are usually quoted as:

- Finite interaction time with the target
- Need for an accurate beam director for ensuring the laser beam remains on the target during period required for the finite interaction time
- Complexity of the entire laser-based system, particularly the beam director
- Lack of any battle-damage assessment capability; however, this aspect is not unique to laser-based countermeasure systems

In the case of a laser-based countermeasure system, using an in-band damage route, a number of the cited drawbacks listed above really do not apply. The major drawback of this approach is that it is reliant on some co-operation from the threat system, *i.e.* the threat system needs to be looking at the asset, such as a military vehicle, being protected by the laser-based countermeasure system. Moreover, the operational wavebands of the threat and the emitted laser line need to be compatible. However, it may be deduced that if the threat system is observing the asset on a

[2] An unintentional demonstration was given to the world during the Apollo 12 mission on the moon when astronaut Bean pointed a colour vidicon camera at the sun. The consequence was a loss of televised pictures from this moon landing.

battlefield then it is likely to have hostile intent. The laser requirements for this class of countermeasure system may be summarised as follows:

- Wavelength matched to the optical pass band of the sensor and in an atmospheric window, but there are not special requirements on spectral purity and avoiding chirping of the emitted laser wavelength.
- Pulse energy needs to be such that when focussed by the optical train of the sensor under attack, the fluence in the focal plane exceeds the laser-induced damage threshold [13, 14] of the component positioned in the focal plane, usually a detector, but it may be any one of a number of other optical components – typical pulse energy requirements of the source are of the order of a joule.
- Pulse-repetition rate may be low, as a single pulse of energy is all that is required. However, owing to the dynamic behaviour of the atmosphere, as described in Section 3 of this chapter, there are many stochastic processes that may reduce the brightness of the laser beam and also modify its direction of propagation. Consequently, a short burst of pulses is often used in a macro pulse of duration of the order of a millisecond. Individual pulses may be many orders of magnitude shorter, in order to fulfil the Bartoli criterion [14].
- Beam divergence is usually a compromise between the efficient use of the laser-resonator configuration that enables the maximum amount of energy to be extracted from the gain medium and the pointing stability of the beam director. A beam-expansion telescope may be used to modify the divergence of the beam to match the characteristics of the beam-pointing system. The actual requirement is beam brightness, as that parameter is a fundamental aspect of achieving the requisite energy density on the optical aperture of a system under attack. Typical requirements for the launched beam divergence are of the order of a milliradian.

This damage route is viewed as a technique for defeating the emerging imaging threat seeker technology. An alternative method concerns the use of flares and a laser to defeat the target-tracking system. In this case the release of a flare may be masked by the short emission from a laser, so that the target profile or centroid appears to change from frame to frame and thus disrupt the tracking process.

Turbulence in the atmosphere can have a significant impact on the effectiveness of this class of system, owing to distortion of the optical image used to track and "identify" the threat, as well as the modification of the spatial characteristics of the laser beam. The beam jitter induced by turbulent motion of the atmosphere can cause the beam to miss its target! Thermal blooming effects are unlikely to be significant except for very high-power systems.

4.4.2 Out-of-band Damage

The in-band approach discussed above is quite sophisticated, as it relies on coincidence of the optical transparency of the targeted sensor and the damaging pulse, as well as the fact that this sensor needs to be observing the laser-weapon system. In contrast, the out-of-band systems do not have such restrictions and can cause damage when the incident energy is absorbed by the outer structure of an object, such as a missile.

Application and use of an out-of-band directed energy weapon has a number of implications, both for system performance and tactics. The interaction of the laser beam with the target is likely to last a number of seconds so the beam director must track the target very accurately and point the beam with commensurate accuracy, often to a few tens of microradians[3]. The effectiveness of this type of directed energy weapon system has some sensitivity to the aspect presented by a target, but it is capable of providing support defence capability, by intercepting crossing targets. In fact, this is the preferred engagement aspect for many directed energy weapon systems, owing to the shorter interaction length in the threat system. This is because of the thin skin and proximity of vulnerable components to the outer skin of a missile.

This type of directed energy weapon system has a significant impact on the laser power required to invoke the defeat mechanism, as it involves melting, vaporisation or shattering of a component. Figure 17 indicates the severe demands on laser performance required to achieve significant target damage effects at tactically meaningful ranges. Directed energy weapon systems require very efficient high-power lasers in order to minimise the amount of supplied power and waste-heat management. Moreover, very high accuracy pointing-system techniques are required to provide "time on target" for the interaction to take place.

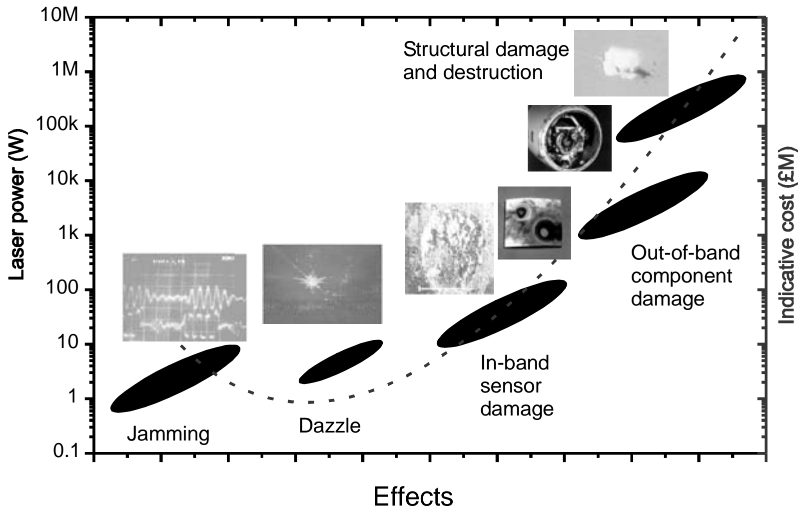


Figure 17. Laser effects summary

The laser requirements for this class of countermeasure system may be summarised as follows:

- Wavelength does not need to be matched to the optical pass band of the sensor, but it should be chosen to minimise absorption and scattering in the atmosphere

---

[3] This exacting requirement may be envisaged as holding a laser beam the size of a dinner plate on a remote target, many kilometres away, whilst the atmosphere shimmers with a time constant of a few milliseconds, over a period of a few seconds.

and the aberrations that follow from non-optimal transmission through the atmosphere.

- There are no special requirements on spectral purity and avoiding chirping of the emitted wavelength, providing it does not incur any detrimental propagation effects.
- Pulsed or continuous-wave emissions may be used to damage a component or cut through the thin skin of a threat. In the case of a pulsed system, the pulse-repetition rate is usually in the kilohertz regime for cutting a hole in a structure, whereas a low repetition rate of high-energy pulses may be used to fracture a dielectric panel or cover.
- Power or average-power requirement is many tens of kilowatts; however, in this case an equally important fundamental consideration is the total efficiency of the laser, often termed the wall-plug efficiency, owing to the large powers involved to operate the laser and the management of the waste heat.
- Beam divergence is the usual compromise between the competing factors, considered above, but the actual requirement is beam brightness. A typical laser will have a very small beam divergence, possibly of the order of a 100 μrad, or smaller, which has a significant impact on the pointing stability of the laser-beam director.

4.5 Destruction

The various defeat mechanisms and techniques considered above may be thought of as tactical countermeasures approaches or systems. The destruction of a target by vaporising a substantial part of it, cutting off a wing of a missile or fracturing the body of a vehicle may be considered to be a strategic defensive system. This is because the laser-based system will be very large and require substantial sub-systems to ensure the optimum laser-beam characteristics are achieved. Such sub-systems include adaptive-optical techniques [9] to correct for aberrations induced by the atmosphere and very sophisticated target-tracking techniques [15] to ensure the beam is maintained and focussed on a specific point on a selected target.

A number of mechanisms may be invoked to destroy a target, such as:

- Vaporisation of a surface
- Spallation or shattering of a surface
- Induced hydraulic shock of a sealed system

The laser requirements for this class of countermeasure system are by far the most demanding of any of the approaches considered in this review; they may be summarised as follows:

- Wavelength does not need to be matched to the optical pass band of the sensor, but it should be chosen to minimise absorption and scattering in the atmosphere and the aberrations that follow from non-optimal transmission through the atmosphere.
- There are no special requirements on spectral purity and avoiding chirping of the wavelength, providing very low-loss transmission is maintained.

- Pulsed or continuous-wave emissions may be used to create damage in a structure. Many of the candidate laser technologies are continuous-wave devices, as their photons are generated from chemical lasers, such as the deuterium-fluoride devices [16]. This form of laser-based system frequently uses chemical reactions to create the population inversion required for laser action, as this in the most efficient and the management of waste heat may be minimised. However, this class of laser action, particularly when using a deuterium fluoride gain medium, may limit the type of vehicle that carries this type of weapon system. Typically these lasers are megawatt-class devices.
- In the case of a pulsed system the pulse-repetition rate may be high or low depending on the defeat mechanism being invoked. For example the explosively-pumped iodine laser is a single-shot device emitting a huge pulse in the megajoule class.
- Beam divergence of the emission is usually very small, in the few tens on microradians regime, but there has to be a compromise between the competing factors considered above and also avoiding ionisation effects that lead to plasma formation.

The Theater High Energy Laser (THEL) system has demonstrated techniques for destroying Katyusha rockets [17] and this is demonstrated in the series of images shown in Figure 18. In this case the damage-class laser is a deuterium fluoride device.

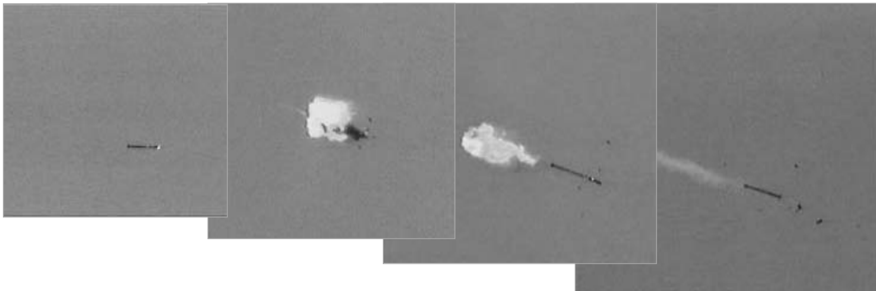


Figure 18. Katyusha rocket defeat [Courtesy Northrop Grumman]

The American Air Force Air-Borne Laser project is an example of a strategic system, with a vast array of sub-systems devised to ensure that the optimum fluence is delivered to a remote target to ensure its destruction. Many of the sub-systems contain lasers to ensure the damage beam hits the selected aim point, remains at that point and is focussed so that the optimum fluence is delivered to the target and it is destroyed in the shortest possible time. A schematic illustration of the Air-Borne Laser system is shown in Figure 19.

Figure 19. The Air-Borne Laser system [Courtesy AFRL]

The impact of the environment of this class of IR-countermeasure system can be one of the limiting features, as all of the phenomena described in section 3 may be invoked. In the limit, thermal blooming may cause focussing leading to plasma formation and de-coupling of the beam from the target.

5 Evolution of IR-jammer Systems

The original jammers were designed to defeat seekers that were tracking and homing on to the emissions from the hot metal in jet engines, as shown in Figure 8. The first jammers used either burning fuel or heated carbon rods to provide the IR-jam energy. The temperature of the source was selected to have a colour temperature that would give a substantial IR emission around 2 μm. Typical colour temperatures were close to 1000 K

The jam waveform, needed to confuse the target-tracking system in the missile, was modulated mechanically with a device known as a chopper. The jam-to-signal ratio was typically of the order of 5:1, to ensure that the seeker was receptive to the jam energy, *i.e.* the jam waveform is more attractive than the thermal signature of the victim. The great attraction of these first-generation IRCM jammer systems was their inherent simplicity. The IR emission was omni-directional, with silicon "windows" making the emission covert in the visible bands. The IRCM system could be activated on take-off, if the aircraft was likely to be attacked, and switched off on its return to base, or leaving the danger zone.

Advances in missile-seeker technology reduced the effectiveness of the first-generation jammers, owing to the move to seekers with longer wavelength sensitivity, as shown in Figure 8. A consequence of the development in detector technology meant that IR systems were required with more jam power, particularly around the 4 μm waveband. Moreover, there was a requirement for greater jam power in order to defeat these newer missiles and protect aircraft with larger thermal

signatures. This requirement led to the development of directed infrared countermeasure (DIRCM) systems. The first effective DIRCM system was the CELEOS-DIRCM system with a caesium arc lamp. The Nemesis system uses a xenon arc lamp with a higher colour temperature and focuses this energy into a narrow beam.

In the first-generation DIRCM system the jam energy from an incoherent source is focussed to increase the radiant intensity of the beam. However, DIRCM systems are more complex than the original omni-directional systems, there is now a need to know the direction of approach of the threat in order to point the narrow jam beam in the appropriate direction (of the threat). The latest stage in the development of DIRCM systems is the transition to laser-based jam sources.

The evolution of jam sources from incoherent to coherent sources and the impact on DIRCM systems and their performance is considered below.

5.1 Incoherent IR Sources

All objects whose temperature is above absolute zero will emit electromagnetic radiation and this was the starting point for the design of IR-jammer sources. The fundamentals are considered below:

- A black-body object is one that absorbs all radiant energy incident upon it, regardless of wavelength and emission depends only on the absolute temperature of the body
- Planck's radiation law gives the spectral composition of the radiation emitted by a black body at an absolute temperature (T), and is shown in Figure 20. The total radiant emittance (area under the curve) increases rapidly with increasing temperature, and that each curve exhibits a maximum at a definite wavelength and that the higher the temperature the shorter that wavelength. Planck's Law can be expressed as

$$W_\lambda = \frac{C_1}{\lambda^5 \left[\exp\left(\frac{C_2}{\lambda T} \right) - 1 \right]}$$

where:

W_λ = spectral radiant emittance (W m^{-2} μm^{-1})
C_1 = first radiation constant ($2\pi hc^2$)
= 3.74×10^8 W m^{-2} μm^4
C_2 = second radiation constant (hc/k)
= 1.44×10^4 μm K
h = Planck's constant
c = velocity of light
k = Boltzmann's constant.

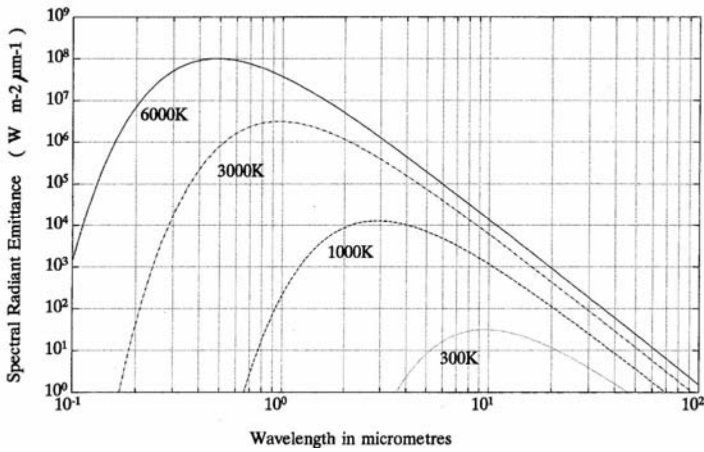


Figure 20. Planck's radiation law

- Stefan-Boltzmann law gives the total radiant emittance (W) of a black-body at a given temperature: $W = \sigma T^4$ where σ = Stefan-Boltzmann constant = $5.67\text{x}10^{-8}$ W m^{-2} K^{-4}.
- Wien's displacement law gives the maximum of the spectral emission for a given absolute temperature: λ_{max} T = constant = 2898 µm K where λ_{max} is the wavelength of maximum spectral radiant emittance. So for T = 6000 K (approximate temperature of the solar photosphere) λ_{max} ~0.5 µm.
- Emissivity (ϵ) represents the “efficiency” factor of the radiant emission from real bodies. The value of emissivity for a real body is the ratio of the radiant emittance of the real body to that of a black body, at the same temperature.
- A gray-body is one where ϵ_λ is independent of λ, and is typical of solids
- A selective radiator, where ϵ_λ varies with λ. and is typical of gaseous emitters.

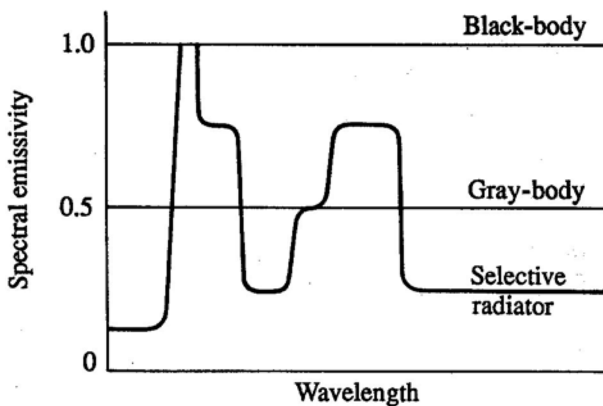


Figure 21. Spectral emissivities

Hence, the radiant emittance from a real body is calculated by multiplying Planck's radiation law by the relevant emissivity values. This may be seen illustrated in Figures 21 and 22.

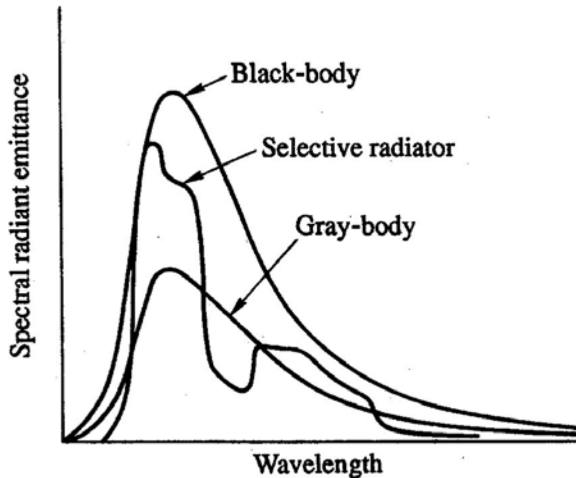


Figure 22. Resultant spectral radiant emittance

5.2 Arc lamps

The development of metal vapour arc lamps as the source of IR energy enabled significantly larger jam powers to be emitted more efficiently compared with heated rods. This emission from an arc lamp has a spectral emission comparable with a black body with a given colour temperature. Typically caesium arc lamps are operated to give a colour temperature around 3000 K and xenon arc lamps are operated with an equivalent colour temperature of 6000 K. Moreover, arc lamps can be modulated electronically to give better modulation depths, leading to more effective jam waveforms, through reduction of un-modulated jam energy (known as "self S") augmenting the platform signature.

The superior performance of arc lamps enabled platforms with a larger thermal signature to be protected from the simpler threat seekers. Additionally, this jam source technology provides some capability for defeating more modern threats, which operate in the longer wavelength band of the mid–IR, referred to as band IV[4], with emission around 4 μm, as indicated in Figure 8.

Further developments were required in order to provide sufficient jam power (radiant intensity) in the new seeker operating band in the "4 μm" region.. Energy from an arc lamp was focussed into a relatively narrow beam and steered over a field of regard using a gimballed beam director [18]. This modification to the emission enhanced the radiant intensity of the "jammer beam", but led to the complication of the jammer system. Figure 23 illustrates this technique and shows how it enhances

[4] Band IV is the spectral region of the mid–IR from 3.8 to 4.7 μm.

the “J” over a narrow beam. The polar diagram on the left shows how the move to the longer seeker wavelengths put the omni-directional jammer at a disadvantage, but when the jam energy is focussed into a narrow beam the radiant intensity of the jammer exceeds that of the aircraft, for a given emitted power.

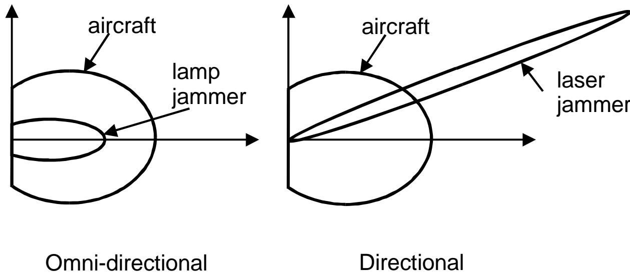


Figure 23. Transition from IRCM to DIRCM

The use of arc-lamp technology allows some adjustment of the power emitted from this device; as the colour temperature increases the spectral radiant emittance increases according to Plank’s law. However, the peak emission moves to shorter wavelengths as the temperature increases, according to Wein’s law.

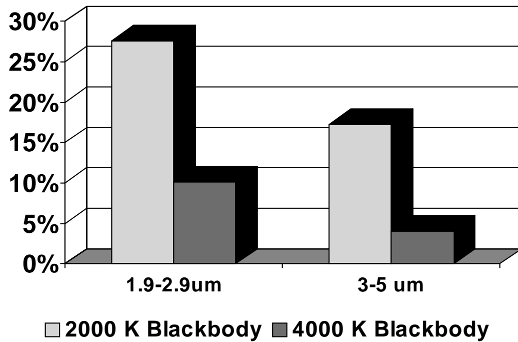


Figure 24. Relative IR emission from arc lamps

Thus, increasing the colour temperature increases the emission at all wavelengths, including longer wavelengths, as required to defeat the more modern threats, but the percentage of the emission in the usable region of the IR decreases. Hence the IR source appears to become less efficient as even at high colour temperature only a small fraction of the total emission is in the mid-wave IR. This is illustrated in Figure 24 showing the fraction of the IR energy in the two principal seeker operating bands as a percentage of the total emission, for two colour temperatures of an arc lamp. The other disadvantages of arc-lamp technology are considered to be:

- Requirement for large power supplies
- Need to shield switching circuits and carriage of arc lamp on gimbal structure to minimise interference with the beam director's control system
- Movement of arc between electrodes that modifies the focussing of the beam, as well as leading to a potential shorting of its operational life
- Relatively large angular divergence, giving less than optimum radiant intensity
- Relatively low total efficiency

5.3 Coherent sources

The inherent characteristics of the laser [19] mean that the beam has a natural low divergence and therefore techniques are not required to "focus" the beam. Consequently, the laser offers the potential of a very bright source and hence is capable of protecting platforms with very high thermal signatures given adequate jam power. The high radiant intensity of a laser source, offering J:S ratios of 1000:1 or more, can be used to exploit new defeat mechanisms that lead to a rapid defeat of the seeker. Additionally, the low divergence means there is a low probability of intercept for off-axis observers.

Laser technology offers the benefit of generating spectral lines in the appropriate seeker operating bands such as around 2 and 4 μm. The use of wavelength tuning techniques enables a number of spectral lines to be generated in the appropriate parts of these seeker operating bands, and moreover, in parts of the atmospheric windows where the absorption is a minimum [6].

Solid-state laser technology has the characteristics demanded by a countermeasure system, such as being compact, efficient, inert and potential for a long shelf life. However, the major snag with this technology today is that very few sources, with the appropriate characteristics, are capable of direct generation, *i.e.* not using non-linear conversion methods. Non-linear conversion adds complexity and causes inefficiency [19].

Various types of semiconductor laser technology [16, 19] offer techniques that may be suitable for laser-based systems, particularly for protection of small-signature platforms. One technique that continues to develop is one involving optical pumping of a semiconductor substrate containing multiple quantum wells, known as an optically pumped semiconductor laser (OPSL). This is a relatively simple approach that offers the possibility of a compact source that has a wavelength diverse output, achieved by minor modifications to the growth pattern used to create the quantum wells. The major challenge for OPSLs is gaining efficient room-temperature operation. Currently, these devices operate around 80 K, and therefore require a cooling engine.

Other semiconductor laser research involves development of diodes, but achieving efficient Band IV emission is still a significant challenge. Approaches involving inter-sub-band transitions in the conduction band of the semiconductor are being developed.

5.4 DIRCM Systems

The first-generation (static) IRCM systems could simply be turned on at some suitable point in a mission and they would emit jam energy until turned off or until the aircraft landed. The modern DIRCM system needs an alerting system to warn the platform protection system that a missile is approaching on a given bearing.

This alerting system, known as a missile approach warning system (MAWS), are devices with a wide field of view and are used to detect the launch of a missile, either from their UV or IR emission in the rocket exhaust. A combination of the "sensed position" of the signal on the sensor array, inertial-measurement techniques [18] and tracking techniques provides the bearing of the approaching missile. This sub-system provides the bearing data of this threat in an appropriate reference frame to the beam director, which is slewed to the missile's approach direction and the thermal scene around this direction is viewed with a thermal camera. This designated threat is then tracked by the thermal camera in the beam director.

There are a number of challenges for these MAWS sensors to provide high sensitivity for long-range detection and acquisition of targets, whilst maintaining a low false alarm rate, so that the potential long-range countermeasure capability may be exploited. Other technical issues concern the harmony between the distributed MAWS sensors and the remainder of the DIRCM system, needed to have a high probability of handover from the MAWS to the beam director. Various techniques [18] may be used to improve the accuracy of the data supplied by the MAWS to ensure a high probability of the target being acquired by the thermal camera, which is used to track the designated missile and the jamming can begin. Figure 25 shows a generic DIRCM-jammer system.

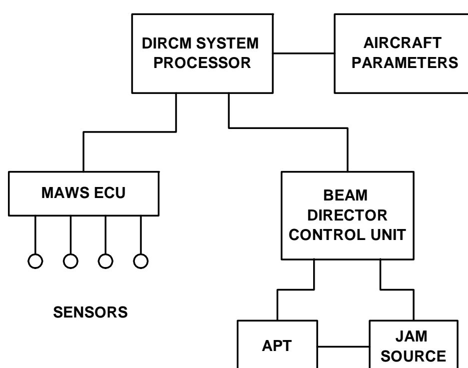


Figure 25. Generic DIRCM system

A typical engagement sequence has an alert provided by the MAWS that gives the bearing to the DIRCM processor that generates the slewing commands for the beam director. The thermal camera searches for the designated target and tracks it and virtually simultaneously applies a jam waveform. The whole process up to the start of jamming may be less than a second or two. The jamming sequence is shown with

a helicopter in Figure 26. Figure 27 shows the small Nemesis system, an in-service system.

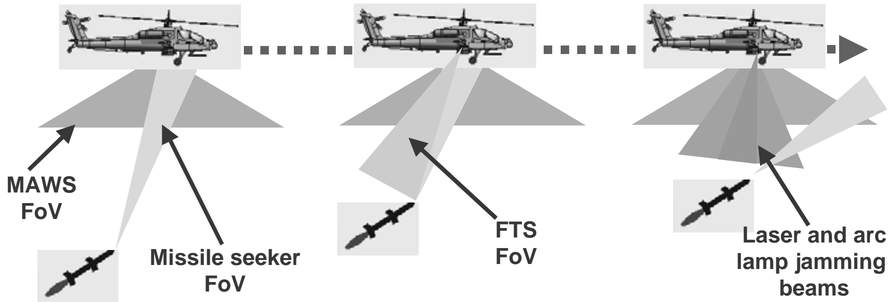


Figure 26. A jamming sequence [Courtesy BAE Systems]

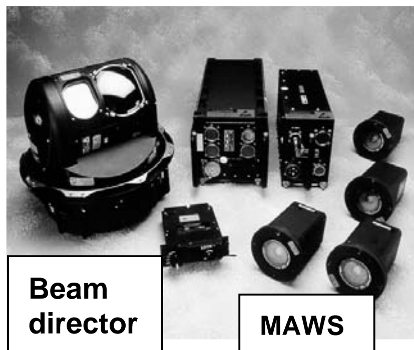


Figure 27. Small Nemesis countermeasure system [Courtesy Northrop Grumman]

The evolution of the DIRCM system technology to a laser-based system by supplementing the arc-lamp technology, or its replacement with laser sources, offers the potential of greater system flexibility and system functionality. These features may be exploited to provide enhanced system performance; hence providing effective protection system to more platforms.

An important systems integration aspect is the fact that a laser can be mounted remotely from the gimballed beam-director system. Consequently, the gimbal need only carry a stabilised mirror system [18] to direct the beam at the approaching target, leading to a miniature beam director. Care has to be taken to ensure that the laser energy is not back scattered into the receiver / target-tracking channel, and thus blinding the thermal camera of the fine-track sensor, during jamming.

One direct system's exploitation of a laser-based DIRCM is the use of the laser energy reflected from the approaching seeker. Once a target is being tracked by the system's beam director, this approaching threat may be interrogated by the countermeasure laser. The "interrogating" laser light is reflected by the seeker in the threat weapon; however, it is modulated by the reticle in the seeker in a

characteristic fashion by each class of threat seeker, hence analysis of the retro-return leads to identification of the approaching threat. An optimised jam waveform can then be used to create a rapid optical break lock (OBL) condition, using information gained from earlier exploitation of the threats.

This is the closed-loop IRCM (CLIRCM) technique [20]; it has been demonstrated by the LIFE and FLASH[5] systems. The retro-return may also be used to enhance the target-tracking performance of the beam director, a technique known as active tracking.

CLIRCM also offers another very important system's effectiveness function, as monitoring of the retro-signal enables effectiveness of the countermeasure technique to be assessed. If the retro-signal disappears then this is a powerful indication that the seeker is no longer looking at its intended victim. However, this assessment may also be made by simply monitoring the returned signal from the seeker as part of the active-tracking system now being investigated for some advanced laser-based DIRCM systems, without the system and processing complexity of a CLIRCM system.

The concept of an active-tracking system is quite simple. In this case the laser energy reflected from objects in the focal plane of the threat seeker is detected by the tracking camera in the laser-pointing system and used as the primary tracking point rather than the heat energy in the missile's plume. This means that the divergence of the laser beam does not need to be enlarged artificially; this requirement is to account for the fact that, with "plume tracking", the tracking point is to the rear of the missile, and so the laser energy is not directly pointed into the seeker. Reduction in the divergence of the laser beam by using the active-tracking technique enables the radiant intensity to be increased and hence offers the prospect of enhanced countermeasure performance. Moreover, this offers a route to countermeasures effectiveness assessment by monitoring the intensity of the laser return, as noted above. Disappearance of the laser return should indicate the loss of the sightline between the seeker and its victim, *i.e.* OBL.

6 Characteristics of Laser-based Jammers

On-board systems offer a number of significant operational advantages over the simpler off-board systems. The commonly cited advantages offered by on-board systems are:

- Compatibility with all types of military operations. Use of jammers complies with a large range of rules of engagement spanning hot war to peace-keeping operations.
- DIRCM systems are, in general, effective over the full spectrum of aircraft operational altitudes. In particular these systems are still effective when the aircraft is close to the ground, so provide protection during take-off and landing from forward operating bases.

[5] Flying Laser self-defence system Against Seeker-Head (missiles), a Franco-German technology demonstration project.

- Jammer systems are covert and so may be used during any operations, however, laser-based systems have a particularly low probability of intercept owing to the high directionality of the beam.
- Laser-based jammers are capable of a virtually unlimited number of engagements so they may be considered to have a very deep magazine.
- Careful design, and the inherent flexibility of this technology, means that laser-based jammers are capable of defeating the entire spectrum of heat-seeking missiles, including the emerging systems.
- The use of intense sources provides additional flexibility for jamming a seeker when the jam source is outside the field of view (FoV) of the seeker.
- The mode and fine control of the jam waveform means that jammers may operate in conjunction with flares to provide enhanced protection.
- Jammers offer a route to so-called role-fit systems that may be integrated with a pod for fitting to legacy platforms. Thus enhancing the survivability potential of all aircraft, including fast jets.
- Laser emissions may be tuned to any part of mid–IR atmospheric window to defeat the entire spectrum of heat-seeking missiles.
- Laser technology offers the only viable route to providing the very bright sources that are required to have the potential to defeat the emerging threats.

7 Future Developments of Laser-based Systems

The biggest challenge facing the laser-based jammer systems is to make the technology affordable. Many of the current devices are very complex and this leads to a high cost, as well as less than desired reliability, particularly when the devices are qualified to operate in a traditionally harsh military environment. Future developments, in this class of countermeasure system, are likely to pursue devices with reduced numbers of components, and more use made of techniques that will allow direct generation of IR photons. Other developments envisaged for laser-based jammer systems are:

- More compact systems
- Enhanced efficiency
- Greater tunability
- Enhanced functionality of the system, so that the entire countermeasures system has greater functionality
- Remote location of the laser source and having a single source for all laser-based functions on a platform

The investment in the American Homeland Security project to protect civil aircraft may revolutionise the cost and reliability of jammer systems. The requirements for operation on civil aircraft are somewhat different from those required for military aircraft.

The future developments of tactical laser-based defensive systems, using the "damage and destruction" method are likely to focus on the development of compact sources and the probable exploitation of fibre lasers [21]. The major challenges with this technology centre on achieving very high efficiency of the photon generation

process and the minimisation of the waste heat. Other requirements will concern the control of the aberrations occurring in the beam, from the point of generation to the point of contact with a designated target.

References

1. Herskovitz, D, Is seeing believing? Journal Electronic Defence 1998: **21,** 41-46
2. AIM 9 development, http://people.howstuffworks.com/sidewinder.htm.
3. Maiman, T H, Stimulated optical radiation in ruby. Nature 1960: **187,** 473
4. Hecht J, Laser weapons go solid-state, Laser Focus World 2004: **40,** No 9, 61-68
5. Hecht J, Optical systems can help stop missiles, Laser Focus World 2004: **40,** No 5: 94-100
6. Smith F G, Atmospheric propagation of radiation, ERIM and SPIE Optical Engineering Press, Ann Arbor, USA, 1993
7. Tennekes, H & J L Lumley, A first course in turbulence, The MIT Press, Cambridge Massachusetts, USA, 2001
8. Kolmogorov, A N, The local structure of turbulence, in incompressible viscous fluid for very large Reynolds' number, Doklady Akad Nauk USSR, Academy of Sciences USSR. New York, 1941: **30,** 301
9. Tyson, R K, Principles of Adaptive optics, Academic Press. Orlando USA, 1991
10. Titterton D H, The development of optical countermeasures, Proceedings SPIE Conference on Technology for optical countermeasures, London 2004, 5615, pp 1-15, 2005
11. Richardson M A, *et al.*, Surveillance and target acquisition systems, second edition, Brassey's Land Warfare **4**, London, UK, 1997
12. Tranchita, C J *et al.*, "Active infrared countermeasures" in the *Infrared and Electro-Optical Systems Handbook* (Ed D H Pollock), **7,** Chapter 3, ERIM and SPIE Optical Engineering Press, Ann Arbor, USA, 1993
13. Wood R M, Laser-induced damage of optical materials, IOP, Bristol, UK, 2003
14. Bartoli, F, L Esterowitz, R Allen, & M Kruer A generalized thermal model for laser damage in infrared detectors, Journal of Applied Physics 1975: **47,** No 7, 2875-2881
15. Nasburg R E, "Tracking and control systems" in the *Infrared and Electro-Optical Systems Handbook* (Ed M C Dudzik), **4,** Chapter 5, ERIM and SPIE Optical Engineering Press, Ann Arbor, USA, 1993
16. Hecht J, The Laser Guidebook, McGraw-Hill, Blue ridge Summit, Pennsylvania, USA, 1992
17. Wall, R, Threat Perception Casts Doubt on Laser Project. Aviation Week and Space Technology, 2002: **156.** No 1, 30-32
18. Titterton D H, & J L Weston Strapdown Inertial Navigation Technology (2nd edition - with applications), The IEE, Stevenage, UK, 2004.
19. Webb C E & J Jones, Handbook of lasers and applications, IOP, Bristol, UK, 2003.
20. Fulgham, D A, Laser can foil SAMs, Air-to-Air Missiles. Aviation week and Space Technology 2001, **154,** No 21: 43-45
21. Durvasula, L N, SPIE Conference Proceedings on Fiber lasers: Technology, Systems & Applications, San Jose 26 to 28 January 2004, **5335,** 2004

Survey of Thermophotovoltaic (TPV) Devices

Michael G. Mauk,
School of Engineering and Applied Science, University of Pennsylvania,
Philadelphia, Pennsylvania, USA.
mgmauk@yahoo.com

1 Introduction and Overview

Thermophotovoltaic (TPV) cells are solid-state *p-n* junction semiconductor devices that directly convert heat into electric power. More specifically, TPV cells absorb the predominantly infrared radiation emitted by a heated element and produce electric power by way of the photovoltaic effect. TPV cells are the critical component of various demonstrated and proposed systems for generating electricity from several types of heat sources including combustion processes, radio-isotopes, concentrated sunlight, process waste heat, and nuclear reactors. TPV systems are envisioned for use in applications ranging from small power supplies to replace batteries to utility-scale co-generation of electricity.

An efficient, practical, and cost-effective means for directly converting thermal energy to electrical energy has been an object of considerable research and development efforts for at least half a century. A scaleable device with neither moving parts nor working fluids that directly converts *heat* - or as in the case of TPV, *infrared radiation* from a heated source - to *electric power* is an appealing concept. Such a device or system would find use for remote, backup, and marine power generation as a more reliable replacement for diesel generators, especially in locations not suitable for photovoltaic solar arrays or storage batteries. Other potential applications include power plants for air or ground vehicle and ship propulsion, utilization of waste heat from industrial processes, and electric power supplies that would replace alternators in cars and trucks. Thermal sources for TPV systems include combustion of hydrocarbons such as gasoline or diesel fuels; propane or butane burners; highly focused sunlight; radioactive decay of radioisotopes such as plutonium-238; and heat exchangers coupled to nuclear reactors, furnaces, exhaust systems, fossil fuel-fired boilers, space heaters, or internal combustion engines. A specific example with some immediate commercial interest are "autonomous" gas- or oil-fired residential heating systems outfitted with TPV devices that convert heat to electricity for powering control circuits, relays, and blower fans, and which could thus operate during electric utility power outages [1]. More generally, there is interest in both miniaturizing TPV systems for powering small electronics systems, and scaling-up TPV for large-scale applications such as power plants, submarines, and buildings; as well as for energy recovery from the glass and paper making industries. In certain military applications, small TPV-based systems are under consideration for replacing battery power supplies for communication devices, laptop computers, and portable lighting. Other advantageous features may be noted: TPV systems are modular, quiet, and expected to be relatively safe, pollution-free, and low-maintenance [2]. In some situations, fuel-powered TPV systems can complement solar power systems for improved reliability and load management, and more favorable economics.

A common design of a TPV system employs an emitter material, such as SiC or a ceramic such as a rare earth oxide, maintained at an elevated temperature using one of the heat sources mentioned above. The emitter produces mostly mid-infrared radiation that impinges on a thermophotovoltaic cell -or more commonly, on an array of interconnected TPV cells -which through the photovoltaic effect generates a current and voltage that can be delivered to an electrical load. For high-performance TPV cells under typical operating conditions, the voltage generated by each TPV cell is very roughly 0.5 V and the current density is 1 to 3 A/cm^2 of cell area, thus implying a power density on the order of 1 W per cm^2 of active TPV array area. Various optical filters and reflectors, as well as emitter materials that absorb radiation at one wavelength and re-emit radiation at a different wavelength, can be incorporated into the TPV system to utilize better the thermal radiant energy. Since the characteristics of the radiation incident on the TPV cell are highly system-specific, it is difficult to quote standardized figures of merit or performance metrics akin to those that readily characterize solar photovoltaic devices, such as the "AM1.5 efficiency" which designates a standardized typical solar spectrum and insolation intensity as measurement conditions. Moreover, the system efficiency is dependent not only on the intrinsic conversion efficiency of the TPV cell, but also on the effectiveness of the combined TPV cell and system in utilizing photon recuperation effects that return any radiant energy unused by the cell back to the emitter or heat source. Specifically, photons which are neither absorbed by the TPV cell and thus do not contribute to the generated current, or photons that are not efficiently utilized by the TPV cell, *e.g.*, because their energy substantially exceeds the TPV cell bandgap, can be redirected to the radiator using reflectors and optical filters and thereby improve the overall thermal-to-electric conversion efficiency. Also, because the performance of semiconductor devices in general, and photovoltaic-type devices in particular, degrade with increasing device temperature, it is important to keep the TPV cell from becoming too hot. Therefore, heat transfer considerations, *e.g.*, avoiding parasitic absorption of thermal radiation and adequate heat-sinking, figure prominently in the design of TPV devices and systems.

While it is fair to say that TPV has not as yet found any substantial commercial or military applications, there has been significant progress in both TPV devices and systems in the last decade, and prospects are excellent for continued technical and cost improvements. Further, for at least several niche applications, TPV appears better suited than competing technologies such as thermoelectrics, thermionics, or fuel cells. TPV *system* efficiencies can be assessed on the basis of the quantity of electrical energy produced from either a unit of heat energy or, taking into account combustion efficiency, from a unit of fuel with a specified energy content. TPV system efficiencies in the 20–30% range appear feasible in the near-term, and modeling studies indicate system efficiencies of 30–40% may be possible [3]. The recent report [4] of a TPV system that demonstrated over 20% radiant heat conversion efficiency is a significant milestone and bodes well for continued and very likely increasing interest in this technology.

A TPV system can be configured so that a heated emitter directly irradiates the TPV cell, in which case the thermal radiation incident on the TPV cell has a spectral distribution characteristic of a blackbody at the emitter temperature, or more generally, a spectrum characteristic of the emitter material emissivity and its temperature. The emitter is typically heated to a temperature in the range of 1000–

2000 K. This emitter temperature range may be regarded as a compromise that is sufficiently high to produce desired radiant energy fluxes (~10 W/cm^2) needed for practical system sizes and cell power densities (~1 W/cm^2), but not too high so as to avoid the many technical problems encountered with operating a system at emitter temperatures in excess of 2000 K. These complications include thermal degradation and oxidation of components, NO_x generation, excessive demands on thermal insulation, difficulties in cooling the TPV cells, and safety. Emitter temperatures of around 1050 ± 50 °C appear to be of particular interest to several U.S. Government projects that fund much of the work in high-efficiency TPV technology. These emitter temperatures are best matched with TPV cells with spectral responses centered about the near-infrared and mid-infrared spectral ranges, thus necessitating the use of TPV devices made in low-bandgap (~0.5 eV) semi-conductors. However, as mentioned, some TPV systems employ spectral shifting techniques using wavelength-selective filters and reflectors, and/or incandescent selective emitters as re-radiators that modify the radiation spectrum utilized by TPV cells, thereby permitting more efficient conversion with wider-bandgap TPV cells such as those made in silicon.

From a device perspective, TPV cells are basically low-bandgap versions of *p-n* junction solar cells and photodiodes, designed and adapted to convert radiation in the mid-infrared wavelength range, in contrast to the insolation that spans the near-IR, visible, and UV wavelength ranges utilized by conventional solar cells. Most TPV cells are made in low-bandgap (0.25 to 0.7 eV) III-V compound semiconductors and their related ternary and quaternary alloys. Although these materials are not mainstays of commercial semiconductor technology, they are nevertheless well-established for mid-infrared optoelectronics technology that includes detectors, LEDs, and lasers. Development work on TPV cells in silicon, germanium, and several II-VI compounds and alloys continues, especially in applications where either low cell cost is an over-riding consideration and/or spectral shifting appears sufficiently workable. Still, research and development has focused mainly on a small subset of low-bandgap III-V compound binaries (GaSb, InAs) and their derivative ternary and quaternary III-V alloys, including InGaAs, InPAs, InAsP, InGaAsSb, InAsSbP, and related materials. For TPV cells based on ternary and quaternary alloys, the alloys are normally formed as epitaxial layers on GaSb, InP, GaAs, and InAs substrate wafers. The *p-n* junction is sited in the epitaxial alloy layer, the bandgap of which determines the spectral response and open-circuit voltage of the TPV cell. However, some TPV devices are made in wafers cut from bulk ternary III-V crystals, or in multicrystalline ingots or deposited polycrystalline films.

In the last decade, efforts to develop high-efficiency TPV cells have largely 'bifurcated' along two parallel paths based on two distinct materials systems: 1. TPV devices fabricated in InGaAs and InAsP epitaxial structures on InP substrates, which is represented by work at NREL (Golden, CO), NASA-Glenn Research Center (Cleveland, OH), Bechtel Bettis (Pittsburgh, PA), Ohio State University, their collaborators, and others; and 2. TPV devices fabricated in InGaAsSb epitaxial structures on GaSb substrates, which is represented by work at MIT Lincoln Labs (Boston), Sarnoff Labs (Princeton, NJ), Sandia National Laboratories (New Mexico), KAPL/Lockheed-Martin (Schenectady, NY), the Ioffe Institute (St. Petersburg, RU), and the Fraunhofer Institute (Freiberg, Germany), among others.

InGaAsSb is also referred to as GaInAsSb. Although the latter designation conforms to the usual rules of chemical nomenclature wherein the cation with the highest atomic number is listed first, both designations are commonly used.

The InGaAs/InP system exploits the availability of semi-insulating InP substrates, which facilitates the realization of monolithic interconnected module configurations for series-interconnected arrays. This feature enables high-voltage, low-current monolithic arrays which reduce resistive losses associated with high currents. Also, from a systems perspective, high-voltage, low-current modules are preferred over low-voltage, high current modules. This materials system has also been highly developed for optical fiber systems components. In earlier work on the InGaAs/InP system, the tolerable lattice-mismatch between the InGaAs epitaxial layer and InP susbstrate limited the bandgap range of InGaAs. InGaAs alloys grown on InP substrates and with bandgaps less than 0.6 eV suffered from compressive strain and defects. There has been considerable progress in this area in the last two years, yielding high-performance InGaAs TPV cells with bandgaps as low as 0.52 eV [5].

The InGaAsSb/GaSb material system provides for lattice-matched TPV cells down to a bandgap of about 0.5 eV (limited by a miscibility gap), thus offering a slightly better spectral match to emitters with operating temperatures in the 1000 to 2000 K range. Unfortunately, there are no semi-insulating GaSb (or other suitable lattice-matched insulating substrates) for this quaternary alloy, thus thwarting a simple approach to monolithic integrated modules analogous to those developed in the InGaAs/InP system. However, some new techniques using wafer bonding (see Section 6) can evidently circumvent this particular deficiency of the InGaAsSb materials system.

To facilitate more specific discussion of TPV device structures, Figure 1 shows cross-sections (not to scale) of several representative low-bandgap TPV devices. All of the devices shown are single-junction -as opposed to tandem/multi-junction-designs. The device specification details, *i.e.*, dopants, dopant concentration levels, metallizations, layer thicknesses, vary according to fabrication methods and intended applications. For most types of TPV cells, both *p-on-n* (*p*-type emitter; *n*-type base) and *n-on-p* (*n*-type emitter; *p*-type base) configurations are used, although in cells where the emitter is formed by diffusion, only *p-on-n* configurations are realizable due to the difficulty in diffusing *n*-type dopants in III-V semiconductors. (The term *emitter* for the top layer of a cell is borrowed from transistor terminology, where it refers to a relatively heavily-doped side of a *p-n* junction that injects minority carriers into the base. Here, the term *emitter* is potentially confusing as it must be distinguished from the heated *radiation emitter* component external to the TPV cell.) Otherwise, the optimum cell structure (configuration, layer thicknesses, and doping levels) is dictated by considerations regarding minority carrier diffusion lengths, emitter sheet resistivity, metallization contact resistances, feasible doping ranges with respect to dopant solubility and compensation, heavy doping effects such as bandgap narrowing and enhanced minority carrier recombination, and free-carrier absorption. Emitter thicknesses in *n-on-p* cells are typically are in the range of 0.2 to 1 μm, with base thicknesses of 2 to 3 mm, while *p-on-n* cells generally use thicker (1 to 3 μm) emitters. This is dictated in part by design rules stipulating that the minority carrier diffusion length should be at least twice the emitter thickness.

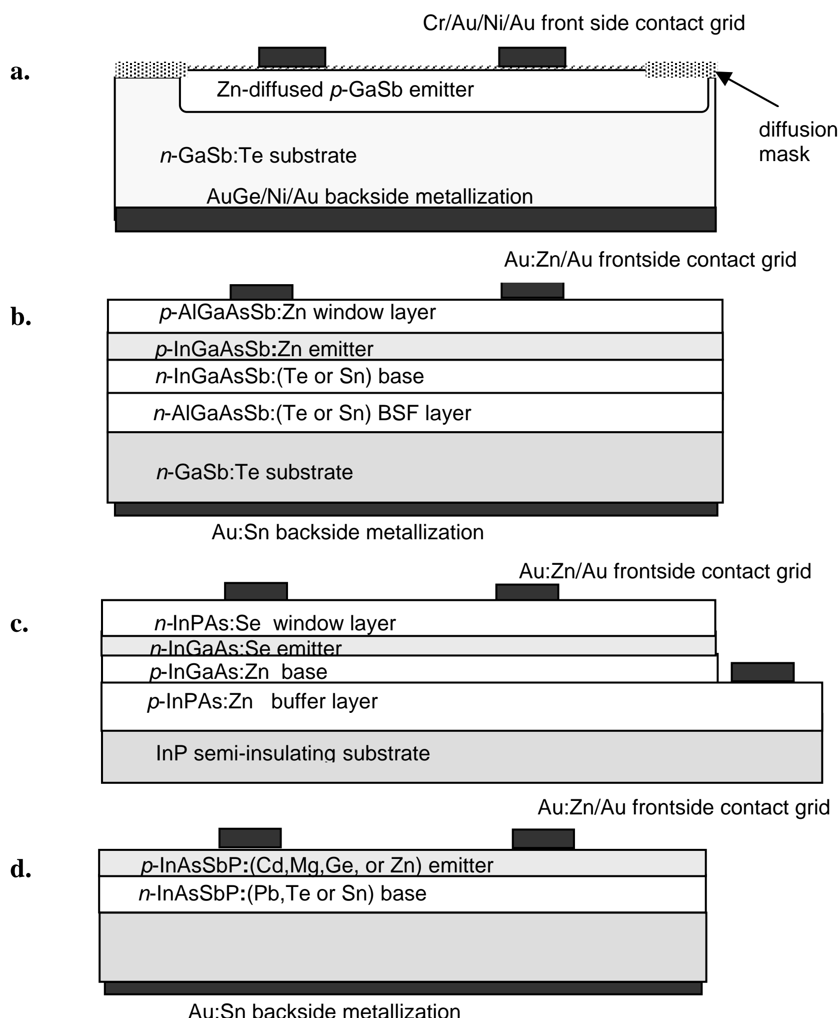


Figure 1. TPV cell structures: **a)** Zn-diffused GaSb TPV cell. (A more sophisticated version of this cell includes an epitaxial emitter in place of the Zn-diffused junction, and window and BSF layers.) **b)** Epitaxial AlGaAsSb/InGaAsSb/GaSb double heterostructure TPV cell, **c)** Epitaxial *p-n* InGaAs/InAsP/InP TPV cell. This cell can be made on a con-ductive InP susbtrate with a backside contact metallization, or on a semi-insulating InP substrate, in which case both emitter and base contacts are made from the front side to form a series-connected mini-module (details not shown), along with appropriate isolation of separate monolithic TPV cells (also not shown). **d)** InAsSbP/InAs *p-n* junction TPV cell. Both epitaxial emitters and Zn-diffused junctions have been used with InAsSbP TPV cells.

There are also design trade-offs related to the dopant-dependent minority carrier diffusion length in the emitter; the emitter sheet resistance contribution to the series resistance of the solar cell; heavy doping effects exacerbating minority carrier recombination and bandgap narrowing, both of which increase the diode dark

current and lower open-circuit voltage; as well as free-carrier absorption due to high doping of the emitter that diminishes the spectral response and photon recuperation capabilities of the TPV cell. These effects can be accurately modeled according to the analysis described in Section 3 and the Appendix.

Figure 1a shows a relatively simple TPV device fabricated by Zn-diffusion in an *n*-type GaSb wafer. This type of TPV cell is used in at least several current or near-term commercial products, usually in combination with a propane heater or furnace (see Section 6). Figure 1b,c shows heterostructure TPV cells that include a wide-bandgap window cladding layer and a wide-bandgap BSF/buffer cladding layer. The cladding layers effect minority carrier confinement to reduce minority carrier surface recombination effects. Minority carrier confinement at an isotype heterojunction requires the cladding layer have a bandgap that is about 0.2 eV higher than the bandgap of the layer in which the minority carriers are confined. Figure 1d shows an InAsSbP/InAs TPV device which is of interest for applications requiring a low-bandgap (< 0.5 eV), long-wavelength (2-4 μm) spectral response cell. Device fabrication techniques may include zinc diffusion; photolithography; reactive-ion, wet chemical or anodic etching; and thin-film deposition of metals, anti-reflection coatings, and passivating layers. Most of these processes are adapted from III-V compound solar cell technology. Epitaxial growth methods include metal organic chemical vapor deposition (MOCVD), molecular beam epitaxy (MBE), and liquid-phase epitaxy (LPE). As the epitaxy and device fabrication methods applied to TPV devices are fairly standard and similar to those used in compound semiconductor optoelectronics and solar cells, this review will focus instead on device structures and design methodologies, especially with regard to aspects unique to TPV technology.

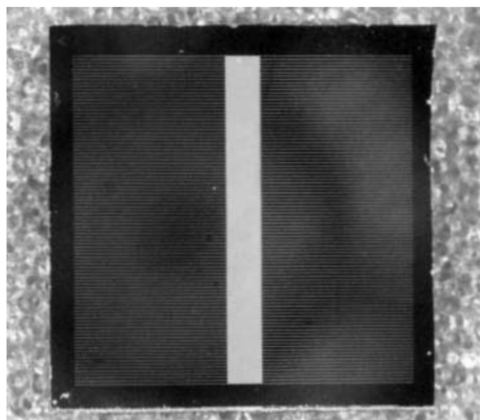


Figure 2. 1 cm × 1 cm GaInAsSb TPV Cell. TPV cell. Center busbar is 1-mm wide. The grid fingers are 10 μm wide and spaced 100 μm apart.

Figure 2 shows a top-view photograph of a 1-cm×1-cm InGaAsSb/GaSb cell similar to that of Figure 1b. The front grid shading is about 20%, indicative of the large metallization coverage needed to accommodate the high currents typical of TPV cell operation. Here again, TPV device designs can profitably adapt features developed

for concentrator solar cells that operate at 100 to 1000 suns intensity and which produce current densities much in excess of 1 A/cm^2.

In addition to the basic single-junction designs shown in Figure 1, more sophisticated tandem or multi-junction devices, as well as monolithic series-interconnected TPV arrays, are also under development. New concepts in TPV technology include micron-gap TPV cell configurations, various thermophotonic schemes, and spectral control using photonic crystals and microstructured surfaces. While these developments are somewhat generic and not specific to any particular TPV cell type or material, their utilization would no doubt alter the material selection with respect to optimum bandgaps, as well as the design of the cell with respect to front surface topography, minimizing parasitic absorption and maximizing photon recuperation, and tailoring the TPV cell spectral response.

2 TPV Literature and Other Sources of Information

The subject matter of thermophotovoltaics is highly diverse and multidisciplinary, encompassing semiconductor materials and device physics, and in particular low-bandgap (0.2–0.75 eV) III-V compound semiconductor materials and devices; TPV cell design; device packaging; conductive and convective heat transfer and radiative transport modeling; corrosion, oxidation, and materials reliability for high-temperature operation; mid-infrared spectral filters and reflectors; and topics incidental to ceramic, semiconductor, metal, and photonic crystal emitters. TPV progress has been archived in six volumes of the *Thermophotovoltiac Generation of Electricity* conference proceedings (vol. 321 [1994], vol. 358 [1995], vol. 401 [1997], vol. 460 [1998], 653 [2002], and vol. 738 [2004]) published by the American Institute of Physics (Melville, NY). Future conferences and additional volumes are expected approximately every two-years. Also, the three major photovoltaics conferences (*IEEE Photovoltaics Specialists Conference*, *European Photovoltaic Solar Energy Conference,* and the (Asian) *Photovoltaic Science and Engineering Conference*), each held at eighteen month intervals, normally include a section devoted to or including thermophotovoltaics. Coutts [2,6] has provided two comprehensive reviews of thermophotovoltaics technology, which give a detailed description of materials, devices, systems, and engineering issues. Andreev [7] has presented a briefer but more recent review of several TPV technologies. In addition, Coutts and Fitzgerald authored a general interest article on TPV in *Scientific American* [Sept. 1998]. *Semiconductor Science and Technology* published a special issue devoted to thermophotovoltaics in May of 2003. Reports of thermo-photovoltaics research and development appear irregularly in *IEEE Transactions on Electron Devices, Journal of Applied Physics, Applied Physics Letters,* and *Progress in Photovoltaics Research and Applications*. The latter journal also published a biography of thermophotovoltaics listing 200 citations covering the period from 1960 to 1995 [8]. There is also relevant work reported at the MIOMD (*Mid-Infrared Optoelectronics Materials and Devices*) conference series, much of which is published in special issues of *IEE Optoelectronics*. Approximately 100 patents related to TPV were found by a search of the U.S. Patent and Trademark Office Database. Work on space and concentrator solar cells, in addition to other mid-infrared optoelectronics such as detectors, LEDs, and lasers, as well as so-called

"third-generation" photovoltaic and thermophotonic concepts, are often directly or indirectly relevant to TPV technology.

3 Basic Operation of TPV Cells

As with conventional solar cells, the performance of a TPV cell can be analyzed in terms of open-circuit voltage V_{oc} (*i.e.*, the voltage produced by the TPV cell under a specified illumination with no electrical load ($R_L \rightarrow \infty$)); short-circuit current density J_{sc} (*i.e.*, the current produced by the TPV cell, per unit area of the cell, under the same specified illumination when the load is short-circuited ($R_L \rightarrow 0$); and fill-factor *FF* which indicates the operating point (J_{max}, V_{max}) on the current-voltage characteristic of the TPV cell that yields maximum electrical power P_{max}.
The efficiency η is thus

$$\eta = \frac{P_{max}}{P_{in}} = \frac{V_{max} J_{max}}{P_{in}} = \frac{FF \cdot V_{OC} J_{SC}}{P_{in}} \tag{1}$$

where P_{in} is the power density per unit area of the radiation incident upon the TPV cell. The incident power density P_{in} is found by integrating the spectral density of the incident radiation over all wavelengths. The spectral power density can be expressed as $P_E(E)$ [power/unit-area/photon-energy] in terms of the photon energy *E,* or equivalently as $P_\lambda(\lambda)$ [power/unit area/photon-wavelength] in terms of photon wavelength λ, where photon energy and wavelength are related by $E = hc/\lambda$. Thus,

$$P_{in} = \int_0^\infty P_E(E)dE = \int_0^\infty P_\lambda(\lambda)d\lambda \tag{2}$$

Alternatively, the incident power can be found by integrating the spectral photon flux $\phi(\lambda)$ [photons/unit-area/unit-time/photon-wavelength] over all wavelengths, *i.e.*,

$$P_{in} = \int_0^\infty \frac{hc}{\lambda} \phi(\lambda) d\lambda \tag{3}$$

The short-circuit current density J_{SC} is then given by

$$J_{SC} = \int_{E_G}^\infty P_E(E)\ SR_E(E)dE \tag{4a}$$

$$= \int_0^{\frac{hc}{E_G}} P_\lambda(\lambda)\ SR_\lambda(\lambda)d\lambda \tag{4b}$$

where $SR_E(E)$ (or $SR_\lambda(\lambda)$) denotes the spectral response [amps/watt-incident radiation] as a function of photon energy *E* (or photon wavelength λ). Equivalently, J_{sc} can be calculated from the photon flux density $\phi(\lambda)$ of the incident radiation and the external quantum efficiency $QE_{ext}(\lambda)$ [dimensionless], as

$$J_{SC} = q \int_0^{\frac{hc}{E_G}} \phi(\lambda)\, QE_{ext}(\lambda)\, d\lambda \tag{5}$$

where q is the electronic charge.

The external quantum efficiency is readily interpreted as the probability that an incident photon of wavelength λ will generate a charge carrier that contributes to the short-circuit current of the TPV cell. The external quantum efficiency is a measure of the reflection and absorption of incident photons, the generation of minority carriers, and the collection of minority carriers by the *p-n* junction. Thus, for example, the external quantum efficiency accounts for the loss due to reflection of incident photons from the front surface of the cell, the loss resulting from sub-bandgap and low-energy photons that are not absorbed, and losses resulting from photogenerated minority carriers that recombine before they are collected by the *p-n* junction.

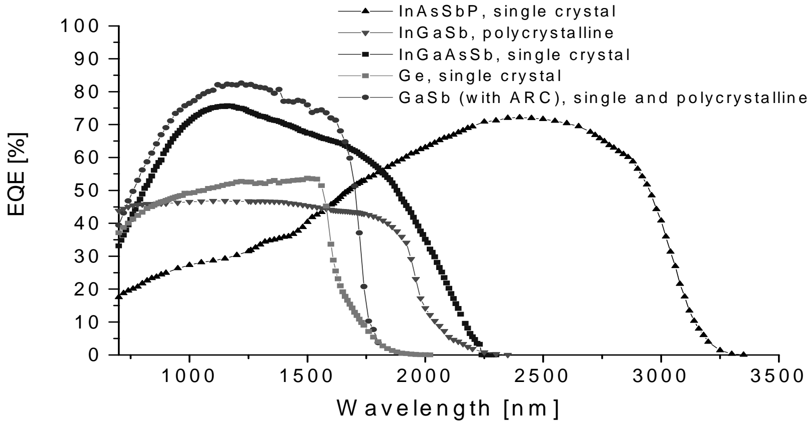


Figure 3. Measured external quantum efficiency spectrum of several types of TPV cells. Antireflection coating (anodic oxide) was applied only to the GaSb TPV cell. Shadowing losses caused by the contact grid were in the 15-20% range. (Figure courtesy of Dr. Oleg Sulima, GE Energy, Newark, Delaware USA).

Figure 3 shows external quantum efficiencies measured for several types of TPV cells made in semiconductors covering a bandgap range of 0.35 to 0.7 eV. The devices include TPV cells made as diffused junctions in monocrystalline and polycrystalline GaSb and monocrystalline Ge wafers, diffused junctions in wafers cut from polycrystalline InGaSb ingots, epitaxial junctions in GaInAsSb heterostructures made on single-crystal GaSb substrate wafers, and diffused junctions in InAsSbP epitaxial layers grown on monocrystalline InAs substrate wafers. The TPV cells shown represent a broad selection of device types in both sophistication and cost. The data show that the current stock of materials in which TPV cells can be made covers a wide spectral range and that the system designer has considerable latitude in selecting a cell type based on spectral response

considerations. For mature TPV technologies such as GaSb, InGaAsSb, and InGaAs, external quantum efficiencies of 50 to 80% are achievable over an appreciable spectral range. In this regard, it is noted that the front contact metallization grid obscures about 10 to 20% of the front surface, imposing a limit on the external quantum efficiency even with a perfect anti-reflection coating and no minority carrier recombination losses. If the grids are made reflective, it may be possible to ameliorate this 'shading' loss at the system level, since reflected photons could be returned to the radiator. This feature has been suggested but it is not clear to what extent it has been realized in actual cells or exploited in TPV systems.

The open-circuit voltage V_{oc} is related to the reverse-saturation current J_0 of the *p-n* junction and the short-circuit current J_0 by way of the diode equation

$$J = J_0 \left[\exp\left(\frac{qV}{mkT_C} \right) - 1 \right] - J_{SC} \tag{6}$$

where m (also often denoted as n) is the diode ideality factor, ideally equal to 1, k is Boltzmann's constant, and T_C is the absolute temperature of the TPV cell. J_0 is also often referred to as the 'dark' current of the cell.

Re-arranging the above equation and solving for $V_{OC} = V$ when $J = 0$, yields

$$V_{OC} = \frac{mkT_C}{q} \ln\left[1 + \frac{J_{SC}}{J_0} \right] \tag{7}$$

For purposes of predicting performance, determining the best bandgap for a particular situation, and assessing the potential of various TPV materials, a generalized formula for J_0 has been developed for solar cell modeling - see for example, Fan *et al.* [9].

$$J_0 = \beta(E_G) \cdot T_C^{\,3} \exp\left(-\frac{E_G}{kT_C} \right) \tag{8}$$

Nell and Barnett [10] fit the above equation to reported values for several solar cells of different bandgaps, arriving at a 'best-fit' constant value of $\beta(E_G) = C =$ 17.9 mA/cm^2/K^3. Wanlass *et al.*. [11] suggest $\beta(E_G) = 3.165 \times 10^{-4} \cdot \exp(2.19 \cdot E_G)$ A cm^{-2}K^{-3} with E_G measured in eV.

More relevant to TPV devices is the compilation by Charache [12] of experimental J_0 values in low-bandgap semiconductor *p-n* junctions shown in Figure 4. J_0 decreases with increasing minority carrier lifetime, but there is a well-known theoretical limit to minority carrier lifetime imposed by radiative recombination phenomena. This limit depends on whether photon recycling effects are operative, as can be effected and enhanced by a backside reflector and other design features. Photon recycling refers to the re-absorption of photons produced by radiative recombination of minority carriers, wherein the re-absorbed photon generates a minority carrier. Photon recycling can increase the effective minority carrier lifetime by about an order of magnitude [13,14].

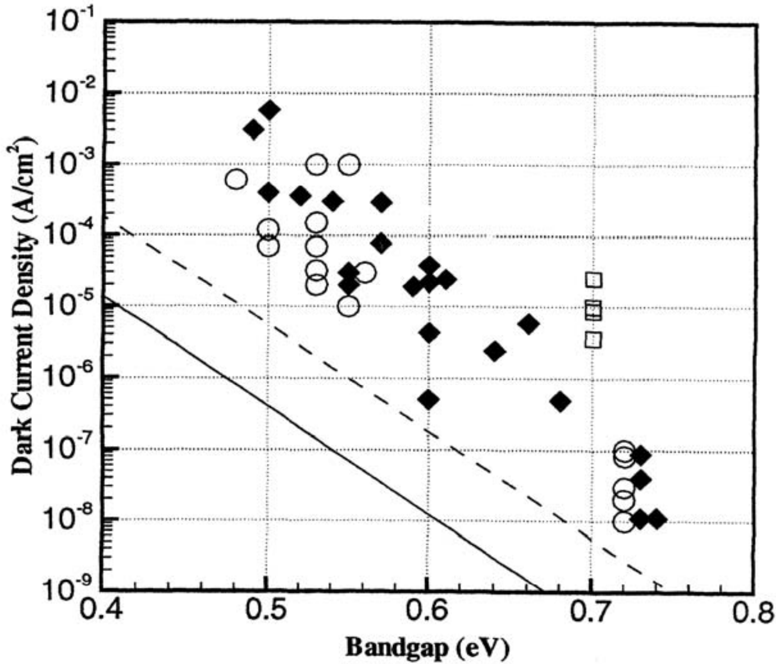


Figure 4. Dark current density J_0 as a function of bandgap E_g for thermophotovoltaic or photovoltaic devices. Experimental results, Key: □ germanium, ○ InGaAsSb and GaSb, ◇ InGaAs, Theoretical modeling predictions, Key: —— radiative recombination minority carrier lifetime limit *with* photon recycling effects, ------- minority carrier lifetime limit *without* photon recycling effects. From Charache [12].

Since open-circuit voltage is largely determined by J_0, it is useful to compare predicted values of J_0 as a function of bandgap E_G and measured values of J_0 in actual devices. As discussed shortly, some modeling results indicate that higher TPV system efficiencies could be achieved using TPV cells with bandgaps less than 0.5eV, which is the lower bound of the mainstream TPV cell technologies based on InGaAsSb and InGaAs. Moreover, tandem TPV cells and so-called microgap TPVcell configurations would benefit from TPV cells with lower bandgaps than those of current interest. The InAsSbP alloy system is a good candidate for TPV cells with bandgaps in 0.3 to 0.5 eV range (see Figure 3), thus extending the spectral response out to wavelengths of around 3.5 μm. Pushing TPV cell technology to lower bandgaps may be hindered by severe Auger recombination effects that become increasingly prevalent in low-bandgap semiconductor devices. Therefore, to what extent the trend indicated in Figure 4 can be extrapolated to low-bandgap (< 0.5 eV) devices remains to be seen. Further, compared to InGaAs and InGaAsSb these low-bandgap materials are much less studied and have less developed technology bases. Moreover, doping of low-bandgap semiconductors can be problematic due to intrinsic doping effects and other electrically-active defects.

In addition to J_{SC} and V_{OC}, the cell efficiency depends on the fill factor FF. There are no simple analytical formulae for estimating the fill-factor. One useful. approximation is

$$FF \cong \frac{v - \ln(v + 0.72)}{v + 1} \quad (9)$$

where

$$v = \frac{qV_{OC}}{kT_C} \quad (10)$$

and T_C is the cell temperature. The fill factor is determined by the diode ideality factor m, shunt and leakage currents, and series resistance effects. For mature TPV technologies, fill-factors in the range of 60–70% are common. This is somewhat lower than those achieved in solar cells and this may be attributed in general to the relative non-ideality and leaky junction behavior of low-bandgap diodes, and also to the comparatively high current levels encountered in TPV devices which increase series resistance effects.

4 TPV System Thermodynamic Limits and Modeling

It is instructive to compare the energy flows in both solar cells and TPV cells in order to appreciate the basic operation of the device, design issues and various loss mechanisms. Further, it is informative to contrast the typical implementation of a TPV cell with that of a solar cell, especially in view of the natural inclination to apply solar cell design paradigms to TPV devices.

In Figure 5, the energy flows in a solar cell and two variants of a TPV system are shown. In Figure 5a, broad-spectrum incident solar radiation (including visible, infrared, and ultraviolet components) illuminates the solar cell. The solar spectrum can be well approximated as a blackbody emitter at 6000 K. The total intensity is nominally 100 mW/cm^2. A certain fraction of this radiation is lost due to front surface reflection, although this can be ameliorated by antireflection coatings and surface texturing to reduce total reflection losses to a few percent or less. Some incident photons are lost due to the obscuring or shading effects of the metal grid lines formed on the front surface of the solar cell, amounting to additional approximate 5% loss. More fundamentally, longer-wavelength photons with energies less than the semiconductor bandgap energy are not absorbed, and thus do not contribute to the photovoltaic effect. For a radiation emitter with temperature T_E, the fraction of incident power comprising sub-bandgap energy photons and that cannot be converted by the solar cell is given by

$$F_{\text{sub-}E_g\ \text{loss}} = \frac{\int_0^{E_G} \left\{E^3 / [\exp(E / kT_E) - 1]\right\} dE}{\int_0^{\infty} \left\{E^3 / [\exp(E / kT_E) - 1]\right\} dE} \quad (11)$$

Photons with energies in excess of the bandgap are absorbed and generate minority carriers. These photo-excited minority carriers quickly thermalize to energies within

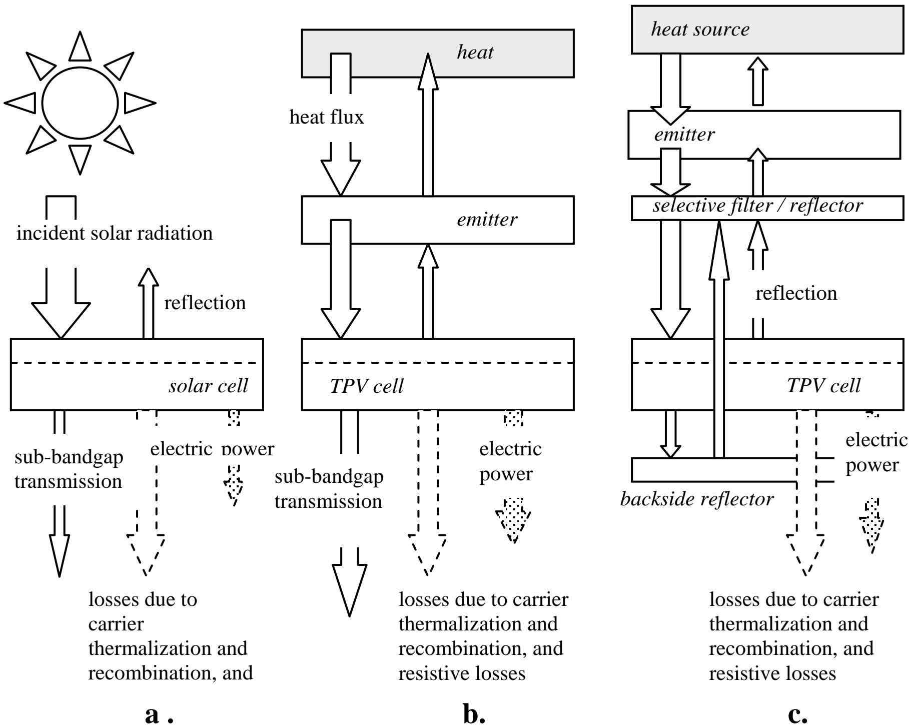


Figure 5. a) Solar flux is incident upon the solar cell, some fraction of which is reflected. Sub-bandgap light does not contribute to the photovoltaic effect, and is either absorbed uselessly in the back contact or reflected back through the front surface. A large portion of the incident energy is lost due to thermalization of carriers created by photons with energies exceeding the energy gap. **b)** A TPV cell presents a somewhat different situation with respect to energy flows. In principle, light reflected from the front surface can be returned to the emitter. Thus, reflection need not be a loss to system efficiency. In a similar regard, and as indicated in **c**), the sub-bandgap radiation can also be returned to the emitter. This is not feasible with solar cells. The efficient recovery of sub-bandgap radiation requires that parasitic (*e.g.*, free-carrier) optical absorption losses be avoided.

$\sim kT_c$ of the conduction and valence band edges. The fraction of incident power that is lost due to thermalization of photo-excited carriers is given by

$$F_{>E_g \text{ thermalization loss}} = \frac{\int_{E_G}^{\infty} \left\{ E^2 (E - E_G) / [\exp(E / kT_E) - 1] \right\} dE}{\int_0^{\infty} \left\{ E^3 / [\exp(E / kT_E) - 1] \right\} dE} \qquad (12)$$

The optimum bandgap of the solar cell is that which minimizes the sum of these two loss fractions.

These two loss mechanisms, *i.e.*, losses associated with sub-bandgap photons and losses due to thermalization of photo-generated minority carrier energy in excess of the bandgap, are practically unavoidable in a conventional solar cell, but can be addressed in TPV system design as will be described. For this reason alone,

single-junction solar cells have efficiency limits in the 30-35% range. The photogenerated minority carriers diffuse or drift to the *p-n* junction where they are collected, cross the *p-n* junction, and are thereby converted to majority carriers that comprise the light-generated (short-circuit) current of the solar cell. There are additional consequent losses due to surface and bulk minority carrier recombination, as well as series and shunt resistance.

Figure 5b shows a particular implementation of a TPV cell in a simple TPV system. An emitter is heated by a thermal source to a temperature between 1000 to 2000 K. The radiation has a characteristic emission spectrum which may either approximate a blackbody, or in the case of selective emitters such as rare earth oxides, may exhibit a relatively narrow spectral emission width. Some of the emitted radiation impinges the TPV cell and is absorbed to generate minority carriers. The energy reflected from the TPV cell can be re-directed to the emitter to help maintain its temperature. Thus, in marked contrast to a solar cell where reflection is an irredeemable loss, reflection of photons in a TPV system is not necessarily a loss in overall system efficiency. Further, if the spectral characteristics of the anti-reflection are tailored to reflect high-energy photons, the losses associated with thermalization of carriers generated by photons with energies in excess of the bandgap energy can be reduced. In the arrangement shown in Figure 5b, sub-bandgap photons still constitute a significant loss, similar to that encountered with solar cells.

Figure 5c shows a TPV system that includes a selective filter or reflector between the emitter and TPV cell, and also a reflector on the backside of the TPV cell. The selective filter/reflector will ideally redirect high-energy photons back to the emitter in order to reduce thermalization losses as discussed above. In addition, a backside reflector can redirect sub-bandgap photons to the emitter. Alternatively, the sub-bandgap photons can be returned to the emitter by siting a reflector between the emitter and cell that exhibits a high reflectivity for photons with energies below that of the TPV cell bandgap. The return of photons to the emitter by the use of selective filters, backside and frontside reflectors, and reflective cell coatings, called *photon recuperation*, provides for efficiency enhancements that are not possible in solar cell systems, and the design and optimization of TPV cells used in systems that incorporate photon recuperation should seek to fully exploit these effects.

Baldasaro *et al.* [16] have developed a formulation of TPV device performance limits for a system that employs radiative sources characterized as blackbody (*i.e.*, Planck's radiation Law) emitters. The TPV efficiency η is the product of several factors:

$$\eta = \overline{QE}\ FF\ \left(\frac{qV_{OC}}{E_G}\right) F_0 \tag{13}$$

where $\overline{QE}$ is the external quantum efficiency, weighted-averaged over the spectral bandwidth of radiation incident on the TPV cell. FF is the fill factor. qV_{OC} / E_G is sometimes referred to as the *voltage factor* and is a measure of the degree to which the device utilizes the potential energy barrier of the *p-n* junction, which is ultimately related to the bandgap of the semiconductor, to produce a working external cell voltage. F_0 is the *photon over-excitation factor* and quantifies the

thermalization losses associated with photons above the bandgap. For TPV systems with perfect photon recuperation of sub-bandgap photons,

$$F_0 = \frac{E_G \int_{E_G}^{\infty} E^2 \left[\exp\left(E / kT_E\right) - 1\right]^{-1}}{\int_{E_G}^{\infty} E^3 \left[\exp\left(E / kT_E\right) - 1\right]^{-1}} \tag{14}$$

If there is no recuperation of sub-bandgap photons, a more general factor F_0' can be used to estimate the efficiency

$$F_0' = \frac{E_G \int_{E_G}^{\infty} E^2 \left[\exp\left(E / k \cdot T_E\right) - 1\right]^{-1}}{\int_{0}^{\infty} E^3 \left[\exp\left(E / k \cdot T_E\right) - 1\right]^{-1}} \tag{15}$$

The above analysis is highly idealized in that it ignores some unavoidable additional losses inherent in photovoltaic and TPV devices. The most fundamental limit to TPV performance is imposed by radiative recombination. Baldasaro *et al..* [15] indicate that the minimum dark current $J_0 > J_{rad}$, as stipulated by considering the radiative recombination of minority carriers, is given by

$$J_{rad} = \frac{q(n^2+1)}{4\pi^2 c^2} \int_{E_G/\hbar}^{\infty} \left[\omega^2 \exp\left(\frac{\hbar\omega}{kT_C}\right) - 1\right]^{-1} d\omega \tag{16a}$$

$$\cong \frac{q(n^2+1)}{4\pi^2 c^2 \hbar^3} \left[E_G{}^2 kT_C - 2\left(kT_C\right)^3 \left(\frac{E_G}{kT_C} - 1\right)\right] \exp\left(\frac{-E_G}{kT_C}\right) \tag{16b}$$

$$\cong \frac{q\left(n^2+1\right) E_G^2 kT_C}{4\pi^2 c^2 \hbar^3} \cdot \exp\left(-E_G / kT_C\right) \tag{16c}$$

where the symbols have their usual meaning, and n is the refractive index of the semiconductor. In a like fashion, the short-circuit current of a TPV cell illuminated by an emitter with temperature T_E is found to be

$$J_{SC} = \frac{q\left(n^2+1\right)}{4\pi^2 c^2 \hbar^3} T_E \exp\left(-E_G / k_B T_E\right) \tag{17}$$

Combining these equations gives a radiative limit for the “voltage factor”:

$$\text{voltage factor} \equiv \frac{V_{OC}}{E_G} = 1 - \frac{T_C}{T_E} + \frac{kT_C}{E_G} \ln\left(\frac{T_E}{T_C}\right) \tag{18}$$

For InGaAsSb TPV cells with $E_G \cong$ 0.5 eV, deployed in a system with a radiator (emitter) temperature of 1500 K, and with the TPV cell temperature T_c maintained at 300 K, the voltage factor is about 0.88, corresponding to a V_{OC} of about 440 mV. As the best-demonstrated InGaAsSb TPV cells in this bandgap range achieve open-circuit voltages of 300–350 mV, there are evidently other factors

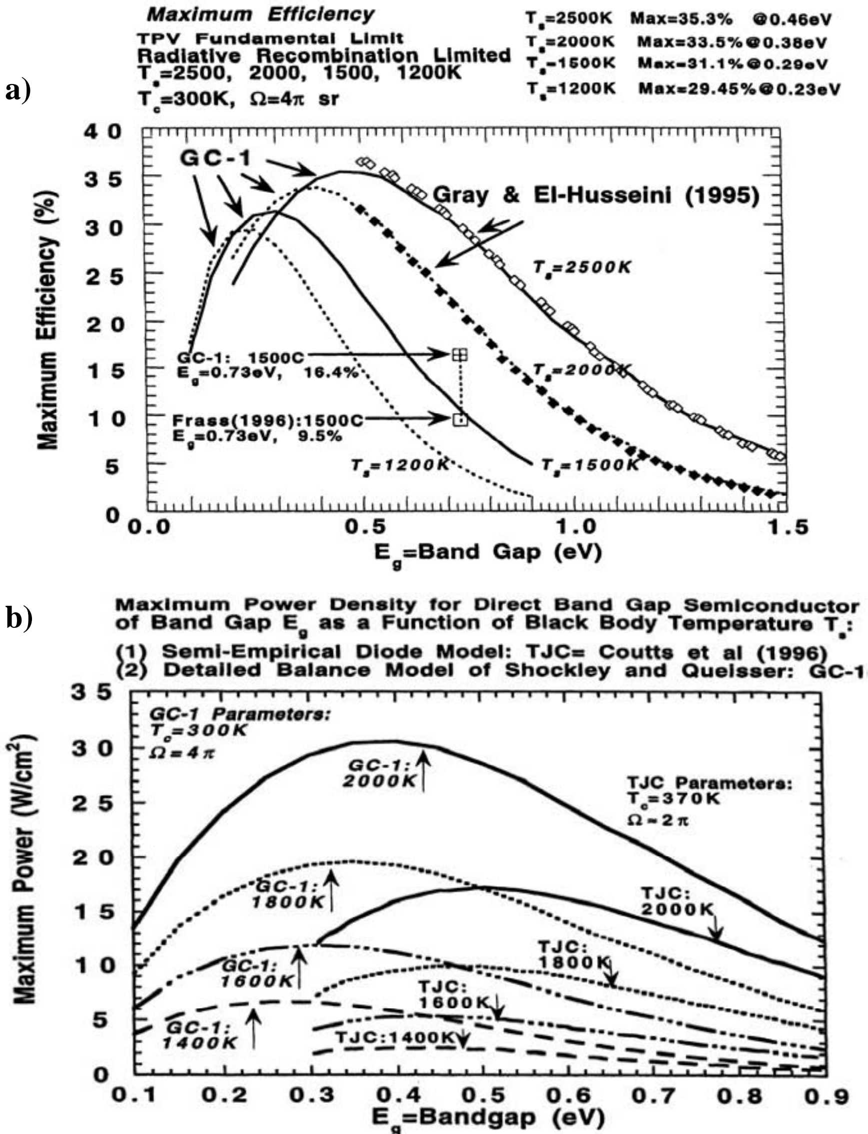


Figure 6. **(a)** Maximum efficiency as function of emitter temperature and TPV cell bandgap. T_S is the thermal source (emitter) temperature, T_C is the TPV cell temperature. GC-1 and Gray and El-Husseini (1995) refer to modeling results compiled by Cody [21], plus experimental device data from Fraas *et al.* (1996). Figure from Cody [21]. **(b)** Maximum power density as a function of thermal emitter temperature and TPV cell bandgap. GC-1 and TJC refer to modeling results compiled by Cody [22] including results from T.J. Coutts [6]. T_S = 370 K is the TPV cell temperature. Figure from Cody [22].

limiting the open-circuit voltage of such TPV cells. This situation is fairly typical of TPV cells in the bandgap range of 0.5–0.7 eV, and in fact, the voltage factor becomes proportionately worse for TPV cells with bandgaps < 0.5 eV. This latter phenomenon is often attributed to Auger recombination. In general, since the $\overline{QE}$ and the FF are already close to their theoretical limits, there is most room for improving overall TPV cell efficiency by means that will enhance the voltage factor.

One objective of modeling TPV devices and systems in a generalized way (independent of extrinsic materials properties) as described above is to determine the optimum bandgap for an emitter operating at a specified temperature or a specified emission spectrum. Modeling is also used to estimate the performance potential for TPV cells as a function of bandgap and radiation characteristics. Such models have been developed by Kittl [3], Woolf [17], Wojtczuk *et al.* [18], Gray and El-Huseini [19], Cody [20,21], Coutts and Ward [22], and Coutts [6], mainly for the purpose of finding optimum bandgaps of TPV cells for maximizing system efficiencies. Some of these results are summarized in Figure 6a, which shows the maximum efficiency assuming a radiative recombination limit for the cell dark current. The maximum power as a function of bandgap and emitter temperature is shown in Figure 6b. Figure 6a,b is based on a black body spectrum, and assume either a semi-empirical diode model [23] or a Shockley-Queisser detailed balance model [21,24]. Subsequent modeling results by Coutts and Ward [22] further corroborate these trends.

5 TPV Cell Modeling

As an illustrative and representative case study, the performance of a double-heterostructure *p-n* homojunction AlGaAsSb/GaInAsSn/GaSb TPV device is analyzed in some detail. This structure is sufficiently general in that its analysis can be readily applied to simpler TPV devices as special cases. The analysis follows that of Hovel [25] and Hutchby and Fudurich [26] as first applied to solar cells, and similar adaptations specific to TPV cells [27-33]. The model is based on closed-form solutions of one-dimensional minority carrier transport equations incorporating optical generation, and bulk and surface recombination of minority carrier in quasi-neutral regions and the space-charge region surrounding the *p-n* junction. The closed-form solutions of the current equations can be programmed on a computer to simulate the current-voltage characteristics and spectral response of the TPV cell. Such models permit optimization of the device structure (*i.e.*, alloy compositions, layer thicknesses and doping levels), provide a sensitivity analysis of performance to variations in materials properties and design parameters, and offer an assessment of practical performance potential for specific designs using realistic materials properties. Further, the simulated spectral response can be compared with measured spectral responses in order to identify losses and estimate minority carrier recombination lifetimes. In this last regard, such models enable solar cells, detectors, and TPV devices to serve as excellent diagnostic tools for evaluating material quality. The next section defines the TPV device geometry. The analysis of the various current components used to describe TPV performance, along with models of relevant materials properties as a function of alloy composition, are included in the Appendix. While it is true computer simulations of device performance based on numerical solution of more rigorous device physics models, including two- and

three-dimensional models, provide better accuracy, the one-dimensional closed-form solutions are nevertheless very useful and provide considerable insight into workings of the TPV device. TPV device performance is most sensitive to minority carrier recombination lifetimes, which can vary by several or more orders of magnitude depending on the nature and concentration of dopants and other impurities and contaminants, defects, bandgap narrowing effects, electric fields, and stress of the constituent semiconductors. TPV device performance is comparatively less sensitive to variations in mobility, optical absorption, bandgap narrowing effects on carrier concentration, and resistivity. Thus, it is informative to use such a model to study device performance as a function of minority carrier lifetime treated as an adjustable parameter. Of particular interest are limits to the minority carrier lifetime, due to Auger and radiative recombination, and the corresponding maximum open-circuit voltage attainable for such lifetimes, as this is mainly where TPV cell performance falls short of its theoretical potential.

5.1 TPV Device Structure and Delineation

Figure 7 shows a generic double heterostructure in which a *p-n* junction, comprised of oppositely-doped emitter and base layers, is sandwiched between two wide-bandgap cladding layers. The rear cladding layer, denoted as a *buffer*, is used to create a back surface field (BSF) effect by way of a combination high-low isotype doping junction and/or confinement due to a minority carrier band-offset created by the discontinuity in bandgaps between the lower-bandgap base and higher-bandgap buffer layer. This function may be achieved if the substrate bandgap and/or doping is higher than that of the base layer, in which case a separate buffer layer is not needed. However, including an epitaxial buffer layer generally improves material quality by blocking defect propagation from the substrate or initial growth interface. The emitter and base serve as the active (*i.e.*, radiation absorbing / minority carrier generating) layers of the TPV cell. In this analysis, the doping types of the emitter and base are not specified since both *p-on-n* and *n-on-p* configurations are common. The window layer is the same doping type as the emitter, and the buffer layer is the same doping type as the base. The bandgap E_G of the base and emitter layers is the primary determinant of the short-circuit current and open-circuit voltage of the TPV cell for a given radiation spectrum. The window layer should have a sufficiently wide bandgap (0.7–1.0 eV) in order to maintain a good transparency for the wavelengths absorbed by underlying emitter and base. Doping the window layer can reduce the sheet resistance of the emitter, lowering the series resistance of the solar cell, but excessive doping creates undesirable free-carrier absorption.

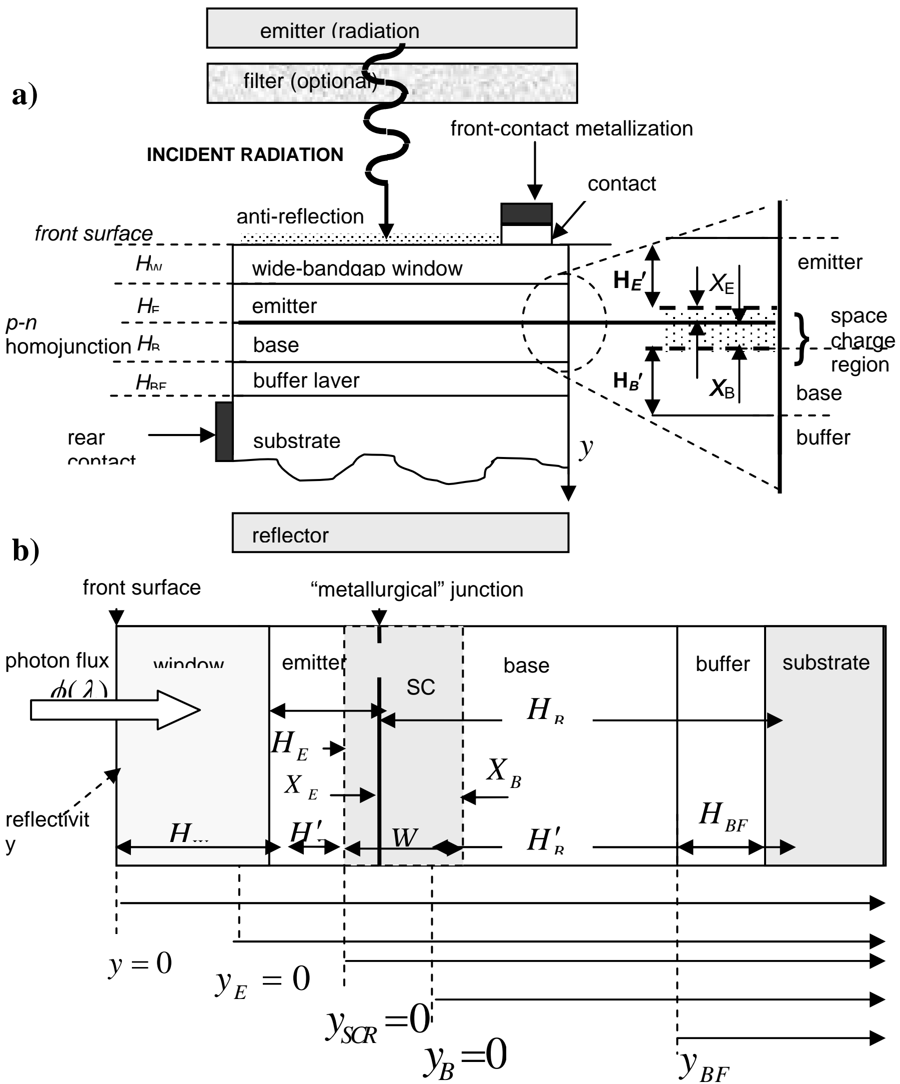


Figure 7. a) General schematic of a double heterostructure TPV cell showing the window, emitter, base, buffer, and substrate, and its positioning in a TPV system with respect to the radiation source and external reflectors and filters. **b)** Delination of device geometry for TPV cell model showing the various layers of the double heterostructure (window H_W, emitter H_E, base H_B, buffer H_{BF}) as formed by an epitaxial growth process on a substrate. Also indicated is the metallurgical junction and quasi-neutral regions (H_E' and H_B') that result from the space-charge region (SCR) of width W formed around the *p-n* junction.

5.2 Minority Carrier Recombination and Limits to Open-circuit Voltage

The analysis presumes that the minority carrier recombination rate $\Re$ in any layer of the device (*e.g.*, emitter, base, space-charge region, window) can be characterized by constant effective minority carrier lifetime τ_{eff} as

$$\Re = \frac{\Delta u}{\tau_{eff}} \tag{19}$$

Actually, this assumption is strictly valid only under low minority carrier injection conditions, *e.g.*, where the minority carrier concentration is less than doping level. Moreover, several recombination mechanisms will generally be operative in determining the effective recombination rate. The effective minority carrier lifetime in a layer of thickness d can be determined by considering contributions from radiative recombination (τ_{rad}), Auger recombination (τ_{Auger}), Shockley-Read-Hall (SRH) type defect-mediated recombination via midgap states (τ_{SRH}), and surface recombination at both bounding interfaces of the layer as characterized by surface recombination velocities S_1 and S_2. The equation for the effective minority carrier lifetime is

$$\frac{1}{\tau_{eff}} = \frac{1}{M \cdot \tau_{rad}} + \frac{1}{\tau_{Aug}} + \frac{1}{\tau_{SRH}} + \frac{S_1}{d} + \frac{S_2}{d} \tag{20}$$

M is a photon-recycling factor that accounts for the regeneration of minority carriers due to absorption of photons created by radiative recombination of minority carriers. M can range from 1 to ~10, depending on the optical properties of the interfaces, the thickness of the layer, and the absorption properties of the layer [13-15,34]. For example, highly internally reflective and/or scattering interfaces enhance photon recycling effects, whereas transmissive interfaces or optically-absorbing supporting substrates tend to foil photon recycling.

The radiative recombination rate is given by

$$\Re_{rad} = Bnp \tag{21}$$

with the proportionality constant B, called the bimolecular recombination coefficient. For low-level injection, a constant radiative lifetime τ_{rad} related to the doping N can be found as

$$\frac{1}{\tau_{rad}} = B\,N \tag{22}$$

A formula derived [13] for B is

$$B = 0.58 \times 10^{-12} \sqrt{\varepsilon} \left(\frac{1}{m_p + m_n} \right)^3 \left(1 + \frac{1}{m_p} + \frac{1}{m_n} \right) \left(\frac{300}{T} \right)^{3/2} E_G^2 \tag{23}$$

where ε is the dielectric constant, E_G is the bandgap, and m_p and m_n are the effective masses of holes and electrons, respectively. A simpler formula [35] approximating Equation (23) is

$$B \cong 3.0 \times 10^{-13} \left(\frac{300\ \text{K}}{T} \right)^{3/2} \left(\frac{E_G}{1.5\ \text{eV}} \right)^{2} \tag{24}$$

Figure 8 shows some B values for semiconductors of various bandgaps.

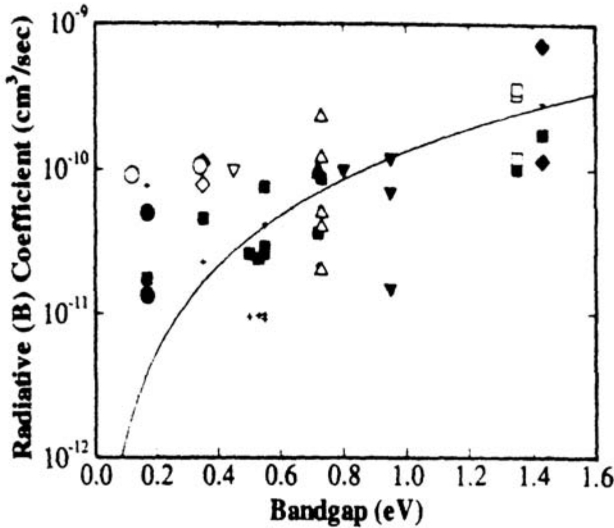


Figure 8. Calculated and experimental values of radiative coefficients (bimolecular recombination coefficient) B for direct bandgap III-V compounds (T = 300 K). ---, ■, + are calculations reported in the literature: experimental data: ◆ GaAs, □ InP, ▲ GaSb, ● InSb, △ InGaAs, ▼ InGaAsP, ▽InGaSb, and O InAsSb. From Charache *et al.*[29], see therein for specific references.

The rate of Auger recombination is given by

$$\Re_{Auger} = C_n n^2 p + C_p p^2 n \tag{25}$$

which at high injection becomes

$$\Re_{Auger} = C_n n^3 + C_p p^3 = Cn^3 = Cp^3 \tag{26}$$

At low injection levels material doped at a concentration N, a constant Auger lifetime can be found as

$$\frac{1}{\tau_{Auger}} = CN^2 \tag{27}$$

Theoretical models of Auger coefficients are rather involved. For low-bandgap III-V compound semiconductors, see for example Gelmont and Ivanov-Omskii [36], Rogalski [37], and Tian *et al.*. [38]. Figure 9 shows some C values for semiconductor with various bandgaps.

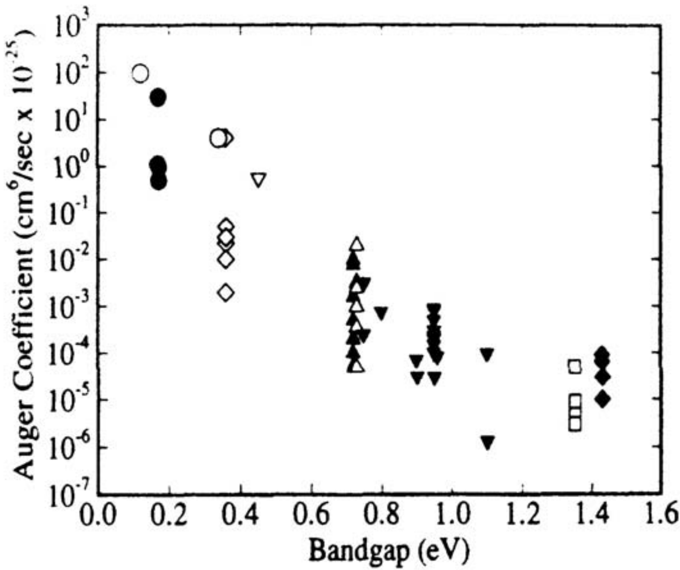


Figure 9. Calculated and experimental values of Auger coefficients C for direct bandgap III-V compounds (T = 300 K). experimental data: ◆ GaAs, □ InP, ▲ GaSb, ● InSb, △ InGaAs, ▼ InGaAsP, ▽InGaSb, and O InAsSb. From Charache *et al.*[29]; see therein for specific references.

The Shockley-Read-Hall (SRH) and surface recombination contributions to effective lifetime are dependent on the concentration of defects that produce midgap energy states. Therefore, it is not possible to derive useful correlations as a function of bandgap analogous to those shown in Figures 8 and 9. In many cases, and especially in high-quality, passivated semiconductors, the effective lifetime is usually dominated by radiative and/or Auger recombination, and the effects of SRH and surface recombination can be neglected for purposes of assessing ultimate efficiency limits. Studies, including experimental data, of minority carrier recombination properties in materials III-V compound alloys used for TPV cells, especially InGaAsSb, are available in [28,39-46].

In predicting and comparing the performance potential of various TPV materials and cells, finding the optimum bandgap for a particular application, and identifying loss mechanisms, it is useful to explore the limitations on dark current J_0 and open-circuit voltage V_{OC} in more detail.

As discussed previously, the dark current derived thermodynamically by the principle of detailed-balance is [16]

$$J_0 = \frac{q \cdot (n^2+1) \cdot E_G^2 \cdot kT}{4\pi^2 \hbar^3 c^2} \exp\left[-\frac{E_G}{kT}\right] \cdot \exp\left(qV/kT\right) \tag{28}$$

This equation indicates the minimum dark current and therefore stipulates an absolute upper limit to TPV cell open-circuit voltage. Equation (28) is for a cell formed on an optically absorbing substrate. For a TPV cell with a perfect back reflector, and thereby reaping the benefits of photon recycling, the term (n^2+1) can

be replaced by 1, corresponding to a more than ten fold reduction in dark current. These two cases are indicated by the two straight-line plots in Figure 4, and serve as a lower theoretical limit to $J_{0,}$ and thus analogously indicate an upper limit to V_{OC}, as a function of the bandgap of the TPV cell material.

More conservatively, the dark currents can be estimated by considering the various recombination mechanisms that all real devices must encounter. Calculating the dark current as due to sum of recombination over a cell thickness H

$$J_0(V) = \int^{H} \Re(V)dx \tag{29}$$

where

$$\Re(V) = \Re_{rad} + \Re_{Auger} + \Re_{SRH} + \Re_{surfaces} \tag{30}$$

The open-circuit voltage V_{OC} is then found by solving

$$J_0\left(V_{OC}\right) = J_{SC} \tag{31}$$

Using the charge neutrality condition and following relationship valid for carrier densities carriers in a device with external applied voltage V,

$$n \cdot p = n_i^2 \exp(qV/kT) \tag{32}$$

expressions for the dark current for each recombination mechanism, under either high or low injection, can be derived.

Table 1. Recombination limitations on diode dark current and open-circuit voltage

Recombination rate		Injection level	
		Low-level injection	High-level injection
Auger	$C_n n^2 p + C_p n p^2$	$J_{Auger} = qH\left(C_n N_D + C_p N_A\right) \times n_i^2 \exp(qV/kT)$ $V_{OC} = \frac{kT}{q}\ln\left[\frac{J_{SC}}{qH\left(C_n N_D + C_p N_A\right)n_i^2}\right]$	$J_{Auger} \cong qH\left(C_n + C_p\right) \times n_i^3 \exp(3qV/2kT)$ $V_{OC} = \frac{2}{3}kT\ln\left[\frac{J_{SC}}{qH\left(C_n + C_p\right)n_i^3}\right]$
Radiative	Bnp	$J_{rad} = qHBn_i^2 \exp(qV/kT)$ $V_{OC} = \frac{kT}{q}\ln\left[\frac{J_{SC}}{qHBn_i^2}\right]$	$J_{rad} = qHBn_i^2 \exp(qV/kT)$ $V_{OC} = \frac{kT}{q}\ln\left[\frac{J_{SC}}{qHBn_i^2}\right]$
SRH	$\frac{np}{\left\{\begin{matrix}\tau_{n0}(p+p_1) \\ +\tau_{p0}(n+n_1)\end{matrix}\right\}}$	$J_{SRH} = \frac{qHn_i^2}{N_D \tau_{n0}}\exp\left(qV/kT\right)$ $V_{OC} = \frac{2kT}{q}\ln\left[\frac{J_{SC}N_D\tau_{n0}}{qn_i^2 H}\right]$	$J_{SRH} = \frac{qHn_i \exp\left(qV/2kT\right)}{\left(\tau_{n0}+\tau_{p0}\right)}$ $V_{OC} = \frac{2kT}{q}\ln\left(\frac{J_{sc}(\tau_{n0}+\tau_{p0})}{q \cdot n_i \cdot H}\right)$
Surface	$\frac{np}{\left\{\begin{matrix}(p+p_1)/S_n \\ +(n+n_1)/S_p\end{matrix}\right\}}$	$J_{surface} = qn_i^2\left(\frac{S_p}{N_D} + \frac{S_n}{N_A}\right)$ $V_{OC} = \frac{kT}{q}\ln\left(\frac{J_{SC}}{q\frac{n_i^2}{N_D}S_p + q\frac{n_i^2}{N_A}S_n}\right)$	$J_{surf} = qn_i\left(S_n + S_p\right)\exp\left(qV/2kT\right)$ $V_{OC} = \frac{2kT}{q}\ln\left[\frac{J_{SC}}{qn_i\left(S_n + S_p\right)}\right]$

Table 1 summarizes various dark current and corresponding open-circuit voltage limitations for the several recombination mechanisms under both high- and low-level injection. For more details and derivation of the entries in Table 1, and discussion of performance limits for photovoltaic and TPV devices in general, reference is made to Shockley and Queisser [24], Henry [47], Green [48], Araújo and Martí [49], Luque [50], Landsberg [51], Mertens *et al.*. [52], Charache *et al.*. [29], Baldasaro *et al.*. [16], and Dashiell *et al.*. [33]. The symbols used in Table 1 have their usual meaning; *H* is the solar cell thickness. The open-circuit voltage is limited by whichever recombination mechanisms gives the lowest value of V_{oc}.

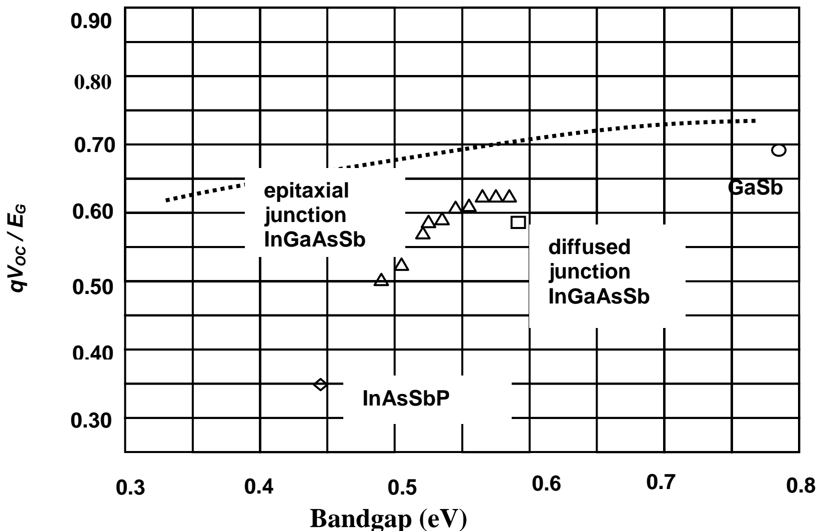


Figure 10. Voltage factor comparison of various TPV cells. After Wang [46].

Wang [53] has compiled open-circuit voltage data in the form of a voltage factor (qV_{OC} / E_G) as a function of bandgap E_G for GaSb, InGaAsSb, and InAsSbP TPV cells (see Figure 10). The V_{oc} data are corrected to $3A/cm^2$ short-circuit current densities for a consistent comparison. Also shown is an estimated limit to the voltage factor (dashed line) as determined by an analysis by Baldasaro *et al.*. [54]. The data underscore the observation that lower-bandgap (< 0.5 eV) TPV cells fall short of their estimated potential open-circuit voltage. This can be attributed to Auger recombination effects that become very pronounced at lower bandgaps, and/or unoptimized devices and immature materials technology that reduce the performance of the lower-bandgap devices. In this regard, very little effort has been directed toward InAsSbP materials and devices compared to that of GaSb and InGaAsSb. As there is interest in developing low-bandgap (< 0.5 eV) devices for tandem TPV cells, micron-gap TPV cells and other applications, as well as LEDs, lasers and detectors, these issues merit more attention from both a modeling and experimental perspective.

6 Survey of TPV Materials and Devices

The foregoing discussion of TPV materials and devices centered on InGaAsSb/GaSb as a representative example of state-of-the-art TPV technology. In the following, other important materials, devices and device configurations are surveyed. We emphasize recent developments in the last several years to supplement topics covered in the extensive reviews by Coutts [2,6].

6.1 Silicon, Germanium, and Si-Ge Alloy TPV Cells

Silicon and germanium have long and venerable histories in semiconductor device technology, and further, silicon photovoltaic devices are the mainstay of the world solar cell industry. Thus, there has been considerable incentive to leverage the technology base and manufacturing infrastructure developed for silicon solar cells for TPV applications. Early TPV work explored the use of silicon solar cells in converting thermal energy produced by concentrated sunlight to electricity. However, without efficient spectral shifting, the relatively high bandgap of silicon solar cells precludes its use in combustion, radio-isotope, or waste-heat recovery systems where high heat-to-electricity conversion efficiencies are needed in order to be economically viable or displace competing technologies. Nevertheless, silicon TPV cells may find niche applications such as co-generation with residential furnaces where low efficiencies can be tolerated and low-cost is a critical requirement. For example, Bitnar *et al.* [55] advocate a combustion-flame TPV system that uses a foam ceramic emitter made from Yb_2O_3 (which provides spectral shifting suited for the silicon solar cell spectral response) and commercially available silicon solar cells. This system is expected to have an efficiency of about 1% which is sufficient for autonomous furnace applications.
Germanium (bandgap ~0.7 eV) cells provide a significantly better spectral response match to most thermal emission sources in comparison to silicon TPV devices. However, despite its use as the low-bandgap scavenger bottom cell of high-efficiency multijunction concentrator and space solar cells, Ge is not a particularly

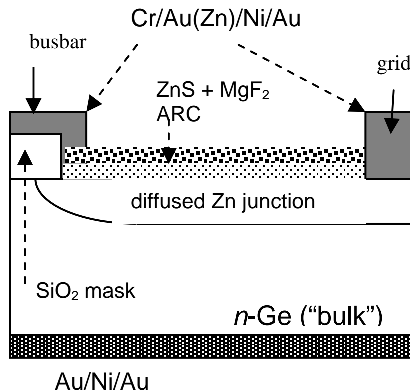


Figure 11. Zn –diffused Ge TPV cell. After [56].

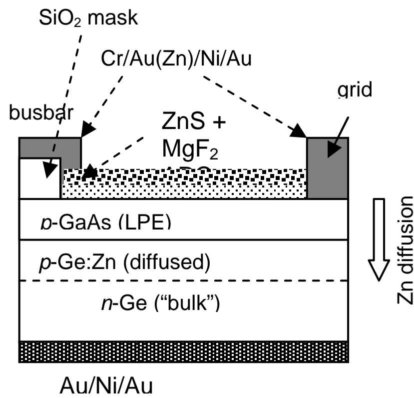


Figure 12. GaAs/Ge heteroface TPV cell. After [56].

good photovoltaic material. This is due, as Coutts [2] pointed out and which was no doubt widely appreciated, to the high electron effective mass of Ge, which compared to III-V semiconductors of similar bandgap, leads to an anomalously high conduction band density of states and consequent high intrinsic carrier concentration density. As a result, Ge diodes exhibit a relatively high reverse-saturation ("dark") current and a low open-circuit voltage compared to III-V compounds of similar bandgap. The low voltage of Ge cells is further exacerbated by the difficulty in passivating Ge surfaces.

Despite these limitations, the Ioffe Institute [7,56,57] has pursued the development of Ge-based TPV cells (Figure 11). *n*-type Ge wafers are coated with Si_3N_4 or SiO_2 films and patterned with openings to serve as a mask for selective Zinc diffusion to form the *p-n* junction. This avoids junction perimeter effects at the edge of the device that create excess dark current and lower open-circuit voltages. An improved Ge-based cell with a surface-passivating GaAs window layer was also developed (Figure 12). The 0.1-μm thick epitaxial GaAs window is grown by liquid-phase epitaxy (LPE) at 380 °C from a lead (Pb) melt. Zn is then diffused through the window and into the *n*-type Ge base to form a *p*-GaAs/*p*-Ge/*n*-Ge heteroface TPV cell. Figure 13 compares the performance (short-circuit current, open-circuit voltage, and fill-factor with a variable intensity AM0 spectrum) for the Zn-diffused homojunction and the GaAs/Ge heteroface TPV cells. It is clear the addition of a GaAs window layer gains a significant boost in open-circuit voltage. The incorporation of an epitaxy step will no doubt add not insignificant expense. Nevertheless, these results have already exceeded the expectations of low-performance traditionally anticipated for Ge TPV cells.

Bittnar [55] comments "whether Ge photocells will ever be applied in TPV systems" remains doubtful. Their efficiency remains lower compared to low bandgap III-V materials such as GaSb. Today, Ge cells are cheaper than GaSb cells by a factor of 6 or 7, but this difference may change when GaSb cells are manufactured on large scale." One potential advantage to Ge is that it can be deposited and re-crystallized in the form of large-grain, thin-films supported on low cost substrates. This could provide a significant advantage in terms of reduced materials and processing costs. It would not be as easy to produce similar large-grain

III-V semiconductor thin films due to the incongruent melting of such compounds and the high volatility of the Group V component.

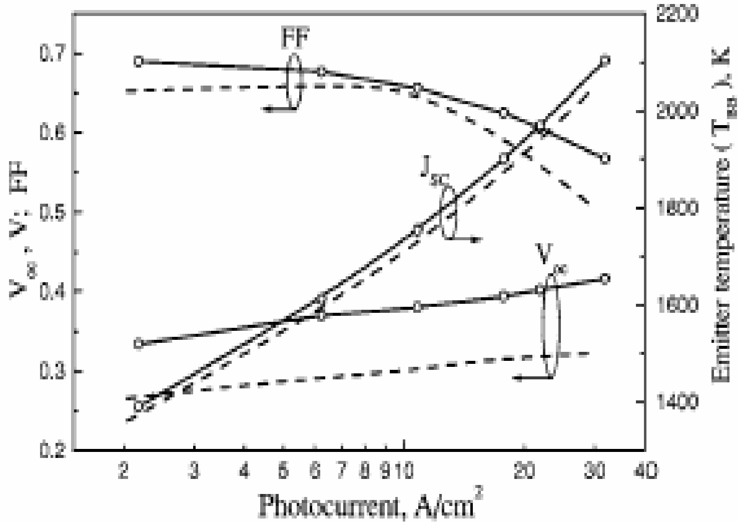


Figure 13. Comparison of Zn-diffused Ge TPV cell (*dashed line*) and GaAs/Ge heteroface TPV cell (*solid line*). From [56].

TPV cells can also be made from Si-Ge epitaxial layers grown on Ge or Si substrates, including strained multilayer quantum wells and quantum dot structures, see for example Bitnar [55]. This approach is still in its infancy and has as yet impacted neither solar nor TPV cell technology.

6.2 GaSb TPV Cells

GaSb has a bandgap close to that of Ge, but GaSb-based TPV cells generally outperform Ge cells. GaSb TPV cells can be fabricated using simple Zn-diffusion in a GaSb wafer to form a *p-n* junction (Figure 14). As discussed by Sulima *et al.*. [58], the resulting Zn-doping profile appears to create a *built-in* electric field effect that improves carrier collection and reduces the junction reverse saturation current. Actually, as Sulima indicates, these favorable diffusion effects are not unique to GaSb, but are operative in Zn-diffused *p-n* junctions formed in InGaSb, InGaAsSb, and InAsSbP as well. Post-diffusion anodic oxidation of the front surface can be used for controlled removal of the heavily Zn-doped "dead" layer at the front surface, as well as passivate the front surface for reduced minority carrier surface recombination and form a dielectric film that acts as an encapsulating layer and anti-reflection coating. Otherwise, the heavily Zn-doped "dead" layer at the front surface of the device degrades performance due to bandgap narrowing effects and increased defect-mediated minority carrier recombination.

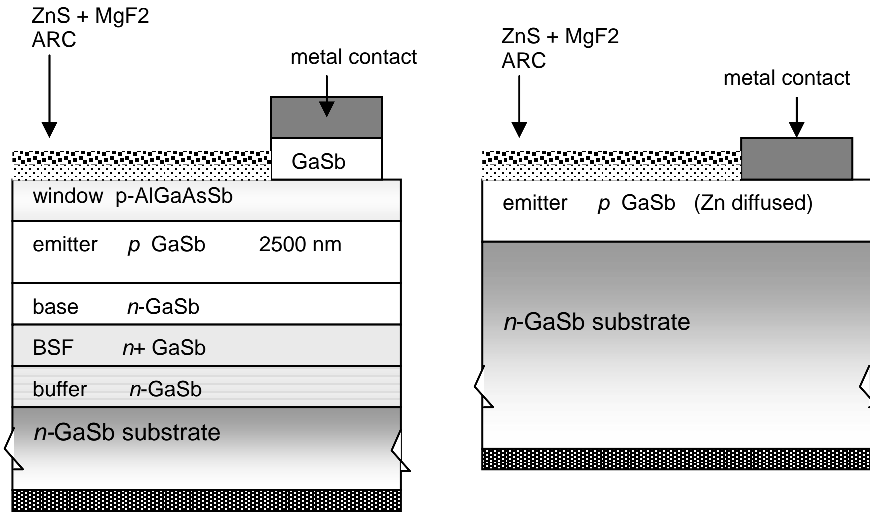


Figure 14. Side-by-side comparison of simple Zn-diffused p-n junction GaSb TPV cell (right) and epitaxial GaSb cell (left). After [59].

To appreciate the performance penalties incurred by using a simple zinc diffusion process to GaSb TPV devices, it is informative to compare the spectral response (Figure 15) of Zn-diffused *p-n* GaSb TPV cells with epitaxial p-n junction GaSb cells passivated with a wide-bandgap, lattice-matched AlGaAsSb window layer (as shown in Figure 14) grown by MOCVD [59].

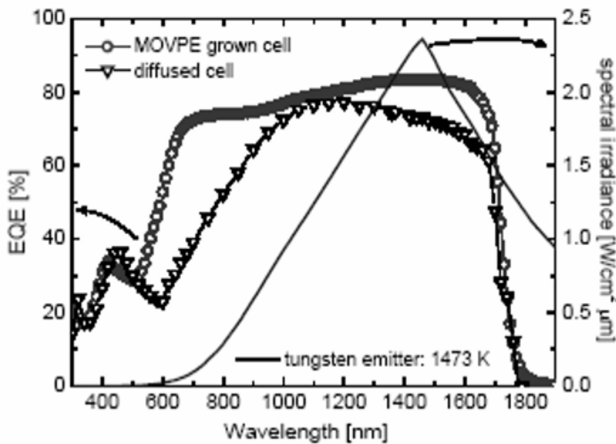


Figure 15. Comparison of external quantum efficiencies for TPV cells shown in Figure 14, superimposed with spectral irradiance curve of tungsten emitter at 1473 K. From [59].

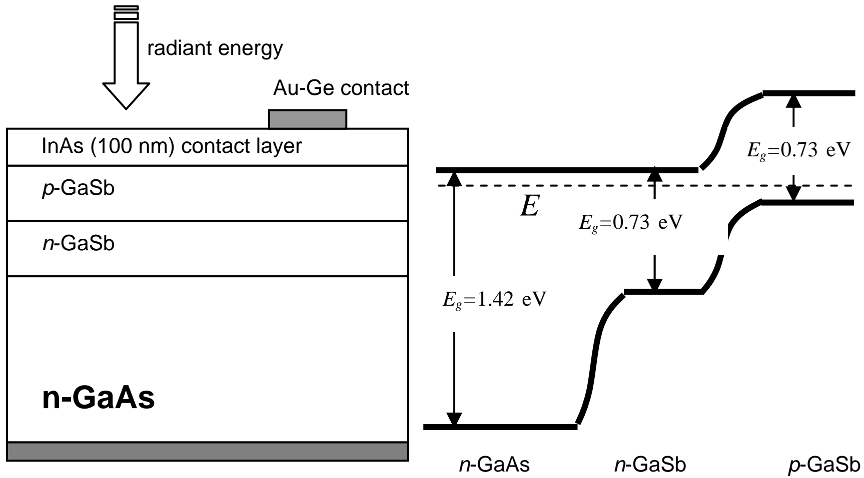


Figure 16. GaSb on GaAs TPV cell. After. [65].

Bett and Sulima [60] have provided a more extensive review of GaSb TPV cells. Stollwerck *et al.* [61] and Martín and Algora [62] describe modeling and simulation, including temperature-dependent effects, of GaSb TPV cells. An investigation of transparent conducting oxides as anti-reflection coatings for GaSb TPV cells has been reported [63], as well as the use of sub-wavelength grating structures formed in the front surface of the GaSb device to provide broadband antireflection [64]. Another interesting development by Bumby *et al.* [65] is a report of GaSb *p-n* junction TPV cells grown by MOCVD on GaAs substrates (Figure 16). GaAs substrates are cheaper and available in larger areas than GaSb. Also, the wider bandgap of GaAs and its ability to be rendered semi-insulating may offer some advantages in device design with respect to photon recuperation and MIM (monolithic interconnected module) configurations. These devices achieved 70% of the efficiency as that of devices made on GaSb wafers, despite the 7% lattice mismatch between GaAs and GaSb. Evidently, the GaSb system is more forgiving of defects than most other III-V semiconductors (GaN excepted). Most of the various TPV systems that appear furthest developed and closest to commercialization utilize GaSb TPV cells [66].

6.3 TPV Cells Based on Bulk III-V Alloy Ternary Crystals

Commercial semiconductor wafers are generally limited to the elemental semiconductors (Si and Ge) or the congruently melting binary compound semiconductors (*e.g.*, GaAs, InP, GaSb, InAs, ZnSe, ZnS, and SiC). Many optoelectronic device technologies, including TPV, rely on the ability to vary semiconductor alloy compositions to tune the bandgap, electron affinity, and valence and conduction band offsets. This, in turn, provides a certain degree of control of optical absorption and emission, improvement in minority carrier injection

efficiency, and effective optical and carrier confinement. Unfortunately, the small set of commercially available substrates restricts the bandgap range of semiconductor alloys due to constraints imposed by lattice matching conditions needed to avoid excessive defects, miscibility gaps, and stress and strain effects. Semiconductor wafers sliced from bulk crystals of ternary (or even quaternary) alloys offer a means to circumvent these restrictions. In one scenario, ternary alloy wafers are fabricated into TPV cells by simple dopant diffusion processes to form a *p-n* junction, similar to that described above for Zn-diffused GaSb TPV cells. In this case, the objective would be to use ternary alloy wafers with bandgaps lower than GaSb, thus avoiding altogether the epitaxy processes currently used to produce TPV cells with bandgaps in the 0.5–0.7 eV range. Alternatively, the ternary wafer might be used as a substrate for lattice-matched epitaxial devices, with the same objective of achieving longer wavelength response or more flexibility with respect to tuning heterostructure bandgaps, electron affinities and band offsets, or to induce strain effects for favorable modification of electrical and optical properties.

The growth of ternary bulk crystals with homogeneous and/or controlled compositions is challenging due to severe segregation effects, which result in compositional grading, stress, nucleation of new phases or misoriented domains, and high defect generation. The amount of useful wafer with a specified target composition netted from an ingot can be small. In the past, many efforts along these lines were abandoned as either uneconomic or technically problematic. Nevertheless, there has been some progress in this area for specific application to TPV devices. Dutta *et al.* [67] report a "self solute feeding" method to grow compositionally-homogeneous GaInSb crystals in the bandgap range of 0.4–0.7 eV. Zn-diffused GaInSb bulk devices showed peak external quantum efficiencies over 60%. Mitric *et al.* [68] and Stelian *et al* [69] describe several crystal growth techniques (vertical Bridgman with magnetic field, vertical Bridgman technique with feeding, double crucible Czochralski, and vertical Bridgman with baffle) to produce GaInSb bulk crystals for TPV applications.

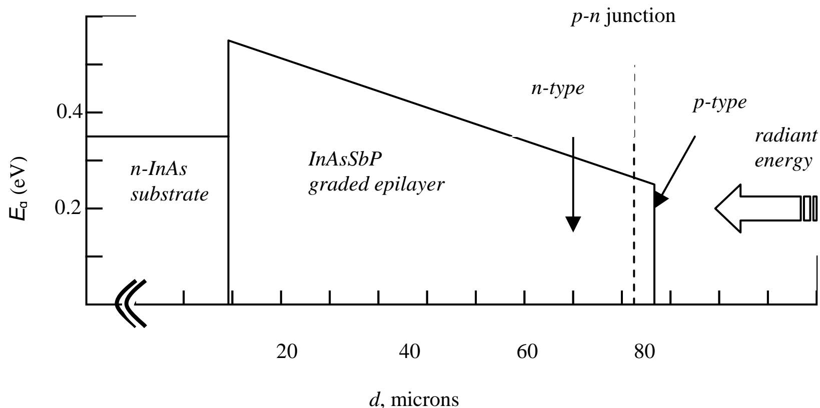


Figure 17. Schematic cross-section bandgap profile of compositionally-graded InAsSbP layer made by liquid-phase epitaxy. After [72].

Another route with similar objectives involves the epitaxial growth of thick (> 100 μm) self-supporting epilayers of ternary and quaternary alloys (InAsSb, InGaSb, InGaAsSb, AlGaAsSb, and InAsSbP), which can be separated from their seeding substrate and/or bonded to a surrogate substrate (Mauk *et al.*, [70,71].) As a specific example relevant to TPV, Mikhailova *et al.* [72] developed thick, compositionally-graded InAsSbP photodiode device structure using liquid-phase epitaxy. A sufficiently gradual compositional grading over a thickness of about 80 μm yields a low-defect quaternary composition with a bandgap (~0.3 eV) that is otherwise precluded with lattice-matched epitaxial layers (Figure 17). Removal of the InAs substrate would reduce parasitic absorption losses and enhance heat sinking of the device.

6.4 InGaAs and InAsP/InP TPV Cells and MIMs

In terms of sophistication and current, best-reported performance, TPV cells made in InGaAs/InP and/or InAsP/InP rival - and in certain respects surpass - those made in GaInAsSb. InGaAs/InAsP/InP TPV cells have been demonstrated with radiant heat conversion efficiencies greater than 20% for emitter radiator temperatures of 1039 °C [4]. This represents the state-of-the-art for TPV technology in general. The common denominator of these and related alloys in their application to TPV is an InP substrate.

The initial work on InGaAs/InP TPV cells used the technology of InGaAs PIN and avalanche photodiodes and IR imaging arrays. InGaAs/InP detectors had already been developed for the low-loss wavelength range of optical fibers (1.3 and 1.55 um) and to detect “eye-safe” lasers. $In_{0.53}Ga_{0.47}As$ is lattice-matched to InP and has a bandgap of 0.73 eV, corresponding to an optical absorption edge of about 1700 nm. So-called ‘extended (wavelength)’ InGaAs/InP detectors have lattice-mismatched (0.75–1.7%) epilayers to achieve lower bandgaps (0.5–0.6 eV) and longer wavelength (1.7–2.5 um) response. In the $In_{1-x}Ga_xAs$ /InP system [73]

$$E_G(x) = 0.555x^2 + 0.505x + 0.356 \text{ [eV]} \tag{32}$$

and the lattice mismatch $M(x)$ between InGaAs epilayer and InP substrate is

$$M(x) = \frac{-0.378 + 0.810x}{11.927 - 0.405x} \tag{33}$$

InGaAs/InP has been especially favored for TPV applications because MOCVD growth of low-bandgap InGaAs epitaxial layers on InP is well developed; high-quality, large-diameter (150-mm) InP substrates are commercially available; and semi-insulating InP substrates have low optical absorption at wavelengths where photon recuperation can be effective [74]. As with InGaAs detectors, there was incentive to develop lattice-mismatched, low-bandgap (0.5–0.7 eV) InGaAs/InP TPV devices in order to extend the spectral response out to ~2500-nm wavelength. A lattice-matched InGaAs/InP TPV device is shown in Figure 18 and a TPV cell with a compositionally-graded layer that bridges the lattice mismatch between the buffer/substrate and “active” InGaAs layers of the device that host the *p-n* junction is shown in Figure 19. This expedient permits the TPV spectral response to be extended out to about 2300 nm. In the early work, the compositional grading to

achieve lower bandgap devices did not completely ameliorate lattice-mismatch effects, and a penalty in quantum efficiency and open-circuit voltage was evident. Initially, there was also some concern that high current levels (~A/cm2) used in TPV devices in combination with such large lattice mismatch and/or excessive strain, would induce defect generation causing accelerated device degradation. However, it appears that considerable progress has been made on this problem, especially with the use of InPAs step-graded buffer layers used to bridge the 1.6% lattice-mismatch between the InP substrate and 0.52-eV-bandgap InGaAs *p-n* junction [5]. For example, Figure 20 is an indicator of progress in developing low-bandgap InGaAs TPV cells as judged by the junction reverse-saturation (dark) current $J_{01}$and diode ideality factor (n). A low value of J_{01} with $n = 1$ is a good indicator junction and material quality. InGaAs/InPAs/InP MIM-0configured TPV devices with 0.52-eV bandgaps have exhibited internal quantum efficiencies of 95%, reverse-saturation currents (J_{01}) of 57 μA/cm^2, and defect densities less than 10^7 cm^{-2}.

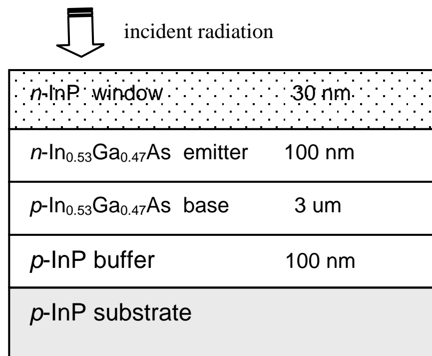


Figure 18. Lattice-matched InGaAs/InP TPV epitaxial device structure. After [75].

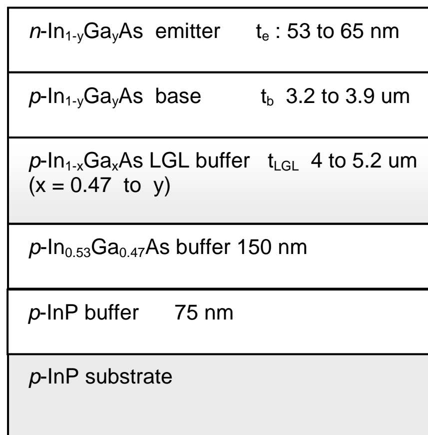


Figure 19. Lattice-mismatched InGaAs/InP TPV epitaxial device structure with graded InGaAs buffer layer. After [75].

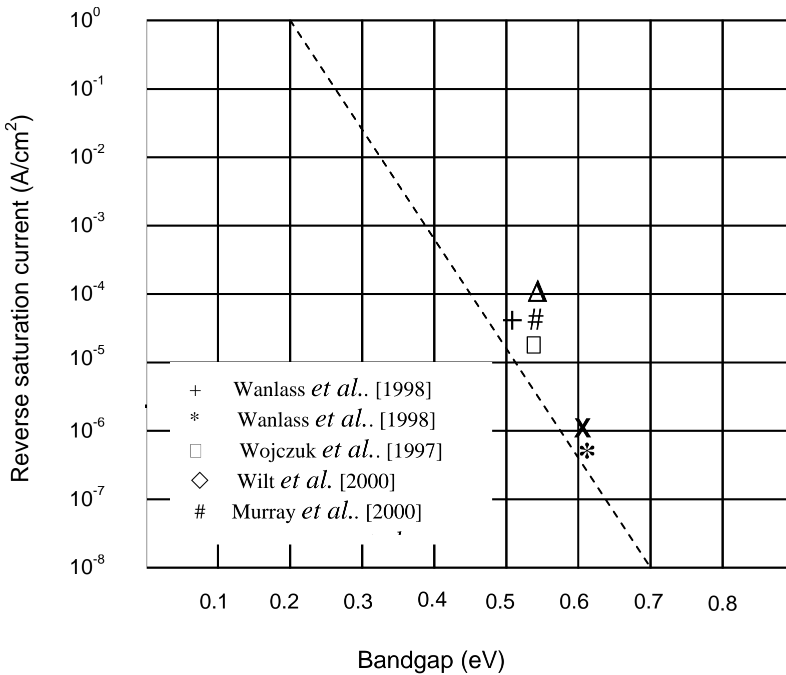


Figure 20. Reported reverse-saturation currents *vs* bandgap for InGaAs TPV cells [5]

InGaAs/InP TPV cells made on semi-insulating (SI) InP substrates configured as monolithic interconnected modules (MIMs) are shown in Figure 21. An important feature of semi-insulating InP is its low optical absorption which permits an effective mirror for sub-bandgap photons by simply depositing reflective films on the underside of the substrate. In general, the series interconnection schemes permit higher open-circuit voltages, better reliability, better heat sinking, and improved spectral response and photon recuperation. The *p*-on-*n* (*p*-type emitter / *n*-type base) TPV device element shown in Figure 21a is part of a 15-element array. The series-interconnection is facilitated by a heavily-doped lateral conduction layer (LCL). An *n*-on-*p* TPV cell (Figure 21b) configuration should reduce free carrier absorption associated with heavily-doped *p*-type emitter of the TPV cell shown in Figure 21a, improving overall optical performance on several counts. The *p*-InAsP base makes electrical contact to the *n*-type LCL *via* a *p/n* InGaAs tunnel junction which serves as a shorting contact between the two oppositely doped layers. A 2 cm × 2 cm, 30-cell *n/p/n* TPV MIM with a bandgap of 0.6 eV has achieved radiant energy-to-electric conversion efficiencies of 23.9% [76].

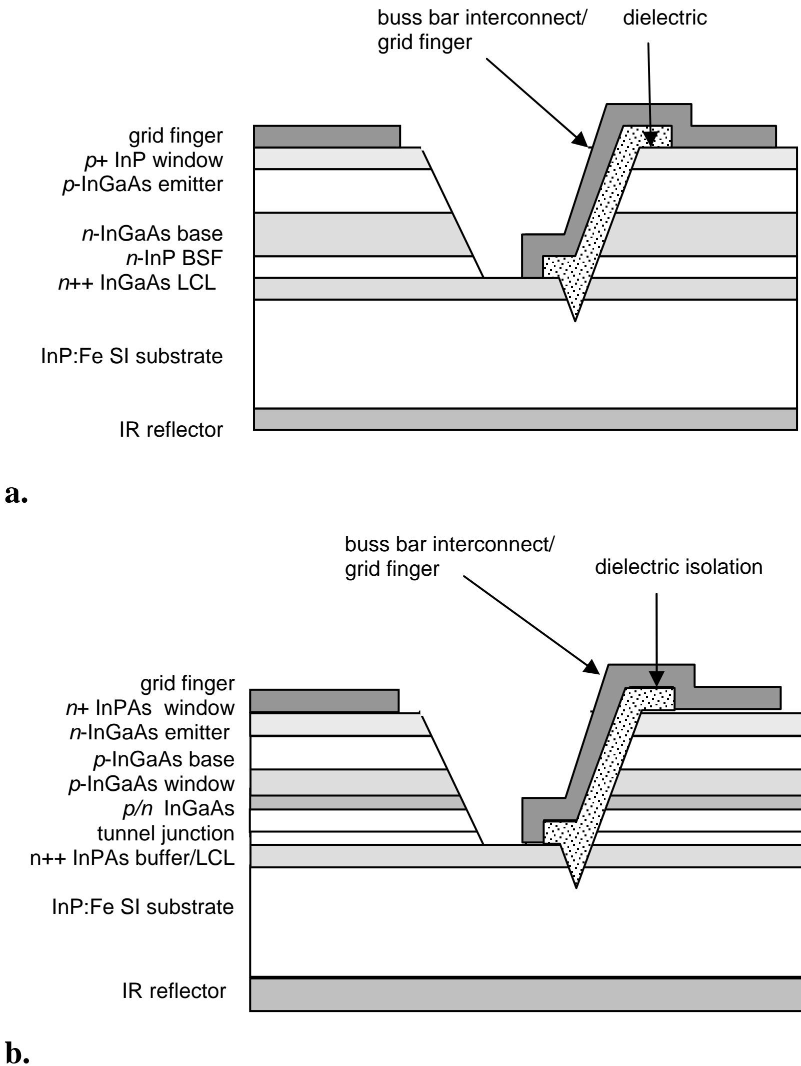


Figure 21. InGaAs/InP (**a**) and InGaAs/InPAs/InP MIM (**b**) TPV cells. After [77].

6.5 Low-bandgap (< 0.5 eV) InAs and InAsSbP TPV Cells

Although the bandgaps and spectral response of the two most developed epitaxial TPV devices (GaInAsSb/GaSb and InGaAsP/InP) are tunable over an appreciable range, there appears to be a practical lower bandgap limit of roughly 0.5 eV for both of these systems for reasons already mentioned. Some theoretical studies [21] suggest advantages in using even lower bandgap (0.2–0.5 eV) TPV cells. Lower

bandgap TPV cells could also find use in tandem TPV cell configurations as the low-photon-energy scavenger cell. Further, micron-gap configurations for TPV cells (see Section 6.10) could benefit from low (<0.5-eV) bandgap TPV cells. Micron-gap TPV seeks to exploit the enhanced radiative transfer between hot and cold surfaces separated by a vacuum gap on the order of the radiation wavelength. Low-bandgap TPV cells imply longer useable wavelengths, which could simplify the mechanics for realizing the micron-gap. Also, low-bandgap TPV cells are a better match to radiators/emitters operating at lower temperatures, and there would be considerable incentive to operate the emitter at as low a temperature as feasible in order to avoid overheating a TPV cell situated very close to the emitter. (On the other hand, low-bandgap TPV cells are more sensitive to heating which argues against their use in micron-gap TPV systems.) At any rate, the availability of TPV cells with bandgaps less than 0.5 eV would be generally useful to systems designers and would provide a means to experimentally assess and exploit new TPV concepts.

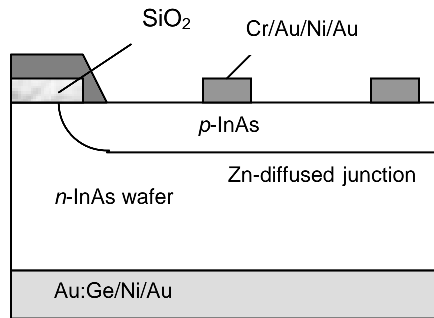


Figure 22. Diffused-junction InAs TPV cell. After [78].

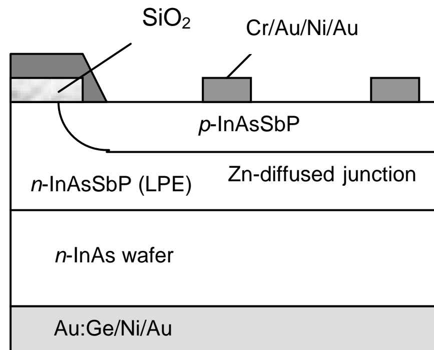


Figure 23. Epitaxial InAsSbP/InAs TPV cell. After [78].

Figures 22 and 23 show InAs and InAsSbP diffused junction TPV cells reported by Andreev *et al.*. [78]. The spectral response of a Zn-diffused, anodic oxidized InAsSbP TPV cell is shown in Figure 24. Other versions of InAsP and InAsSbP based TPV cells are shown in Figure 25. All of the InAsSbP TPV cells described

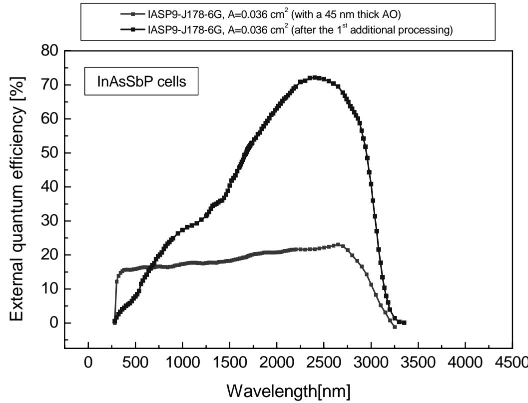


Figure 24. Spectral response of Zn-diffused InAsSbP/InAs TPV cell. Bottom curve is before anodic oxidation, top curve is after anodic oxidation. From [79].

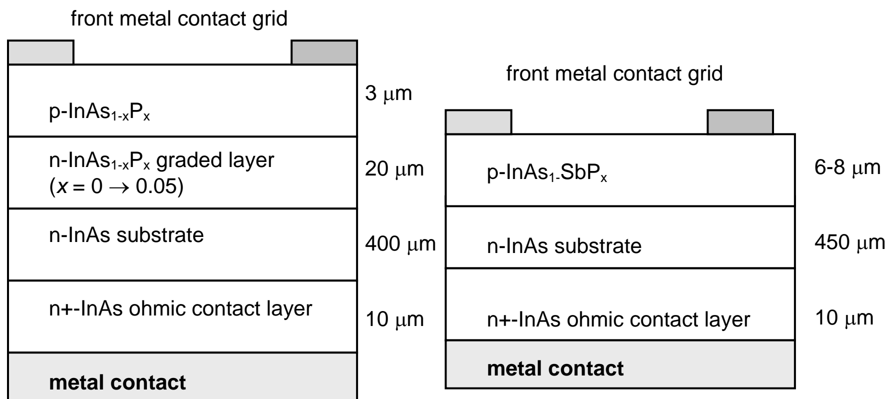


Figure 25. InAsP and InAsSbP TPV cells. After [80].

here were made by liquid-phase epitaxy (LPE), as it appears this technique is comparatively well developed to grow a quaternary alloy that contains three distinct Group V elements (As, Sb, and P). InAsSbP cells have low open-circuit voltages, as evidenced by the small cut-in voltage in the current-voltage characteristic of an InAsSbP TPV cell shown in Figure 26. To what extent this is due to fundamental limitations imposed by Auger recombination, or is simply reflective of their relatively primitive state of technical development remains to be seen. The low voltages may be due mainly to non-optimized doping and inadequate surface passivation. It has proved difficult to dope the low-bandgap InAsSbP materials. This may be due to intrinsic doping and compensating effects associated with impurities and other defects. An interesting sidelight on this issue is the remarkable improvement in quantum efficiency that can be gained with rare-earth gettering, which would implicate extrinsic effects are limiting the minority carrier lifetimes in these materials.

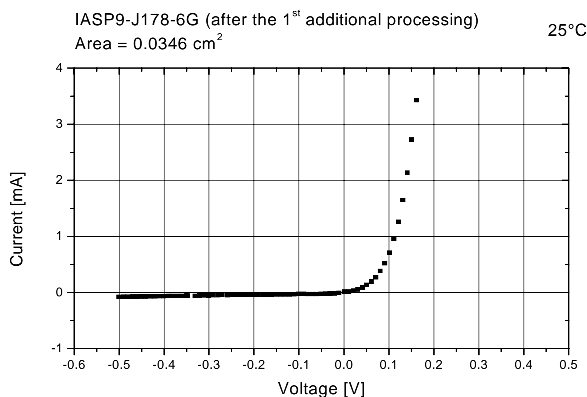


Figure 26. Current-voltage characteristic of Zn-diffused InAsSbP TPV cell after anodic oxidation (same device shown in Figure 24, from [79])

The tendency to dismiss the prospects for low-bandgap (<0.5 eV) TPV cells on the basis of limitations associated with Auger recombination should be tempered by the following observations:

1) While Auger recombination may severely limit the performance of light-emitting devices such as LEDs and lasers, it does not necessarily impose the same penalty on TPV cells. In light-emitting devices, the efficiency is reduced because Auger recombination competes with radiative recombination. However, this in itself does not preclude achieving a long minority carrier lifetime and diffusion length in these materials, but merely says Auger recombination is the dominant mode of recombination. The luminescence efficiency is not directly relevant to TPV efficiency. Thus, because much of the literature on this subject is largely concerned with laser diode performance issues, the impression that Auger recombination effects will severely curtail the performance of *all* optoelectronic devices, including TPV cells, is overly pessimistic.

2) High spectral response can be obtained in InAsSbP diodes and TPV cells, as demonstrated in Figure 24. This shows that InAsSbP with good minority carrier diffusion lengths, and thus sufficiently high lifetimes are feasible.

3) The room-temperature performance of InAsSbP LEDs is remarkably high. Krier and Mao [81] demonstrated InAsSbP LEDs with a 2.3% internal quantum efficiency at room temperature. This luminescence efficiency is consistent with the relative magnitudes of Auger and radiative recombination that produce sufficiently low dark currents for respectable open-circuit voltages. In more recent work, Kane *et al.*. [82] reported InAs/AlAsSb LEDs with a room-temperature internal radiative efficiency of 24%. These results surpassed previous theoretical predictions, showing that much longer minority carrier lifetimes can be attained than had been assumed possible in low-bandgap III-V compounds. They attributed these dramatic improvements in luminescence efficiency to effective surface passivation of the diodes. Evidently surface recombination, not bulk Auger recombination, has limited the performance of InAs LEDs.

4) Mikhailova *et al.*. [72] determined an effective lifetime of 5 ns in low-bandgap InAsSbP *p-n* junction photodiodes.
5) Chen *et al.*. [83] characterized $InAs_{0.86}As_{0.5}P_{0.09}$ epitaxial diodes. This composition has a 300 K bandgap of 0.39 eV. The measured dark current density was 12.6×10^{-3} A/cm^2. With an incident irradiation corresponding to a short-circuit current of 3 A/cm^2, an open-circuit voltage of 0.196 V is implied, for which $qV_{oc}/E_g = 0.5$.

6.6 TPV Cells Based on Other Semiconductors

TPV cells fabricated from still other semiconductors, in addition to silicon, germanium, and the III-V compounds and alloys already discussed, also merit mention. Mercury cadmium telluride (MCT) alloys might appear as viable candidates in light of the considerable work on MCT infrared detectors. This has not proved to be the case. Evidently, the well-known material instabilities, fabrication difficulties, low production yields, and consequent high expense of this semiconductor alloy has been sufficient to deter its exploration for TPV applications, especially in light of the progress demonstrated for low-bandgap III-V antimonides such as InGaAs, InGaAsSb, and InAsSbP. On the other hand, in systems employing spectral control, or where cell cost is an over-riding consideration, such as applications where silicon TPV cells are considered viable, then in fact, some of the wider-bandgap II-VI compound semiconductor materials developed for low-cost solar cells, *e.g.*, CdTe and $CuInSe_2$, might be considered for TPV as well. The advantage to thin film solar cell materials is their amenability to conformal deposition on curved surfaces, which could prove useful in TPV system design, as for instance, in forming the TPV cell as an integrated module on the inner surface of a cylindrical heat shield that surrounds the radiator, see for example Bitnar *et al.* [84]. The crucial feature of thin-film II-VI materials, and which distinguishes them from silicon, germanium, and III-V materials is related to the fact that polycrystalline films that result from their (non-epitaxial) deposition on crystallographically dissimilar substrates such as metals and plastics can support devices with adequate photovoltaic performance. On a different tack, Liu and Wang [85] have suggested multiple-intermediate level interband and intersubband quantum dot thermophotovoltaic devices, and their fabrication in the GaN/AlGaN/GaInN system. The use of bulk III-V nitrides (with nominally large bandgaps), with the aim of exploiting the controlled band offsets that are possible in these semiconductors in order to make quantum dot and quantum cascade TPV devices, stands out from the almost exclusive focus on low-bandgap III-V compound semiconductors for TPV.

6.7 Quantum Well TPV Cells

Quantum well TPV cells have been explored to assess their potential advantages relative to conventional TPV device structures. The basic quantum well TPV (or solar) cell structure is comprised of a *p-i-n* structure where the intrinsic (undoped) region is sandwiched between *p-* and *n*-doped regions. The intrinsic region is commonly comprised of multiple quantum wells, *i.e.*, alternating semiconductor layers of two distinct bandgaps. Figures 27 and 28 show two examples of quantum

well TPV cells, both of which are made from multilayer epitaxial InGaAs structures on an InP substrate. The chief anticipated advantages of quantum well solar and TPV cells are related to 1) an adjustable effective bandgap to tailor the absorption and spectral response, and 2) reduced minority carrier recombination yielding lower dark currents and higher open-circuit voltages than conventional cells with corresponding or similar spectral response. Both of these features are crucial in TPV devices.

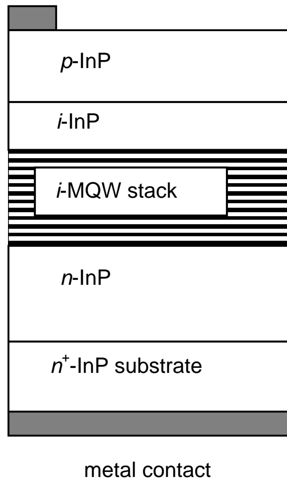


Figure 27. Multi-quantum well (MQW) TPV cell. After [87].

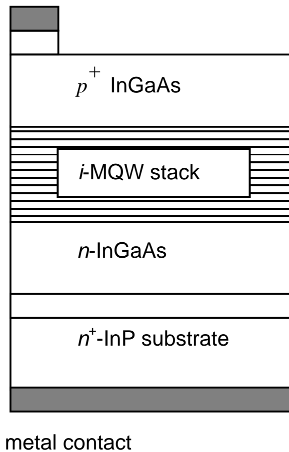


Figure 28. Multi-quantum well (MQW) TPV cell. After [88].

As quantum well and superlattice solar cells have yet to replace conventional solar cells based on bulk materials and thin films, even in specialized applications such as space power and concentrators, their application to TPV is probably premature. Nevertheless, the favorable advantages already demonstrated, as reviewed by Connolly and Rohr [86] and indicated in Figures 29 and 30, are compelling. Further, this approach should benefit from the considerable efforts in developing quantum wells and superlattices for other types of optoelectronics devices.

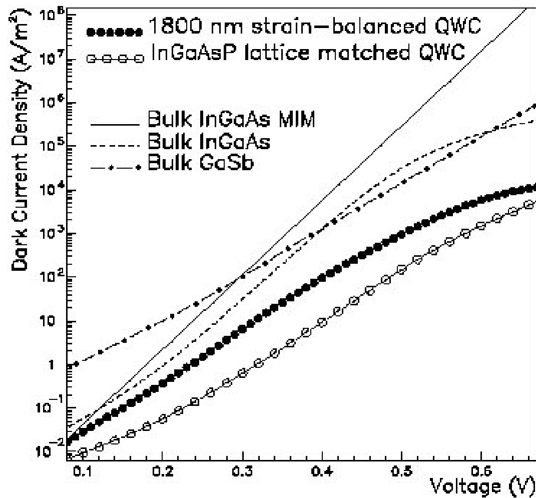


Figure 29. Comparison of dark currents for bulk and quantum well TPV cells. From Connolly and Rohr [86].

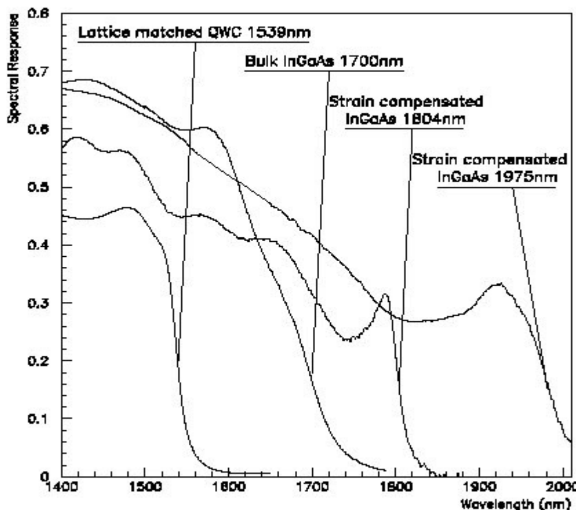


Figure 30. Comparison of dark currents for bulk and quantum well TPV cells. From Connolly and Rohr [86].

6.8 Isolation for Series Interconnection, Integrated Reflectors, and Wafer-bonding for TPV Cells

The monolithic series-interconnection of TPV elements to produce high-voltage modules described for InGaAs/InAsP/InP MIM TPV devices in Section 6.4 cannot be implemented in InGaAsSb/GaSb TPV devices because semi-insulating GaSb substrates are not available. GaSb cannot be rendered semi-insulating probably for a combination of reasons including its relatively low bandgap, and defect doping mechanisms that produce intrinsic doping. This is a serious drawback for the InGaAsSb/GaSb system relative to InGaAs/InAsP/InP.

One approach to realize electrical isolation of device elements employs an epitaxial structure that provides for reverse-biased cell-isolation diodes (CID) as shown in Figure 31.

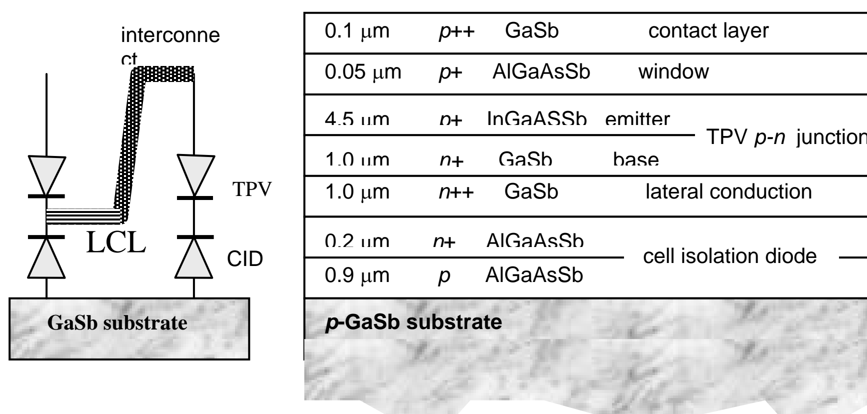


Figure 31. Wiring diagram and epitaxial structure for InGaAsSb/GaSb MIM with CID [77]

In order to utilize photon recycling and photon recuperation effects, a backside reflector is useful. In InP-based TPV cells, a reflective layer is deposited on the backside of the substrate (as shown in Figure 21). This works particularly well because semi-insulating InP has a very low free-carrier absorption, thus avoiding parasitic optical losses. This can also be done with InGaAsSb/GaSb-based TPV cells by depositing an SiO_2/Au reflective layer on the backside of the GaSb substrate [89]. However, the relatively high absorption (even for sub-bandgap photons) in GaSb makes this design less than ideal.

Another approach for isolation and realization of backside mirrors is to eliminate the GaSb substrate altogether, as in the wafer-bonding techniques (Figure 32) and TPV cell designs (Figure 33) reported by Wang *et al.*. [90-92]. An epitaxial device structure with an epitaxial lateral conduction layer and an epitaxial etch-stop layer to facilitate selective removal of the substrate. Post-growth, a bonding layer for attachment of a “handle” superstrate, and an insulating layer for electrical isolation is deposited on the substrate. The stack is attached to a handle (*e.g.*, a GaAs wafer), the GaSb susbstrate is then removed producing a InGaAsSb TPV device structure on an electrically isolating support.

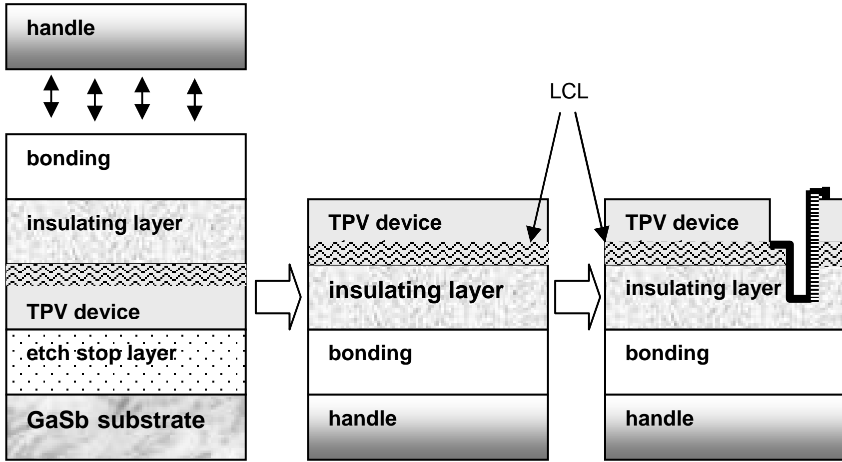


Figure 32. Wafer bonding process for InGaAsSb/GaSb TPV cells. After [77].

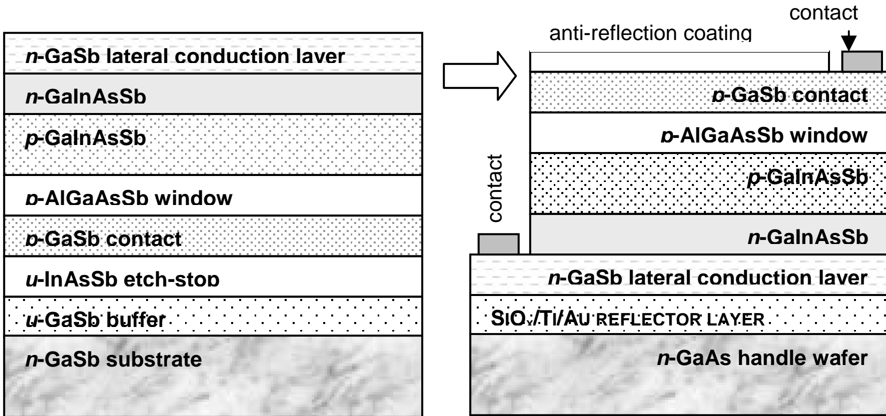


Figure 33. InGaAsSb TPV MIM produced by wafer bonding. After [91].

6.9 Tandem TPV Cells

By analogy with space solar cells and concentrator solar cells, the development of tandem TPV cell configurations would appear a logical next-step in the evolution of TPV technology. Of course, if the spectral control techniques proposed for TPV systems were highly effective, the need for tandem TPV cells would be less compelling. Further, there has been some concern that the spatial and temporal

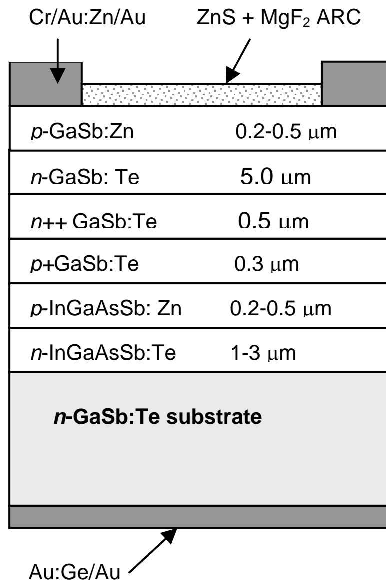


Figure 34. GaSb/InGaAsSb tandem TPV cell. After Andreev *et al.* [93].

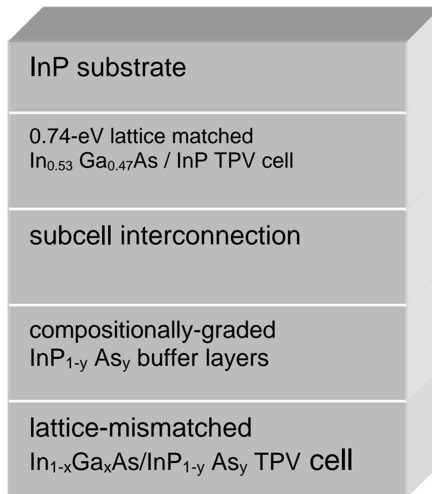


Figure 35. Inverted InGaAs/InPAs tandem TPV cell. After Wilt *et al.* [77].

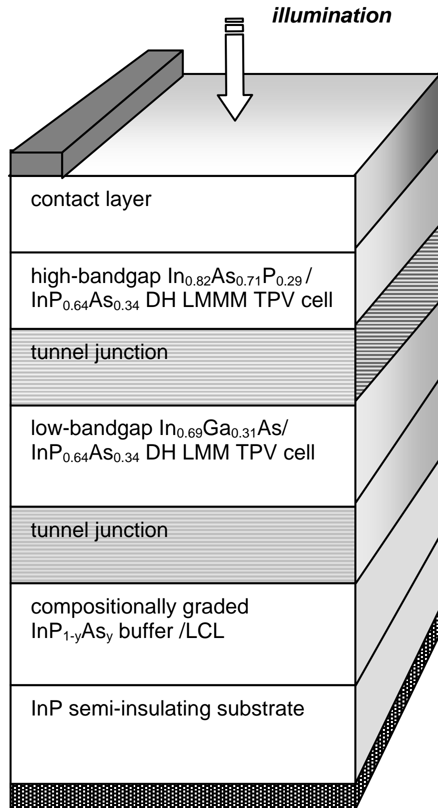


Figure 36. Schematic of 0.72/0.60 eV InGaAsP tandem TPV cell. LMM denotes lattice mismatch. LCL is lateral conduction layer. After [94].

spectral homogeneity of radiators would create current matching problems in the component TPV cells of a tandem TPV configuration. Nevertheless, there appears to be significant gains in TPV cell performance that can be realized with tandem TPV cell approaches. Andreev *et al.*. [93] reported one of the earlier versions of a tandem TPV cell in the InGaAsSb/GaSb materials system (Figure 34). This device was in part an adaptation of components first developed for tandem concentrator solar cells. Figure 35 shows one schematic version of a tandem TPV cell in the InGaAs/InP system. This inverted configuration is illuminated through the InP susbtrate. Another realization of a two-junction InGaAs/InP tandem TPV cell is shown in Figure 36, along with the current-voltage characteristics in Figure 37.

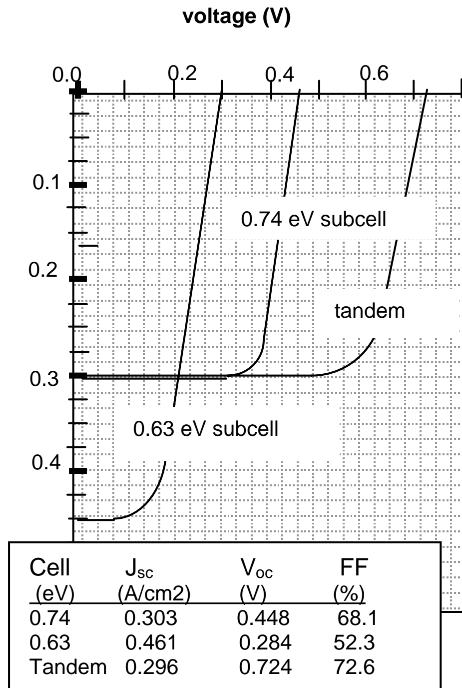


Figure 37. Current-voltage characteristics of 0.0.74/0.63 eV tandem InGaAs/InP TPV cell. After [95].

6.10 Micron-gap TPV Cells

An interesting new direction in TPV technology is the development of *micron-gap* (also called *micro-gap*) TPV configurations [96-102] in which the TPV cell is situated very close (< 1 μm) to the heated infrared emitter. Modeling results (as shown in Figure 38), and supporting experiments indicate that such a sub-micron juxtaposition of the radiation source and TPV cell can result in enhanced near-field energy transfer between the emitter and TPV cell relative to the far-field energy transfer operative in conventional TPV systems. Moreover, the spectral distribution of the transferred radiative energy is also narrowed, which is always beneficial to TPV performance.

Several potential technical difficulties are evident with this approach. First, it would appear very challenging to maintain a uniform separation on the order of 1 μm or less between the cell and emitter surface. Along these same lines, the front contact grids on the TPV cell would need to be redesigned, or better yet, eliminated altogether with an all-back-contact cell design. Moreover, parasitic absorption must be avoided that would result in cell heating. In principle, the micron-gap TPV concept is not limited to any restricted range of emitter temperature, bandgap, or particular TPV cell material, although secondary considerations related to cell

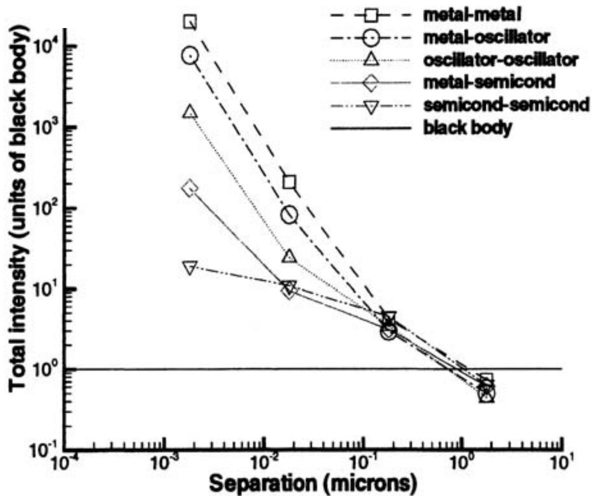


Figure 38. Total intensity of energy transfer between hot (T = 1000 ºC) and cold bodies relative to a black body at same temperature (*horizontal solid line*). From [99].

heating, coupling efficiency and micron-gap size relative to radiation wavelengths and so forth may lead to a favored bandgap for the TPV cell.

To prove the viability of micron-gap TPV, the Draper Labs / MIT group [96, 100] demonstrated enhancement of photogenerated current with a 2×2mm InAs TPV cell separated from a silicon heater chip by silicon dioxide mesa spacers. They were able to show greater than three-fold increases in photocurrent by varying the spacing between the top cell surface and heater. Interestingly, the thermal emitter temperatures investigated ranged from 348 to 408 K, producing current densities of ~50 mA/cm^2. For larger areas, a new TPV cell design is needed since the conventional front metal contact grid will interfere with the required close gap spacing, and will also provide an unwanted thermal conductive path between the TPV cell and heated thermal source.

Microgap TPV cells will require about 0.2 µm obstruction-free clearance of the front surface of the TPV cell [103]. The standard front metal contact grid, which is typically much thicker than this allowed clearance, will impede the close spacing needed to form the microgap. Further, the contact grid will cause thermal shunting between the emitter and cell, forming hot spots on the cell surface and degrading its performance. This makes the conventional front metallization of TPV cells highly problematic. All Back Contact (ABC) solar cells were originally developed for the purpose of eliminating shading losses due to the front contact grid obscuration [104]. This approach would also avoid front contact interference problems in microgap TPV cells. In GaAs ABC solar cell technology, an example of which is shown in Figure 39, the epitaxial GaAs solar cell structure is bonded to a high-temperature glass superstrate, the GaAs seeding substrate, upon which the epitaxial device structure is grown, is removed by selective etching. The back of the n-type

GaAs base layer is masked with an oxide that serves as an insulator and diffusion barrier, and patterned with openings, through which p-type channels are formed by selective zinc diffusion. The channels contact the p-type emitter. The structure is then metallized, resulting in a solar cell where all metallization contacts are from the backside of the cell. The ABC TPV designs could also exploit features and fabrication technology developed for wafer-bonded and MIM TPV cell technology.

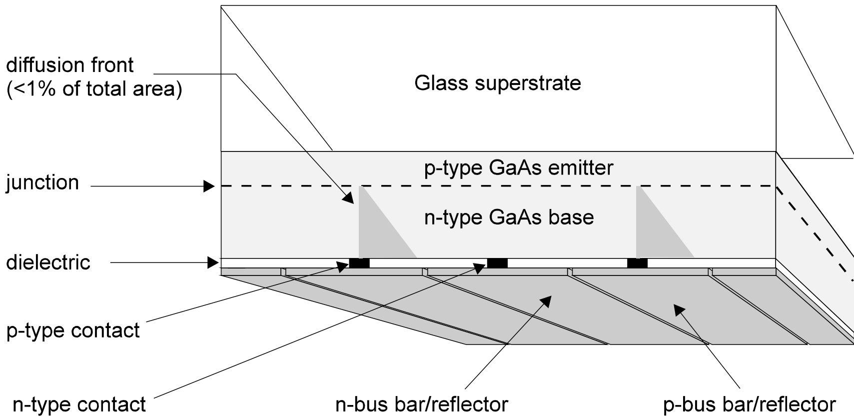


FIGURE 39. All Back Contact (ABC) GaAs solar cell with zero front grid obscuration

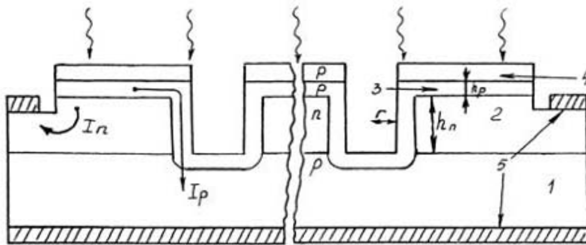


Figure 40a. Cross-sectional view of cell with no front contact grid made by etching vias and conformal diffusion. From [105].

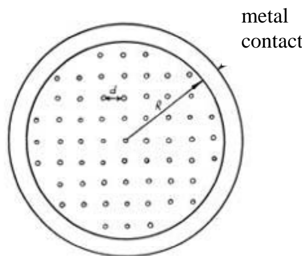


Figure 40b. Top plan view of structure shown in Figure 40a. From [105].

An alternative approach [105] to providing a unobscured front surface is shown in Figure 40. In this design, the epitaxial structure is perforated with a pattern of vias which penetrate into the underlying *n*-type base layer. A conformal diffusion layer forms a wrap-around *p-n* junction such that the emitter current can be collected through the *p*-type substrate. The base current is collected through a contact ring. This design was actually developed for GaAs concentrator solar cells, but should also be applicable to TPV cells for eliminating the front contact grid, thus facilitating formation of an unobscured microgap.TPV cell configuration.

6.11 Spectral Control, Microstructured Emitters, Filters, and Photonic Crystals

At this stage of TPV development, spectral control is now perhaps the most promising route to higher TPV system efficiencies [106,108]. Spectral control can be effected with selective emitters, reflectors, or filters. In the first decades of development, spectral control for TPV relied on selective emitters such as rare-earth oxides, and also filters, reflectors, and anti-reflection coatings made as single- or multi-layer stacks of dielectrics, semiconductors, and metals. New developments in spectral control utilize 1-D and 2-D surface-relief structures [109], such as microstructured tungsten [110,111], and 3-D photonic crystals [112,113].

The realization of effective spectral control will impact both the general viability of TPV systems, as well as the design of TPV cells. Naturally, the emission spectra, transmission characteristics of filters, and spectral characteristics of reflectors need to be matched to the spectral response of a particular TPV cell. Normally, the selective emitter or filter is developed separately from the TPV cell and they are typically considered as distinct, discrete components of the TPV system. A contrary example is the formation of surface relief structures on TPV cells for broadband anti-reflection, as has been reported by Yugami *et al.,* [64] who describe the use of electron-beam lithography and fast atom beam etching to pattern the front surface of a GaSb TPV cell with a sub-wavelength grating. To what extent it is feasible or desirable for surface gratings, photonic crystals and other microstructured elements to be monolithically integrated with and/or fabricated together with TPV device structures for purposes of spectral control is an open question.

6.12 Thermophotonics

Thermophotonics (TPX) is a newly proposed concept to achieve heat-to-electric conversion efficiencies above those of conventional TPV systems, especially for solar energy conversion [114,115]. In one such TPX system (Figure 41), the passive thermal emitter is replaced by a heated, electrically-biased light-emitting diode (LED). The LED is highly forward-biased and produces infrared luminescence which is coupled to a photovoltaic device. The heated LED extracts heat from the heat source and acts as a selective emitter (*i.e.*, narrow spectral width) for which the photovoltaic device is very efficient. As envisioned, thermophotonics approaches would permit the use of wide bandgap cells such as made in GaAs. This would represent a significant departure from the low-bandgap semiconductor device

technology currently under development for TPV systems. Nevertheless, the basic system and device concepts and supporting technology, related to coupling (thermo)photovoltaic devices to heated radiation sources, would be immediately relevant to TPX systems as well.

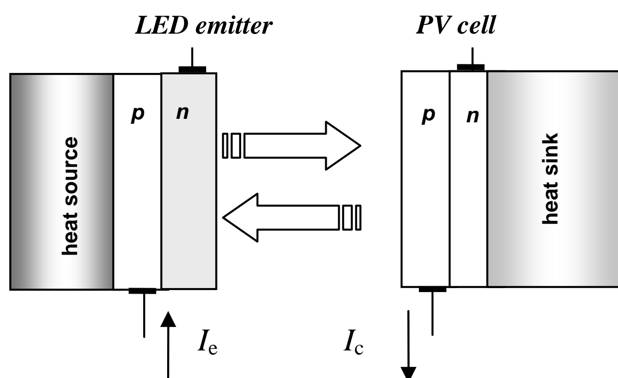


Figure 41. Thermophotonics: luminescence from a heated LED is coupled to a solar cell. After [114].

7. **Concluding Remarks and Outlook for TPV Technology**

In the late 1990s, state-of-the-art TPV cells - as represented by InGaAsSb/GaSb and InGaAs/InP heterostructures - had reached a level of sophistication comparable to that achieved with GaAs-based space solar cells, high-brightness LEDs, laser diodes, and infrared photodiodes. The application of high-efficiency solar cell design concepts to low-bandgap III-V compound TPV cells was enabled by materials and device technology first established for infrared detectors, LEDs, and lasers. Single-junction TPV cells were approaching their theoretical performance limits. There was seemingly little room for improvement in a TPV cell that had already achieved a combination of near-perfect internal quantum efficiencies for photons with energies in excess of the bandgap, fill-factors of around 70%, and open-circuit voltages at levels delimited by Auger or radiative recombination. Further significant advances in TPV would then be relegated to balance-of-systems components such as the heat source, emitter, reflectors and filters, packaging, and issues related to system reliability, safety, manufacturability and cost.

This situation proved to be brief. Unlike solar cell systems, photon recuperation and spectral control are feasible in TPV systems and can have tremendous impact on TPV system efficiency. In order to improve overall system performance, TPV cells were designed to incorporate features to exploit these phenomena. Tandem and multi-junction cells, and novel means to reduce parasitic optical absorption, provide better heat transfer and temperature control, and facilitate new radiative coupling schemes characterize much of the current activity in TPV technology. These tandem

TPV cells and monolithic interconnect modules encounter new technical problems specific to low-bandgap semiconductor materials.

TPV technology is meeting these challenges by better utilizing spectral control, photon recuperation, and photon recycling; by addressing the exacting demands related to heat transfer, high current densities and high injection levels; and by meticulous control of optical absorption, transmission, and reflection in devices. To this end, TPV researchers are pioneering innovations that have no specific antecedent in other semiconductor device technologies, as for example, integrated reflectors and filters, microstructured surfaces, wafer bonding for series interconnection and/or integrated backside reflectors, isolation for monolithic interconnection, spectral control using photonic bandgap materials, and micron-gap radiative coupling of system elements. The flow of innovation may well reverse in the next few years: technologists may be eager to apply new concepts first developed for TPV cells to LEDs, laser diodes, detectors, and solar cells. Along these lines, recent proposals for thermophotonics and high-brightness LEDs may prefigure a kind of convergence for semiconductor energy conversion devices, wherein these optoelectronic devices share many common features with regard to the coupling and conversion of electrical and radiant energy. These systems would utilize a common 'tool box' of materials, device designs, and operating principles exploiting physical effects enabled by surface relief microstructuring, photonic bandgaps, quantum well phenomena, low-dimensional devices, sub-micron-gap coupling of radiation fields, and the like.

TPV technology appears on track to achieve efficiencies in the 20-25% range within the next few years [32], and for at least several applications, TPV seems to have established a clear edge over competing technologies such as thermoelectric and thermionic generators. (On the other hand, thermoelectric devices have a distinct advantage over TPV in that they can be used for active cooling, and therefore, will receive additional research and development efforts for their use as solid-state refrigerators.)

There is no clear "winner" among the two competing high-efficiency TPV technologies (InGaAsSb/GaSb and InGaAs/InAsP/InP). Several years ago, each technology had its unique advantages as well as its limitations. Lattice-matched InGaAsSb could be tuned to bandgaps as low as 0.52 eV, but MIM approaches were hindered by the lack of semi-insulating GaSb. On the other hand, while MIM devices were pioneered with InGaAs/InP structures, extending the spectral response in the InGaAs/InP system seemed limited by lattice-matching constraints and possible reliability issues. However, improved designs and epitaxy techniques appear to be overcoming this limitation. By the same token, wafer-bonding techniques are providing for the realization of InGaAsSb MIMs with efficient photon recuperation capabilities. At this stage, the InGaAs/InAsP/InP system seems to have a lead in the development of tandem MIM devices, but the author is not aware of any fundamental difficulties in following a similar path to make III-V antimonide-based tandem TPV cells, especially if progress is made on low-bandgap InAsSbP TPV cells.

The last decade has witnessed a "coming of age" of photovoltaics and windpower. Whether TPV follows suit and expands into distributed and portable power, vehicle propulsion, and co-generation from industrial waste heat; or else remains a specialized military technology for niche applications is an open question.

Like virtually all alternative energy technologies of the last half century, only time will tell their fate. A safer prediction is that developments in TPV technology will have a significant and wider impact on many areas of semiconductor optoelectronics, as well as conventional solar cells.

Appendix. Modeling of InGaAsSb TPV Cells

A1 Materials Properties

InGaAsSb TPV cells are comprised of InGaAsSb alloy epitaxial layers lattice-matched to GaSb substrates with wide-bandgap lattice-matched epitaxial AlGaAsSb cladding layer(s) that serve as a transparent front-side ‘window’ to passivate the emitter. The basic bandstructure-derived semiconductor properties of the InGaAsSb (and InAsSbP) alloys have been studied by Joullié *et al.* [116]. The bandgap of the $In_xGa_{1-x}As_ySb_{1-y}$ quaternary alloy is given by

$$E(x,y) = 0.726 - 0.961x - 0.501y + 0.08xy + 0.451x^2 + 1.2y^2 + 0.021x^2y + 0.62\cdot xy^2 \tag{A1}$$

Other interpolation schemes for estimating bandgaps of ternary and quaternary alloys used in TPV are also available [117,118]. For example, for a quaternary alloy of the form $A_{1-x}B_xC_yD_{1-y}$, a quaternary material property $Q(x,y)$ as

$$Q(x,y) = (1-x)yB_{\mathrm{AC}} + (1-x)(1-y)B_{\mathrm{AD}} + xyB_{\mathrm{BC}} + x(1-y)B_{\mathrm{BD}} + C_{A/B}x(1-x) + C_{C/D}y(1-y) \tag{A2}$$

or more generally,

$$Q(x,y) = (1-x)yB_{\mathrm{AC}} + (1-x)(1-y)B_{\mathrm{AD}} + xyB_{\mathrm{BC}} + x(1-y)B_{\mathrm{BD}} + C_{\mathrm{AB/C}}x(1-x)y + C_{\mathrm{A/CD}}xy(1-y) + C_{\mathrm{AB/D}}x(1-x)(1-y) + C_{B/CD}(1-x)(1-y) \tag{A3}$$

where the various C terms are interaction or bowing parameters.

For effective masses and mobilities, a more appropriate interpolation scheme is [117]

$$\frac{1}{m(x,y)} = \frac{(1-x)y}{m_{\mathrm{AC}}} + \frac{(1-x)(1-y)}{m_{\mathrm{AD}}} + \frac{xy}{m_{\mathrm{BC}}} + \frac{x(1-y)}{m_{\mathrm{BD}}} \tag{A4}$$

The lattice matching condition to GaSb imposes the additional constraint that x and y are related as

$$y = \frac{0.867}{(1-0.048x)} \tag{A5}$$

Most TPV work has focused on alloys close to the composition $In_{0.15}Ga_{0.85}As_{0.15}Sb_{0.85}$. which have bandgaps ranging from 0.53 to 0.57 eV. There is a miscibility gap precluding alloys with lower bandgaps. As this composition is

close to that of the better-characterized binary GaSb, some measured properties of GaSb can stand-in for the InGaAsSb alloy in cases where no data are available.

A basic electrical property of the semiconductor alloy is the intrinsic carrier concentration [61]

$$n_i = 2\left(2\cdot\pi\cdot m_0 kT_C/h^2\right)^{3/2}\sqrt{\left(m_L^{3/2}e^{-\Delta\Gamma L/k\cdot T_C}+m_\Gamma^{3/2}\right)\left(m_{hh}^{3/2}+m_{lh}^{3/2}\right)} \times e^{-E_G/2\cdot k\cdot T_C}e^{-\Delta E_{BGN}/2\cdot k\cdot T_C} \quad \text{(A6a)}$$

$$= 2.51\times10^{19}\left(\frac{T_C}{300}\right)^{3/2}\sqrt{\left(m_L^{3/2}e^{-\Delta\Gamma L/k\cdot T_C}+m_\Gamma^{3/2}\right)\left(m_{hh}^{3/2}+m_{lh}^{3/2}\right)} \times e^{-E_G/2\cdot k\cdot T_C}e^{-\Delta E_{BGN}/2\cdot k\cdot T_C} \quad \text{(A6b)}$$

For GaSb, m_Γ = 0.047, m_L = 0.58, m_{lh} = 0.053, m_{hh} = 0.32, in units of m_0; and $\Delta\Gamma L$ (the energy difference between the minima of the L and Γ bands) = 0.08 eV [61]. More accurate values of the effective masses and band parameters can be found by interpolating values for the binary constituents (GaAs, InAs, InSb, and GaSb) of the quaternary alloy [119, 120]. Heavy doping effects include bandgap narrowing ΔE_G which can be modeled generally as

$$\Delta E_G = AN^{1/3} + BN^{1/4} + CN^{1/2} \text{ [eV]} \quad \text{(A7)}$$

A simpler expression for bandgap narrowing in n-type and p-type GaSb [62] is

$$\Delta E_{g,n} = 1.36\times10^{-8} N_D$$
$$\Delta E_{g,p} = 10^{-8} N_A \quad \text{(A8a)(A8b)}$$

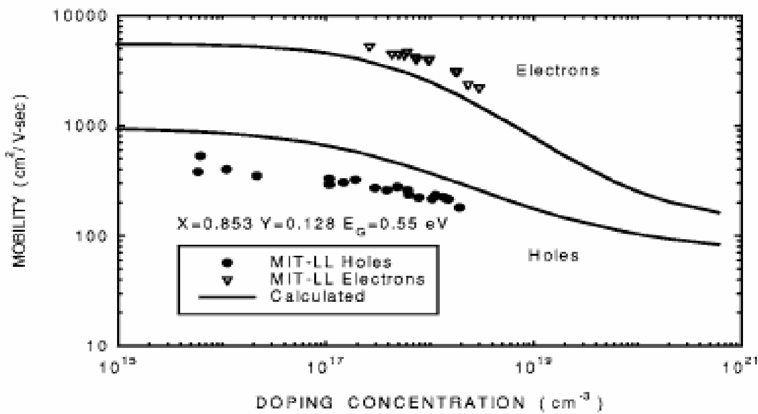


Figure A1. The electron and hole mobilities for $In_{1-x}Ga_xAs_{1-y}Sb_y$ ($x \cong 0.2$; $y \cong 0.9$; $E_G \cong$ 0.53 eV). From Borrego *et al.*. [31]

The electron and hole mobilities for $In_{1-x}Ga_xAs_{1-y}Sb_y$ ($x \cong 0.2$; $y \cong 0.9$; $E_G \cong 0.53$ eV) as a function of doping are plotted in Figure A1. The doping dependence of electron or hole mobility can be well fit to an equation of the form [121]

$$\mu = \mu_{\min} + \frac{\mu_{\max} - \mu_{\min}}{\left(1 + \frac{N}{N_{ref}}\right)^{A}} \tag{A9}$$

Some fitting parameters for Equation (24) are given in Table A1.

Table A1. Mobility parameters for some III-V binaries. From Zierak *et al.*, (1997)

Binary, Carrier type	$\mu_{\min}$ [cm^2/V-s]	$\mu_{\max}$ [cm^2/V-s]	$N_{\rm ref}$ [cm^{-3}]	***A***
InSb, holes	100	750	6×10^{17}	0.6
InSb, electrons	5000	7800	7×10^{16}	0.7
GaSb, holes	70	1400	2×10^{17}	0.5
GaSb, electrons	100	4500	8×10^{17}	0.8

A mobility equation that includes the effects of temperature and doping has been given by Martín and Algora [62]:

$$\mu(N,T) = \mu_{\min,e/h} + \frac{\mu_{\max,e/h}\left(\frac{300}{T}\right)^{\theta_{1,e/h}}}{1 + \left[\frac{N_{D/E}}{N_{ref,e/h}\left(\frac{T}{300}\right)^{\theta_{2,e/h}}}\right]^{\alpha_{e/h}}} \tag{A10}$$

The electron and hole mobility for GaSb can be calculated using Equation (28) and the parameters in **Table A2**.

Table A2. Mobility parameters from Zierak *et al.*, (1997)

Carrier type	$\mu_{\min}$ [cm^2/V-s]	$\mu_{\max}$ [cm^2/V-s]	$N_{\rm ref}$ [cm^{-3}]	θ_1	θ_2	α
Electrons	1050	5650	2.8×10^{17}	2.0	2.8	1.05
Holes	190	875	9.0×10^{17}	1.7	2.7	0.65

The minority carrier diffusion coefficients are related to carrier mobilities as

$$D_p = \left(\frac{kT}{q}\right)\mu_p \tag{A11a}$$

$$D_n = \left(\frac{kT}{q}\right)\mu_n \tag{A11b}$$

The optical absorption of photons with energies near or above the bandgap energy results in photogeneration of electrons and holes. Optical absorption is characterized by a wavelength-dependent optical absorption coefficient $\alpha(\lambda)$. Further, a distinction can be made between absorption that generates electron-hole pairs $\alpha^{eh}(\lambda)$ and extraneous absorption $\alpha_{FC}(\lambda)$ such as due to free carriers, that attenuates the photon flux but does not contribute to the photogeneration of minority carriers. In general, since absorption is a function of bandgap, doping, and electric field, each region (window, emitter, space charge region, base) of the heterostructure will be characterized by distinct absorption coefficients. We emphasize this point because free-carrier absorption, especially in a thick substrate and heavily-doped emitter, can be important when optimizing device structures for photon recuperation. For the heterostructure denoted in Figure 7b, carrier generation functions can be written as:

$$G_W(\lambda, y) = \alpha_W^{eh}(\lambda)\left[1 - R(\lambda)\right]\phi(\lambda)\exp\left\{-\left[\alpha_W^{eh}(\lambda) + \alpha_W^{fc}(\lambda)\right]y\right\} \tag{A12a}$$

$$\begin{aligned} G_E(\lambda, y_E) = {} & \alpha_E^{eh}(\lambda)\left[1 - R(\lambda)\right]\phi(\lambda)\exp\left\{-\left[\alpha_W^{eh}(\lambda) + \alpha_W^{fc}(\lambda)\right]H_W\right\} \\ & \times \exp\left\{-\left[\alpha_E^{eh}(\lambda) + \alpha_E^{fc}(\lambda)\right]y_E\right\} \end{aligned} \tag{A12b}$$

$$\begin{aligned} G_{SCR}(\lambda, y_{SCR}) = {} & \alpha_{SCR}^{eh}(\lambda)\left[1 - R(\lambda)\right]\phi(\lambda)\exp\left\{-\left[\alpha_W^{eh}(\lambda) + \alpha_W^{fc}(\lambda)\right]H_W\right\} \\ & \times \exp\left\{-\left[\alpha_E^{eh}(\lambda) + \alpha_E^{fc}(\lambda)\right]H_E\right\}\cdot\exp\left\{-\left[\alpha_{SCR}^{eh}(\lambda) + \alpha_{SCR}^{fc}(\lambda)\right]y_{SCR} \right. \end{aligned} \tag{A12c}$$

$$\begin{aligned} G_B(\lambda, y_B) = {} & \alpha_B^{eh}(\lambda)\left[1 - R(\lambda)\right]\phi(\lambda)\exp\left\{-\left[\alpha_W^{eh}(\lambda) + \alpha_W^{fc}(\lambda)\right]H_W\right\} \\ & \times \exp\left\{-\left[\alpha_E^{eh}(\lambda) + \alpha_E^{fc}(\lambda)\right]H_E\right\}\cdot\exp\left\{-\left[\alpha_{SCR}^{eh}(\lambda) + \alpha_{SCR}^{fc}(\lambda)\right]W_{SCR} \right. \\ & \times \exp\left\{-\left[\alpha_B^{eh}(\lambda) + \alpha_B^{fc}(\lambda)\right]\cdot_B\right\} \end{aligned} \tag{A12d}$$

The free-carrier absorption has the form

$$\alpha_{FC}(\lambda, N) = \alpha_{FC.0}\left(\frac{N}{N_0}\right)\left(\frac{\lambda}{\lambda_G}\right)^2 \tag{A13}$$

where N is the doping and λ_G is the wavelength corresponding to the bandgap energy. Although no data is generally available for the low-bandgap III-V alloys, by analogy with the better-studied wider-bandgap III-V compounds, it is reasonable to estimate $N_0 \sim 10^{18}$ cm^{-3} and $\alpha_{FC.0} \sim 10$ to 100 cm^{-1}.

Most TPV materials are direct bandgap semiconductors, in which the interband absorption that creates electron-hole pairs can be expressed as [26]

$$\alpha(\lambda) = A_D \left[hc/\lambda - E_G\right]^{1/2} + A_{Ex} \exp\left[\frac{hc/\lambda - E_G}{E_0}\right] \tag{A14}$$

The first term in the above equation is the well-known relationship for optical absorption due to direct bandgap transitions. The second term in the above equation is the Urbach or exponential "tail" in absorption often observed at the energies slightly less than the band edge threshold, especially in heavily-doped materials. The window layer is often an indirect bandgap semiconductor material in which case the optical absorption follows an equation of the form

$$\alpha(\lambda) = A_{ID} \left(hc/\lambda - E_G\right)^2 \tag{A15}$$

Figure A2 shows absorption coefficient data for an InGaAsSb alloy (E_G = 0.53 eV) lattice-matched to GaSb.

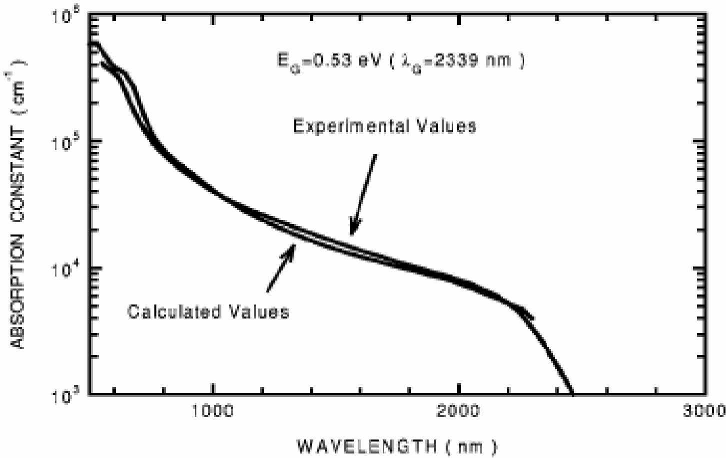


Figure A2. Optical absorbtion data for InGaAsSb alloy. From Borrego *et al.* [31]

The minority carrier lifetime for electrons and holes is a function of quaternary alloy composition, temperature, doping level, and injection level. For purposes of modeling device performance, an effective minority carrier recombination lifetime, τ_{eff} , expression of the form is determined as

$$\frac{1}{\tau_{eff}} = \frac{1}{\tau_{rad}} + \frac{1}{\tau_{Auger}} + \frac{1}{\tau_{SRH}} \tag{A16}$$

This represents contributions from radiative recombination, Auger recombination, and defect-mediated Shockley-Read-Hall recombination *via* mid-bandgap states, for each of which a lifetime can be assigned. These recombination mechanisms exhibit the following doping dependence

$$\frac{1}{\tau_{eff}} = BN + CN^2 + \frac{1}{\tau_{SRH}} \tag{A17}$$

where N is the doping level. Kumar *et al.*. [43] experimentally determined a best fit for B and C in 0.55 eV bandgap InGaAsSb as $B = (3 \pm 1.5) \times 10^{-11}$ cm^3/s and $C = (1 \pm 0.4) \times 10^{-28}$ cm^6/s.

The minority carrier diffusion length is related to the minority carrier lifetime as $L = \sqrt{D\tau}$. The diffusion length has the appealing attribute that it is interpreted as the average distance a photogenerated or injected minority carrier will travel before recombining. As a design rule of thumb, if the minority carrier diffusion length is at least twice the thickness of the quasi-neutral region in which it is photogenerated, it has a high probability of being collected by the junction and contributing to the photogenerated current of the TPV cell, provided that interfacial (surface) recombination is also sufficiently low. Typical minority carrier diffusion lengths in InGaAsSb range from 1 to 10 μm.

Reported surface recombination velocities observed in InGaAsSb heterostructures are in the 1000 cm/s range. It is estimated that the interface recombination at an AlGaAsSb/InGaAsSb interface can be as low as 100 cm/s [41-46].

A2 TPV Device Model

The built-in voltage of the *p-n* junction is given by

$$V_0 = \frac{k_B \cdot T}{q} \ln\left\{\frac{N_E \cdot N_B}{n_i^2}\right\} \quad \text{(A18)}$$

The width W of the depletion or space-charge region (SCR) surrounding the *p-n* junction is

$$W = \left[\frac{2 \cdot \varepsilon \cdot V_0}{q}\left(\frac{N_E + N_B}{N_E \cdot N_B}\right)\right]^{1/2} \quad \text{(A19)}$$

The penetration of the space-charge region into the emitter (X_E) and base (X_B) are

$$X_E = W \cdot \frac{N_B}{N_E + N_B} \quad \text{(A20a)}$$

$$X_B = W \cdot \frac{N_E}{N_E + N_B} \quad \text{(A20b)}$$

giving an effective (quasi-neutral region) emitter thickness of

$$H'_E = H_E - X_E \quad \text{(A21)}$$

The minority carrier transport in the base and emitter, which determines the quantum efficiency and diode dark current, are found by solving the linear minority carrier transport equations in the quasi-neutral regions of the emitter and base. In the emitter

$$D_E \frac{d^2 \Delta u_E}{d{y_E}^2} - \frac{\Delta u_E}{\tau_E} = -G_E(\lambda, y_E) \tag{A22}$$

subject to the boundary conditions

$$-D_E \left. \frac{d\Delta u_E}{dy_E} \right|_{y_E=0} = S_E \, \Delta u_E \big|_{y_E=0} \tag{A23}$$

$$\Delta u_E \big|_{y_E=H'_E} = \frac{n_{i,E}^2}{N_E} \left[\exp\left(\frac{qV}{kT} \right) - 1 \right] \tag{A24}$$

Likewise, in the base

$$D_B \frac{d^2 \Delta u_B}{d{y_B}^2} - \frac{\Delta u_B}{\tau_B} = -G_B(\lambda, y_B) \tag{A25}$$

subject to the boundary conditions

$$-D_B \left. \frac{d\Delta u_B}{dy_B} \right|_{y_B=H'_B} = S_B \, \Delta u_B \big|_{y_B=0} \tag{A26}$$

$$\Delta u_B \big|_{y_{B=0}} = \frac{n_{i,B}^2}{N_B} \left[\exp\left(\frac{qV}{kT} \right) - 1 \right] \tag{A27}$$

where Δu is the excess minority carrier concentration due to optical absorption or injection.

Equations (A22) and (A25) are solved as

$$\Delta u_E = A_E \cosh\left(\frac{y_E}{L_E} \right) + B_E \sinh\left(\frac{y_E}{L_E} \right) + \frac{\tau_E G_E(\lambda, y_E)}{\left(\alpha_E^{lh} L_E \right)^2 - 1} \tag{A28}$$

$$\Delta u_B = A_B \cosh\left(\frac{y_B}{L_B} \right) + B_B \sinh\left(\frac{y_B}{L_B} \right) + \frac{\tau_B G_B(\lambda, y_B)}{\left(\alpha_B^{lh} L_B \right)^2 - 1} \tag{A29}$$

The constants A and B are found using the boundary conditions specified above. The current contribution from the emitter is given by

$$J_E(\lambda) = \pm q D_E \left. \frac{d\Delta u_E}{dy_E} \right|_{y=H'_E} \tag{A30}$$

and from the base

$$J_B(\lambda) = \pm q D_E \left. \frac{d\Delta u_B}{dy_B} \right|_{y_B=0} \quad \text{(A31)}$$

The space charge region also contributes to the current according to

$$J_{SCR}(\lambda) = q \int_0^W G_{SCR}(\lambda, y_{SCR}) dy_{SCR} \quad \text{(A32)}$$

It is convenient to separate the above contributions into a light-generated short-circuit ($V = 0$) current and a voltage-dependent "dark" current ($\phi(\lambda) = 0$). When this is done, the short-circuit current of the TPV cell is

$$J_{sc} = q \int_0^\infty \phi(\lambda) QE(\lambda) d\lambda \quad \text{(A33)}$$

The quantum efficiency is the sum of three contributions

$$QE_{ext}(\lambda) = QE_E(\lambda) + QE_{SCR}(\lambda) + QE_B(\lambda) \quad \text{(A34)}$$

wherein

$$QE_E(\lambda) = \left\{ \frac{[1 - R(\lambda)] \cdot \alpha(\lambda) \cdot L_E}{(\alpha(\lambda) \cdot L_E)^2 - 1} \right\} \cdot$$

$$X \left\{ \begin{array}{l} \dfrac{\left[\dfrac{S_E \cdot L_E}{D_E} + \alpha(\lambda) \cdot L_E\right] - \exp(-\alpha(\lambda) \cdot x_j) - \exp(-\alpha(\lambda) \cdot x_j) \cdot \left[\dfrac{S_E \cdot L_E}{D_E} \cosh\left(\dfrac{x_j}{L_E}\right) + \sinh\left(\dfrac{x_j}{L_E}\right)\right]}{\dfrac{S_E \cdot L_E}{D_E} \cdot \sinh\left(\dfrac{x_j}{L_E}\right) + \cosh\left(\dfrac{x_j}{L_E}\right)} \ldots \\ \ldots - \alpha(\lambda) \cdot L_E \cdot \exp(-\alpha(\lambda) \cdot x_j) \end{array} \right\} \quad \text{(A35)}$$

$$QE_B(\lambda) = \left\{ \frac{[1 - R(\lambda)] \cdot \alpha(\lambda) \cdot L_B}{(\alpha(\lambda) \cdot L_B)^2 - 1} \cdot \exp[-\alpha(\lambda) \cdot (x_j + W)] \right\} \cdot X$$

$$\left\{ \alpha(\lambda) \cdot L_B - \frac{\left[\dfrac{S_B \cdot L_B}{D_B} \cdot \cosh\left(\dfrac{H'}{L_B}\right) - \exp(-\alpha(\lambda) \cdot H)\right] + \sinh\left(\dfrac{H'}{L_B}\right) + \alpha(\lambda) \cdot L_B \cdot \exp(-\alpha(\lambda) \cdot x_j)}{\dfrac{S_B \cdot L_B}{D_B} \sinh\left(\dfrac{H'}{L_B}\right) + \cosh\left(\dfrac{H'}{L_B}\right)} \right\} \quad \text{(A36)}$$

$$QE_{SCR}(\lambda) = [1 - R(\lambda)] \cdot \exp\left[-\left(\alpha_W^{lh}(\lambda) + \alpha_W^{fc}(\lambda)\right) \cdot H_W\right] \cdot \exp\left[-\left(\alpha_E^{lh}(\lambda) + \alpha_E^{fc}(\lambda)\right) \cdot H_E\right]$$
$$\times \left\{1 - \exp\left[-\left(\alpha_W^{lh}(\lambda) + \alpha_W^{fc}(\lambda)\right) \cdot W\right]\right\} \quad \text{(A37)}$$

The above equations give the *external* quantum efficiency of the TPV cell. When $R(\lambda)$ is set identically equal to zero, and the effects of contact grid shading are accounted for, the *internal* quantum efficiency is determined.

The "dark" current is called the injection-diffusion current and is given by

$$J_{01} = q \cdot \frac{D_E}{L_E} \frac{n_i^2}{N_E} \left[\frac{\left(\frac{S_E \cdot L_E}{D_E}\right) \cosh\left(\frac{x_j}{L_E}\right) + \sinh\left(\frac{x_j}{L_E}\right)}{\left(\frac{S_E \cdot L_E}{D_E}\right) \sinh\left(\frac{x_j}{L_E}\right) + \cosh\left(\frac{x_j}{L_E}\right)} \right] + q \cdot \frac{D_B}{L_B} \frac{n_i^2}{N_B} \left[\frac{\left(\frac{S_B \cdot L_B}{D_B}\right) \cosh\left(\frac{H}{L_E}\right) + \sinh\left(\frac{H}{L_E}\right)}{\left(\frac{S_B \cdot L_B}{D_B}\right) \sinh\left(\frac{H'}{L_E}\right) + \cosh\left(\frac{H}{L_E}\right)} \right] \quad \text{(A38)}$$

An additional component of the dark current is due to recombination in the space-charge region and is given by

$$J_{rg} = \frac{q \cdot n_i \cdot W}{\sqrt{\tau_n \cdot \tau_p}} \cdot \frac{2 \cdot \sinh\left(q \cdot V / 2kT\right)}{q \cdot \left(V_0 - V\right) / k \cdot T} \quad \text{(A39)}$$

The resulting closed-form equations for quantum efficiency, fill-factor, dark current, and open-circuit voltage, along with the optical properties and minority carrier transport and recombination models discussed herein, form the basis of accurate computer one-dimensional models used to predict TPV cell performance including spectral response and current-voltage characteristics. The models can also be used to evaluate and optimize new designs, assess the sensitivity of TPV cell performance to variations in materials properties or design specifications, identify losses, compare competing cell technologies, and predict the ultimate potential of specific devices. The analysis described here can also facilitate the use of more sophisticated computer models such as PC-1D and two- or three-dimensional numerical models of TPV performance.

References

1. R.E. Nelson, TPV Systems and State of the Art Development, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 3-17
2. T.J. Coutts, A review of progress in thermophotovoltaic generation of electricity, *Renewable and Sustainable Energy Reviews,* **3** (1999) 77-184.
3. E. Kittl, Unique correlations between blackbody radiation and optimum energy gap for a photovoltaic conversion device, *Conf. Rec. 10th IEEE Photovoltaics Specialists Conf.* (1974) 103-106.
4. B. Wernsman, R.R. Siergiej, S.D. Link *et al.*., Greater than 20% radiant heat conversion efficiency of a thermophotovoltaic radiator/module system using reflective spectral control, *IEEE Trans. Electron Devices* **51, 3** (2004) 512-515.
5. R.J. Wehrer, M.W. Wanlass, D. Taylor *et al.*., 0.52 eV InGaAs/InPAs thermophotovoltaic cells, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, and J. Luther, eds., American Institute of Physics, 2004, 445-452.
6. T.J. Coutts, Thermophotovoltaic generation of electricity, Ch. **11** in *Clean Electricity from Photovoltaics*, M.D, Archer and R. Hill, eds. (Imperial College Press, 2001).
7. V.M. Andreev, Solar cells for TPV converters, in *Next Generation Photovoltaics: High Efficiency though Full Spectrum Utilization*, A. Marti and A. Luque, eds., Inst. Physics Publishing, Bristol, (2004) 246-273.

8. L. Broman, Thermophotovoltaics biography, *Progress in Photovoltaics: Research and Applications*, **3** (1995) 65-74.
9. J.C. Fan, B.-R. Tsaur, and B.J. Palm, Optimal design of high efficiency tandem cells, *Conf. Record 16th IEEE Photovoltaics Specialists Conf.* (1982) 692-701.
10. M. Nell and A.M. Barnett, The spectral p-n junction model for tandem solar cells, *IEEE Trans. Electron Devices* **ED-34** (1987) 257-266.
11. M. W. Wanlass, K.A. Emery, T.A. Gessert, Practical considerations in tandem cell modeling, *Solar Cells*, **27** (1989) 191-204.
12. G. Charache, GaSb-related materials and devices, *Thermophotovoltaic Generation of Electricity, Fourth NREL Conf.*, American Inst. Physics Conf. Proc., **460**, T.J. Coutts, C.S. Allman, and J.P. Benner, eds., 1998, pp 10-11.
13. R. Arhenkiel, Minority carrier lifetime in compound semiconductors, Ch. **1** in *Current Topics in Photovoltaics*, vol. **3**, T.J. Coutts, J.D. Meakin, eds., (Academic Press, 1988)
14. J.E. Parrott, Radiative recombination and photon recycling in photovoltaic cells, *Solar Energy Materials* **30** (1993) 221-231.
15. K.L. Miller, H.Z. Fardi, R.E. Hayes, Effect of multiple reflection propagation on photon recycling in GaAs/AlGaAs double heterostructures, *J. Applied Physics* **75, 12** (1994) 8158-8162.
16. F. Baldasaro, J.E. Raynolds, G.W. Charache *et al.*., Thermodynamic analysis of thermophotovoltaic efficiency and power density tradeoffs, *J. Applied Physics* **89, 6** (2001) 3319-3327.
17. L.D. Woolf, Optimum efficiency of single and multiple bandgap cells in thermophotovoltaic energy conversion, *Solar Cells: Their Science, Technology, Applications, and Economics*, **19, 1** (1986) 19-38.
18. S. Wojtczuk, Low bandgap InGaAs thermophotovoltaic cells, *Proc. Intersoc. Energy Convs. Eng. Conf.* **2** (1996) 974-978.
19. J.L. Gray and A. El-Husseini, A simple parametric study of TPV system efficiency and output power density including a comparison of several TPV materials, *Second NREL Conf. on Thermophotovoltaic Generation of Electricity*, J.P. Benner, T.J. Coutts, D.S. Grinley, eds. American Inst. Phys. Conf. Proc. **358** (1996) 3-15.
20. G.D. Cody and T. Tiedje, The potential for utility scale photovoltaic technology in the developed world: 1990-2010 in *Energy and the Environment*, B. Abeles, A.J. Jacobson, and P.Sheng, eds., (World Scientific, Singapore, 1992) 147-217.
21. G.D. Cody, Theoretical maximum efficiencies for thermophotovoltaic devices, *Thermophotovoltaic Generation of Electricity, Fourth NREL Conf.* T.J. Coutts, J.P. Benner, and C.S. Allman, eds., American Inst. Physics Conf. Proc. **460** (1999) 58-67.
22. T.J. Coutts and J.S. Ward, Thermophotovoltaic and photovoltaic conversion at high-flux densities, *IEEE Trans. Electron Devices* **46, 1**0 (1999) 2145-2153.
23. J.J. Loferski, Theoretical considerations governing the choice of the optimum semiconductor for photovoltaic solar energy conversion, *J. Applied Physics* **27** (1955) 777-784.
24. W. Shockley and H. Quiesser, Detailed balance limit of efficiency of *p-n* junction solar cells, *J. Applied Physics* **32** (1961) 510-519.
25. H.J. Hovel, *Solar Cells*, vol. **11** of *Semicoductors and Semimetals*, R.K. Willardson and A.C. Beer, eds. (Academic Press, New York, 1975).
26. J.A. Hutchby and R.L. Fudurich, Theoretical analysis of $Al_xGa_{1-x}As$-GaAs graded band-gap solar cell, *J. Applied Physics* **47, 7** (1976) 3140-3151.
27. S. Wojtczuk, P. Colter, G. Charache, D. DePoy, Performance status of 0.55 eV InGaAs thermophotovoltaic cells, *Thermophotovoltaic Generation of Electricity, Fourth NREL Conf.*, American Inst. Physics Conf. Proc., **460**, T.J. Coutts, C.S. Allman, and J.P. Benner, eds., 1999, 417-426

28. V.B. Khalfin, D.Z. Garbuzov, H. Lee *et al.*., Interfacial recombination in In(Al)GaAsSb/GaSb thermophotovoltaic cells, *Thermophotovoltaic Generation of Electricity, Fourth NREL Conf.*, American Inst. Physics Conf. Proc., **460**, T.J. Coutts, C.S. Allman, and J.P. Benner, eds., 1999, 247-255
29. G. Charache, P.F. Baldasaro, L.R. Danielson *et al.*., InGaAsSb thermophotovoltaic diode: physics equation, *J. Applied Physics* **85, 4** (1999) 2247-2252.
30. B. Wernsman, R.J. Wehrer, D.M. Wilt *et al.*., Effect of graded doping on MIM performance, *Thermophotovoltaic Generation of Electricity, Fifth Conf.*, T.J. Coutts, G. Gauzzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 488-497.
31. J.M. Borrego, C.A. Wang, P.S. Dutta *et al.*., Performance limits of low bandgap thermophotovoltaic antimonide-based cells for low temperature radiators, *Thermophotovoltaic Generation of Electricity, Fifth Conf.*, T.J. Coutts, G. Gauzzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 498-507.
32. P.F. Baldasaro, M.W. Dashiell, J.E. Oppenlander *et al.*., System performance projections for TPV energy conversion, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** 61-70.
33. M.W. Dashiell, J.F. Beausang, G. Nichols *et al.*., 0.52 eV quaternary InGaAsSb thermophotovoltaic diode technology, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 404-414.
34. M. Borrego, S. Saroop, R.J. Gutmann *et al.*., Photon recycling and recombination processes in 0.53 eV p-type InGaAsSb, *J. Applied Physics* **89, 7** (2001) 3573-3759.
35. E. F. Shubert, *Light-Emitting Diodes*, Cambridge University Press, 2003.
36. B.L. Gelmont and V.I. Ivanov-Omskii, Electron and hole recombination in narrow gap semiconductors, in *Semiconductor Physics*, V.M. Tuchkevich and V. Ya. Frenkel, eds., (Consultants Bureau, New York, 1986, 169-189.
37. A. Rogalski, Band-to-band recombination in $Ga_xIn_{1-x}Sb$, *Infrared Physics* **27, 6** (1987) 353-360.

38. Y. Tian, T. Zhou, B. Zhang, H. Jaing, and Y. Jin, The effect of Auger mechanism on n+-p GaInAsSb infrared photovoltaic detectors, *IEEE Trans. Electron Devices* **46, 4** (1999) 656-660.
39. R.K. Ahrenkile, R. Ellington, S. Johnston, J. Webb, J. Carapella, and M. Wanlass, Recombination lifetime of In_xGa_{1-x} As alloys used in thermophotovoltaic converters, *Thermophotovoltaic Generation of Electricity, Fourth NREL Conf.*, American Inst. Physics Conf. Proc., **460**, T.J. Coutts, C.S. Allman, and J.P. Benner, eds., 1999, 282-289
40. S. Anikeev, D. Donetsky, G. Belenky, S. Luryi, C.A. Wang, J.M. Borrego, and G. Nichols, Measurement of the Auger recombination rate in p-type 0.54 eV GaInAsSb by time resolved photoluminescence, *Applied Physics Letters,* **83, 16** (2003) 3317-3319.
41. D. Donestsky, S. Anikeev, N. Gu *et al.*., Analysis of recombination processes in 0.5-0.6 eV epitaxial GaInAsSb lattice-matched to GaSb, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 320-328.
42. D. Donestsky, S. Anikeev, G. Belenky *et al.,* Reduction of interfacial recombination in GaInAsSb/GaSb double heterostructures, *Applied Physics Letters* **81, 25** (2002) 4769-4771
43. R.J. Kumar, R.J. Gutmann, J.M. Borrego *et al.*., Recombination parameters in InGaAsSb epitaxial layers for thermophotovoltaic applications, *Mat. Res. Soc. Symp. Proc.* **763** (2003) B2.4.1-B2.4.6.

44. R.J. Kumar, J.M. Borrego, P.S. Dutta *et al.*., Auger and radiative recombination coefficients in 0.55-eV InGaAsSb, *J. Applied Physics* **97** (2005).
45. R.U. Martinelli, D.Z. Garbuzov, H. Lee *et al.*., Minority-carrier transport in InGaAsSb thermophotovoltaic diodes, *Thermophotovoltaic Generation of Electricity, Third NREL Conf.* , J.P. Benner and T.J. Coutts, eds., American Inst. Physics Conf. Proc. **401** (1997) 389-395.
46. I. Reich, M.L. Gomez-Herrera, P. Diaz *et al.*., Measurement of Auger lifetime in GaInAsSb/GaSb heterostructures using the photoacoustic technique, *Applied Physics Letters* **79, 7** (2001) 964-966.
47. C.H. Henry, Limiting efficiencies of ideal single and multiple energy gap terrestrial solar cells, *J. Applied Physics* **51, 8** (1980) 4494-4500.
48. M.A. Green, "Limits on the open-circuit voltage and efficiency of silicon solar cells imposed by intrinsic Auger processes, *IEEE Trans. Electron Devices* **ED-31, 5** (1984) 671-678.
49. G.L. Araújo and A. Martí, Limiting efficiencies of GaAs solar cells, *IEEE Trans. Electron Devices* **37, 5** (1990) 1402-1405.
50. A. Luque, The requirements of high efficiency solar cells, in *Physical Limitations to Photovoltaic Energy Conversion*, A. Luque and G.L. Araújo eds., Adam Hilger, Bristol, 1990, 1-42.
51. P. Landsberg, Limitations to open circuit voltage by recombination, in *Physical Limitations to Photovoltaic Energy Conversion*, A. Luque and G.L. Araújo eds., Adam Hilger, Bristol, 1990., 134-147.
52. R. Mertens, R. Girisch, and M. Ghannam, Surface recombination in high efficiency silicon solar cells in *Physical Limitations to Photovoltaic Energy Conversion*, A. Luque and G.L. Araújo eds., Adam Hilger, Bristol, 1990, 148-173.
53. C.A. Wang, Antimony-based III-V thermophotovoltaic materials and devices, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 255-266.
54. P. Baldasaro, E.J. Brown, D.M. DePoy, B.C. Campbell, and J.R. Barrington, *The First NREL Conf. on Thermophotovoltaic Generation of Electricity*, T.J. Coutts and J.P. Benner, eds., American Inst. Physics Conf. Proc. **321** (1994) 29-43.
55. B. Bitnar, Silicon, germanium and silicon/germanium photocells for thermophotovoltaics applications, *Semiconductor Science and Technology* **18** (2003) S221-S227.
56. V.M. Andreev, V.P. Khvostikov, O.V. Khovostikova, E.V. Oliva, V.D. Rumyantsev, and M.Z. Shvartz, Low-bandgap Ge and InAsSbP/InAs-based TPV cells, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 383-391
57. V.P. Khostikova, V.D. Rumyantsev, O.A. Khvostikova, M.Z. Shvarts, P.Y. Gazaryan, S.V. Sorokina, N.A. Kaluzhniy, and V.M. Andreev, Thermophotovoltaic cells based on low-bandgap compounds, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 436-444.
58. O.V. Sulima, A.W. Bett, M.G. Mauk, B. Ya. Ber, and P.S. Dutta, Diffusion of Zn in TPV materials: GaSb, InGaSb, InGaAsSB and InAsSbP, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 402-413.
59. T. Schlegl, F. Dimroth, A. Ohm, and A.W. Bett, TPV modules based on GaSb substrates *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 285-293.
60. A.W. Bett and O.V. Sulima, GaSb photovoltaic cells for applications in TPV Generators, *Semiconductor Science and Technology* **18, 5** (2003) S184-S190.

61. G. Stollwerck, O.V. Sulima, and A.W. Bett, Characterization and simulation of GaSb device-related properties, *IEEE Trans. Electron Devices* **47, 2** (2000) 448-447.
62. D. Martín and C. Algora, “Temperature-dependent GaSb material parameters for reliable thermophotovoltaic cell modeling” *Semiconductor Science and Technology* **19** (2004) 1040-1052.
63. C.M. Ruiz, O. Vigil, C. Algora, D. Martín, V. Bermúdex, and E. Diéguez, “Transparent conducting oxides as antireflection coatings for GaSb TPV Cells” *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 221-229.
64. H. Yugami, K. Kobayashi, and H. Sai, “A broadband antireflection for GaSb by means of subwavelength grating (SWG) structures” *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 482-486.
65. C. Bumby, P.A. Shields, R.J. Nicholas, Q. Fan, G. Shmavonyan, L. May, and S.K. Haywood, “Improved efficiency of GaSb/GaAs TPV cells using an offset p-n Junction and off-axis substrates” *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 353-359.
66. L.M. Fraas, J.E. Avery, W.E. Daniels *et al.*., TPV tube generators for apartment building and industrial furnace applications, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 38-48.
67. P.S. Dutta, J.M. Borrego, H. Ehsani *et al.*. GaSb and GaInSb thermophotovoltaic cells using diffused junction technology in bulk substrates, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 392-401
68. A. Mitric, J. Vincent, R. Caillard, V. Bermudez, E. Dieguez, and T. Duffar, GaInSb bulk crystal growth for thermophotovoltaic application” *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 377-386.
69. C. Stelian, A. Mitric, V. Corregidor, L.C. alves, N.P. Barradas, and T. Duffar, Bridgman solidification of concentrated GaInSb alloys with variable growth rate, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 415-423.
70. M.G. Mauk, A.N. Tata, and J.A. Cox, Solution growth of thick III-V antimonide alloy epilayers (InAsSb, InGaSb, InGaAsSb, AlGaAsSb, and InAsSbP) for ‘virtual substrates’” *J. Crystal Growth* **225** (2001) 236-243.
71. M.G. Mauk, A.N. Tata, J.A. Cox, O.V. Sulima, and S. Datta, Ternary and quaternary alloy III-V substrates made by liquid-phase epitaxy” *IEE Proc.- Optoelectronics* **150, 4** (2003) 395-398.
72. M.P. Mikhailova, S.V. Slododchikov, N.D. Stoyanov, N.M. Stus, Yu.P. Yakovlev, Noncooled InAsSbP/InAs photodiodes for the spectral range 3-5 um, *Technical Physics Letters*, **22, 8** (1996) 672-673.
73. M.W. Wanlass, J.J. Carapella, A. Duda, K. Emery, L. Gedvilas, T. Moriarty, S. Ward, J. Webb, and X. Wu, High-performance, 0.6-eV GaInAs/InAsP thermophotovoltaic converters and monolithically interconnected modules, , *Thermophotovoltaic Generation of Electricity, Fourth NREL Conf.*, T.J. Coutts, C.S. Allman, and J.P. Benner, eds., American Inst. Physics Conf. Proc., **460**, 1999 (1998).
74. S.L. Murray, F.D. Newman, C.S. Murray, D.M. Wilt, M.W. Wanlass, P. Ahrenkiel, R. Messham, and R.R. Siergiej, “MOCVD growth of lattice-matched and mismatched InGaAs materials for thermophotovoltaic energy conversion” *Semiconductor Science and Technology* **18, 5** (2003) S202-S208.

75. M.W. Wanlass, J.S. Ward, K.A. Emery, M.M. Al-Jassim, K.M. Jones, and T.J. Coutts, GaInAs thermophotovoltaic converters, *Solar Energy Materials and Solar Cell*s **41/42** (1996) 405-417.
76. R.R. Siergiej, B. Wernsman, S.A. Derry *et al.*., 20% efficient InGaAs/InPAs thermophotovoltaic cells, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 414-423.
77. D. Wilt, R. Wehrer, M. Palmisiano, M. Wanlass, and C. Murray, Monolithic interconnected modules (MIMs) for thermophotovoltaic energy conversion, *Semiconductor Science and Technology* **18** (2003) S209-S215.
78. V. Andreev *et al.*., Low-Bandgap PV and Thermophotovoltaic Cells, *Proc. 3rd World Conf. on Photovoltaic Solar Energy* (2003) 15-18.
79. M.G. Mauk, O.V. Sulima, J.A. Cox, and R.L. Mueller, Low-bandgap (0.3 to 0.5 eV) InAsSbP thermophotoovoltaics, *3rd World Conf. on Photovoltaic Solar Energy* (2003)
80. V.A. Gevorkyan, V.M. Aroutiounian, K.M. Gambaryan, M.S. Kazaryan, K.J. Touryan, and M.W. Wanlass, Liquid-phase electro-epitaxial growth of low-bandgap p-InAsPSb / n-InAs and p-InAsP / n-InAs diodes heterostructures for thermo-photovoltaic application" *Thin-Solid Films* **451-452** (2004) 124-127.
81. A. Krier and Y. Mao, 2.5 um light-emitting diodes in InAsSbP/InAs for HF detection, *IEE Proc.- Optoelectronics* **144, 5** (1997) 355-359.
82. M.J. Kane *et al.*., Bulk and surface recombination in InAs/AlInSb 3.45 um light-emitting diodes, *Applied Physics Letters*, **76, 8** (2000) 943-945.
83. L.-C. Chen, M.-C. Wu, and W-J Chen, Thermal properties of InAssbP homostructure diodes, *Japanese J. Applied Physics* **37** (1998) 5622-5624.
84. B. Bitnar, W. Durisch, A. Meyer, and G. Palfinger, New flexible photocell module for thermophotovoltaic applications, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 465-473
85. F. Liu and K.L. Wang, Low radiation temperature thermal photovoltaic cells, *J. Applied Physics,* **97** (2005).
86. J.P. Connolly and C. Rohr, Quantum well cells for thermophotovoltaics, *Semiconductor Science and Technology* **18** (2003) S216-S220.
87. P. Griffin, I. Ballard, K. Barnham, J. Nelson, A. Zachariou, J. Eppler, G. Hill, C. Button, and M. Pate, The application of quantum well solar cells to thermophotovoltaics, *Solar Energy Materials and Solar Cells* **50** (1998) 213-219.
88. Al. Freundlich and A. Ignatiev, Quantum well thermophotovoltaic energy converter, U.S. Patent 6,150,604 (November 21, 2000)
89. R.K. Huang, C.A. Wang, M.K. Connors, G.W. Turner, and M. Dashiell, Hybrid back surface reflector GaInAsSb thermophotovoltaic devices, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 329-336.
90. C.A. Wang *et al.*., Fabrication and characterization of wafer-bonded GaInAsSb epitaxy for monolithically interconnected thermophotovoltaic devices, *Mat. Res. Soc. Symp.* **763** (2003).
91. C.A. Wang, P.G. Murphy, P.W. O'Brien *et al.*., Wafer-bonded internal back-surface reflectors for enhanced TPV performance, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 473-481.
92. C.A. Wang, R.K. Wang, M.K. Connors *et al.*., Monolithic series-interconnected GaInAsSb/AlGaAsSb thermophotovoltaic devices wafer bonded to GaAs, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 294-302.

93. V.M. Andreev, V.P. Khovostikov, V.R. Larionov *et al..*, Tandem GaSb / InGaAsSb thermophotovoltaic cells, *Conf. Rec. 26th IEEE Photovoltaics Specialists Conf.* (1997) 935-938.
94. R.R. Siergiej, S. Sinharoy, T. Valko *et al..,* InGaAsP/InGaAs tandem TPV device, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 480-488.
95. D.M. Wilt, R.J. Wehrer, W. Maurer, P.P. Jenkins, B. Wernsman, and R.W. Schultz, Buffer layer effects on tandem InGaAs TPV devices, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 453-461.
96. R. DiMatteo, Method and apparatus for the generation of charge carriers in semiconductor devices, U.S. Patent 6,084,173 (2000).
97. J.L. Pan, H.K.H. Choy, and C.G. Fonstad, Jr., "Very large radiative transfer over small distances from a black body for thermophotovoltaic applications" *IEEE Trans. Electron Devices* **47, 1** (2000) 241-249.
98. J.E. Raynolds, Enhanced electromagnetic energy transfer between a hot and cold body at close spacing due to evanescent fields, Thermophotovoltaic Generation of Electricity, Fourth NREL Conf., T.J. Coutts, J.P. Benner, C.S. Allman, American Inst. Physics Conf. 460 (1999) 49-57.
99. J.E. Raynolds, Enhanced electromagnetic energy transfer between a hot and cold body at close spacing due to evanescent fields, Report from US Gov't KAPL-P-000112 (K98151).
100. R. DiMatteo, P. Grieff, D. Seltzer, *et al..*, Micron-gap thermoPhotoVoltaics (MTPV), *Thermophotovoltaic Generation of Electricity, Sixth Conference*, A. Gopinath, T.J. Coutts, and J. Luther, eds., AIP Conf. Proc. 738 (American Inst. Physics, 2004) 42-51.
101. M. Whale, The influence of interference and heterojuntions on the performance of microscale thermophotovoltaic devices, *Microscale Thermophysical Engineering* **5** (200) 89-106.
102. M.D. Whale and *E.G.* Cravallo, Modeling and performance of microscale thermophotovoltaic energy conversion devices, *IEEE Trans. Electron Devices* **17, 1** (2002) 130-142.
103. A. Meulenberg, personal communication, 2003.
104. M.H. Hannon, M.W. Dashiell, L.C. DiNetta, A.M. Barnett, Lightweight, light trapped, thin GaAs solar cell for spacecraft applications: progress and results update, *Conf. Rec. IEEE Photovoltaics Specialists Conf.* (1996) 191-194.
105. O.V. Sulima, Ph.D. Dissertation, Leningrad Technical University (1985) (in Russian).
106. P.M. Fourspring, D.M. DePoy, J.F. Beausang *et al..*, Thermophotovoltaic spectral control, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 171-179.
107. A.Gombert, An Overview of TPV emitter technologies, *Thermophotovoltaic Generation of Electricity, Fifth Conference*, T.J. Coutts, G. Guazzoni, J. Luther, eds., American Inst. Physics Conf. Proc. **653** (2003) 123-131.
108. W.E. Horne, M.D. Morgan, W.P. Horne, and V.S. Sundaram, Frequency selective filters applied to thermophotovoltaic generators, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 189-197.
109. A.Narayananaswamy, J. Cybulski, and G. Chen, 1D Metallo-Dielectric Photonic Crystals as selective emitters for thermophotovoltaic applications, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) (2004) 215-220.

110. H. Sai, H. Yugami, Y. Kanamori, and K. Hane, Spectrally selective thermal radiators and absorbers with periodic microstructured surface for high-temperature applications, *Microscale Thermophysical Engineering* **7** (2003) 101-115.
111. H. Sai, T. Kamikawa, Y. Kanamori, K. Hane, H. Yugami, M. Yamaguchi, Thermophotovoltaic generation with selective tungsten emitters, *Thermophotovoltaic Generation of Electricity, Sixth Conf.*, A. Gopinath, T.J. Coutts, J. Luther, eds., American Inst. Physics Conf. **738** (2004) 206-214.
112. J.M. Gee, J.B.Moreno, S-Y. Lin, and J.G. Fleming, Selective emitters using photonic crystals for thermophotovoltaic energy conversion, *IEEE Photovoltaics Specialists Conf.* (2002) 896-899.
113. J.M. Gee, S-Y. Lin, J.G. Fleming, J.B. Moreno, Thermophotovoltaic energy conversion using photonic bandgap selective emitters, U.S. Patent 6,583,350 (2003).
114. N.-P. Harder and M.A. Green, Thermophotonics, *Semiconductor Science and Technology* **18, 5** (2003) S270-S278.
115. I.Tobías and A. Luque, Ideal efficiency and potential of solar thermophotonic, converters under optically and thermally concentrated power flux, *IEEE Trans. Electron Devices* **49, 11** (2002) 2024-2030.
116. A. Joullié, F. Jia Hua, F. Karouta, H. Mani, and C. Alibert, III-V alloys based on GaSb for optical communications at 2.0-4.5 μm, *Optical Fiber Sources and Detector,* SPIE Proceedings of SPIE **587**, (1985) 46-57.
117. S. Adachi, *Physical Properties of III-V Semiconductor Compounds: InP, InAs, GaAs, GaP, InGaAs, and InGaAsP* (Wiley, New York, 1992).
118. K. Shim, Principal of energy band gaps of the quaternary alloy AlGaSbAs, *Solid State Communications*, **134, 6** (2005) 437-441.
119. J. Bhan, A. Joullié, H. Mani, A.-M. Joullié, and C. Alibert, III-V heterostructures for laser emission in the 2.55 μm wavelength range *Materials and Technologies for Optical Communications* SPIE Proc. **866** (1987) 126-134.
120. M.P.C.M. Krijn, Heterojunction band offsets and effective masses in III-V quaternary alloys, *Semiconductor Science and Technology* **6** (1991) 27-31.
121. M. Zierak, J.M. Borrego, I. Bhat, R.J. Gutmann, and G. Charache, Modeling of InGaSb thermophotovoltaic cells and materials, *Thermophotovoltaic Generation of Electricity, Third NREL Conf.,* J.P. Benner and T.J. Coutts, eds. American Inst. Physics Conf. **401** (2003) 55-65.

Index

Acceptance 605
Active layer mesa diameter 402–403
Active regions
  band-edge optimization 21–46
  bound to continuum 330–332
  cascaded 17–21
  cavity-integrated design 53–59
  design 3–4
  double phonon resonance 330
  final state optimization 47–53
  optimal parameters 7
  optimization 53–56
  quantum cascade laser 326–336
  single-stage 8–14
  thickness 14–16
  three-well 328–330
  two-well 328–330
Ag
  thermal conductivity 248
  thermal expansion coefficient 248
Air-Borne Laser system 661
AlGaAsSb 132
(AlGaIn)(AsSb)-based diode lasers 136–140
  crystal growth 136–137
  device processing 139–140
  epitaxial layer characterization 137–139
All Back Contact solar cells 718–719
All-hollow-waveguide laser delivery system 624–625
AlSb 40–46
AM1.5 efficiency 674
Antimonide type-II "W" lasers 189–217
  interband cascade 201–203
  MBE growth 191–195
  optically-pumped 196–199
  performance improvement 208–212
  single-stage 199–201
  VCSEL and PCDFB 203–207
Applied power density 19
Arc lamps 664–666
Asymmetric band offset confinement 219–236
Asymmetric quantum wells 445–446
  normal incidence absorption 445
  voltage tuning via Stark effect 445–446
Atmospheric windows for counter-measure systems 639–645
  aero-optical effects 644–645
  extinction 643
  haze, fog, cloud and rain 641
  humidity 40–41
  ionisation8 643–644
  thermal blooming 643
  turbulence 641–643
  wakes and plumes 644
Auger coefficients 195, 209
Auger rates of zinc-blende crystals 109–110
Auger recombination 4–5, 11, 13, 27, 29, 73–75, 81–82, 106–109, 319, 362, 375
Auger suppression 50–52, 226–227
  dilute III-V nitrides 111–112
  Pb salts 110
  quantum wells 112–114
Auger transition matrix element 227

BaF_2
  Bragg mirrors 270
  substrates 242–245
  thermal conductivity 248
  thermal expansion coefficient 248

Band structure 94–100
  calculation by **k.p** method 93–105
  lead salts 99–100
  temperature dependence 103
  zinc-blende crystals 96–99
Band-edge optimization 21–46
  InAs/GaSb/AlSb materials 40–46
Bernard-Duraffourg condition 109
Bioimaging 627–628
Biomedical delivery fibers 622–624
Biomedical lasers 620–621
  all-hollow-waveguide delivery system 622–624
  carbon dioxide 620
  free-electron 620
  solid-state 620
Biophotonics, MIR 615–634
  all-hollow-waveguide laser delivery system 624–625
  basic delivery system 619
  bioimaging 627–628
  biomedical delivery fibers 622–624
  biomedical lasers 620–621
  fundamentals of 617–619
  incoherent light sources 621–622
  laser surgery and tissue ablation 625–627
  spectroscopy and biosensors 628–631
Biosensors 628–631
  breath analysis 629–630
  non-invasive blood glucose monitoring 628–629
BLIP IR detectors, uncooled IR radiation shields 478–479
Bloch's theorem 95, 97
Blood glucose sensors 628–629
Bolometers 602
Boltzmann constant 107, 118
Bound to continuum active regions 330–332
Bound-to-bound transition 431, 432
Bragg mirrors 168, 171
  materials 268–270
  Pb salt-based 268–272
Breath analysis 629–630
Brillouin zone 24, 61, 68, 82, 94, 120, 249, 338
Bulk-like PbTe VCSELs 289–291

Camouflage, negative luminescent devices 482
Carbon dioxide
  exhaled 239–240
  laser 620
Carrier mobilities 82–84
Carrier recombination rates 80–82
Cascaded active regions 17–21
Cavity ring-down spectroscopy 600, 609
CdHgTe-based VCSELs 169–170
CEA, CdHgTe-based VCSELs 169–170
CELEOS-DIRCM system 662
Chalcogenide glass fiber 413–419, 422, 424
Chemical mechanical polishing 536
Chemical sensing 419–424
Cladding 56–59
Closed-loop IR countermeasures 669
Coherent light sources 666
Concave/plane laser cavity 165–167
Conduction band density of states 35
Continuous wave operation 241, 324
Conventional microcavity VCSELs 160–163
Conversion efficiency 411
Coulomb matrix elements 81
Countermeasure systems 190, 635–671
  approaches to 636
  defeat mechanisms 645–661
  heat-seeking missiles 637
  historical development 637–639
  IR jammers 661–669
  laser jammers 669–670
  mirrors 637
  propagation and atmospheric windows 639–645
  smoke screens 637

Cu
thermal conductivity 248
thermal expansion coefficient 248
Current crowding 403
Current to gain factor 7
Current threshold current density 35
CW-VCSELs 293–298
laser emission 294–298
structure and optical properties 293–294

2D-electron channel 220
Damage countermeasures 636, 654–659
in-band damage route 655–657
out-of-band damage 657–659
Dark current 501–502, 507
avalanche photodiodes 570–572
measurements 542–543
Dazzle countermeasures 636, 653–654
Deception countermeasures 636, 647–652
expendables (flares) 649–651
heat-seeking 648
jammers 651–652
Defeat mechanisms for countermeasure systems 636, 645–661
damage 636, 654–659
dazzle 636, 653–654
deception 636, 647–652
denial 636, 645–647
destruction 636, 659–661
Denial countermeasures 636, 645–647
large-area smoke screening system 647
pyrotechnic smoke 647
rapid-bloom obscurant 647
smokes and obscurants 64
Density of states 493
calculations 77
Density of states imbalance 21–25
effects of 25–30
reduction by quantum confinement 30–32
reduction by strain 30, 31
Destruction countermeasures 636, 659–661
Diamond
thermal conductivity 248
thermal expansion coefficient 248
Dielectric waveguide design 309–315
connection between confinement factor and beam quality 309–313
ghost modes in MIR-OPSL 313–315
Differential absorption LIDAR (DIAL) 639
Differential index of refraction 79–80
Dilute III-V nitrides, Augur suppression 111–112
DIRCM systems 638–639, 662, 667–669
CELEOS-DIRCM system 662
laser-based 652
Directed energy weapons 638
Directed infrared countermeasure systems *see* DIRCM systems
Distributed Bragg reflectors 197
Distributed feedback 139
Dot-in-well heterostructures 500
Double phonon resonance active region 330
Double-barrier quantum wells 433–436
on GaAs substrates 434
in InP substrates 435–436
Drude contribution 209, 210
Dynamic infrared scene projection 476–477

Edinburgh Instruments Gascard gas detector 606–608
Effective mode width 19
Eigenvalues 102
Eigenvectors 68
EL *see* electroluminescence
Electric loss 140–143
Electrical pumping 162, 176–177

Electroluminescence 220–224
single p-GaIn0.17AsSb/n-In0.83GaAsSb heterostructure 228–229
Electron beam lithography 535
Electron impact ionization 557–561
InAs and InSb type semiconductors 560–561
multi-valley GaAs or InP semiconductors 557–558
multi-valley GaSb and alloy semiconductors 558–560
Electrophonon resonance 249–253
Empirical pseudopotential 307
Empirical tight-binding method 515
Envelope function approximation 101
Epitaxial layer structure 137–139
Étendue 605
Ethane, real-time recording 630
EuSe, Bragg mirrors 270
EuTe, Bragg mirrors 270
Excess noise factor 579
Exhaled biomarker molecule breath testing 238–240
External-cavity VCSELs 163–167
geometrical stability 165–167
structure description 163–165
Extinction 642

Fabry-Perot cavity 266, 315
Fabry-Perot emission 205
Fermi energies 5, 24
Fermi levels 22–23
Fermi-Dirac distribution functions 107
Fermi's Golden Rule 77, 106, 141, 327
Fiberoptic evanescent wave spectroscopy 422
Figures of merit 7, 8, 10–11, 29
bulk GaAs 12–14
Final state optimization 47–53
Flares 649–651
Flat Valence Band Approximation 109
Fourier transformation 66–67
Free-electron biomedical laser 620
Full width at half maximum 170, 193, 205, 245, 365, 383, 526

GaAlAsSb, ionization coefficient 565–567
GaAlSb, ionization coefficient 565–567
GaAs
double-barrier QWs 434–435
figures of merit 12–14
multi-valley semiconductors 557–558
quantum cascade lasers 333–336
slope efficiency 17
thickness 14–16
Gain coefficient 20
GaInAsSb *see* InGaAsSb
Gas absorption spectra 595–597
Gas chromatography mass spectroscopy 240
Gas detection 595–613
gas absorption spectra 595–597
IR sources 598–601
laser techniques 609–610
methods of 597–598
semiconductor sources 600–601
system design 604–609
thermal sources 598–600
Gas sensors
IR 601–604
IR sources *see* IR sources for gas detection
long-path with semiconductor source and detector 608–609
non-imaging with thermal source and detector 606–608
photovoltaic 602
see also gas detection
GaSb 40–46
electronic band structure 23
ionization coefficient 565–567
nanopillars 537–538
TPVs 699–701

GaSb-based VCSELs 173–184
electrically-pumped 176–177
Naval Research Laboratory/Hughes Research Laboratory 173–175
optically-pumped 177–184
University of Montpellier 176–184
Gd gettering 364–365
Ge TPVs 697–699
Geometric extent 605
Gettering
Gd 364–365
rare earth 363–364
Ghost modes 313–315
totally internally reflected 313

Hakki-Paoli gain measurements 11, 318
Hardware-in-the-loop simulation 476
Heat-seeking technologies 648
Heavy-hole mass 27–28
Heterostructures 5–6
k.p formalism 63–68
HgCdTe
magnetoconcentration effect 459–461
photodiodes 467–471, 490–491
High pressure studies 114–125
MIR LEDs 374–377
near-infrared applications 114–120
High-resolution X-ray diffraction 137, 138, 245, 247
High-speed avalanche photodiodes 547–592
2–2.5 μm spectral range 564–581
3–5 μm spectral range 581–586
impact ionization 548–557
ionization coefficients 557–564
noise and response speed 554–557, 585–586
High-temperature luminescence 233–234
Hitran database 239
Hole impact ionization 561–564
in semiconductors with band gap “resonance” 562–564
Hot wall epitaxy 242
Hughes Research Laboratory, GaSb-based VCSELs 173–175
Humidity 640–641
Hydrostatic pressure 103–105

IA configuration 197
III-Sb-based laser system 132–136
AlGaAsSb 132
device self-heating 148
high-power performance 146–150
laser structure 134–136
long-wavelength 150–151
optical gain and electric loss 140–143
strained GaInAsSb layers 132–134
Image dancing 641
Impact ionization 548–557
threshold energy 549–550
InAs 40–46
band structure 64
electroluminescence emission spectrum 366, 368
ionization coefficient 582–584
LEDs 121–122
MIR LEDs 368–370
photoluminescence spectrum 365
purification using Gd gettering 364–365
semiconductors 560–561
TPVs 706–710
InAs-AlSb 344–347
InAs/GaAs, quantum dot infrared photodetectors 490–491
InAs/GaSb
nanopillars 538
superlattice 471–472, 516–521
InAsP/InP TPVs 703–706
InAsSb
grown by liquid phase epitaxy 381–386
ionization coefficient 582–584

MIR LEDs 36–38
quantum dot LEDs 381–386
InAsSbP TPVs 706–710
Incoherent IR sources 662–664
Incoherent light sources 621–622
Infrared *see* IR
InGaAs
ionization coefficient 582–584
TPVs 703–706
InGaAs-AlAsSb-InP 340–344
(In)GaAs/AlGaAs on GaAs 432–433
InGaAs/AlInAs-InP 433
strain compensated 338–340
InGaAsSb
band structures 36
density of states 37
diodes 32–40
figure of merit curve 33, 34
ionization coefficient 567–570
lasers 122–125
TPVs 72308
InGaAsSb/GaAlAsSb avalanche
photodiodes 572–576
noise and response speed 576–581
InP
double-barrier QWs 435–436
multi-valley semiconductors 558–560
quantum cascade lasers 333–336
VCSELs 171–173
InSb
electronic band structure 23
magnetoconcentration 459–461
photodiodes 464–466
semiconductors 560–561
Interband
cascade W lasers 201–203
detectors 429–431
mid-IR devices 121–125
optical transition rates 106
transitions 4
Interface
bonds 68–70
electroluminescence 220–224
Interface lasers 219–236
improved temperature dependence 228–234
type II broken-gap p-GaInAsSb/ p-InAs heterointerface 220–227
Internal quantum efficiency 362–363
Intersubband absorption 56
suppression of 48–50
Intersubband population inversion 327–328
Intersubband quantum cascade lasers 20
Intersubband transitions 4, 441–442
Intra-cavity laser absorption spectroscopy 165, 600, 609, 610
Intraband absorption 495–496
Intrinsic material properties 4–5
Ionization 643–644
Ionization coefficients 557–564
dependence on electric field 551–552
electron impact ionization 557–561
experimental investigation 582–584
GaInAsSb 567–570
GaSb, GaAlSb and GaAlAsSb 565–567
hole impact ionization 561–564
and multiplication coefficients 553–554
separation using quantum structures 586–588
IR countermeasures *see* IRCM
IR focal plane array 516
IR gas detectors 601–604
optical immersion 602–604
IR jammers 661–669
arc lamps 664–666
coherent sources 666
DIRCM systems 667–669
incoherent IR sources 662–664
IR sources for gas detection 473–476, 598–601
long wavelength 474–476
semiconductor 600–601
temperature stabilisation 474
thermal 599–600

IRCM 651, 661, 665, 667
closed-loop 669
IV-VI semiconductors 237–264
BaF_2 242–245
growth and characterization 241–248
self-heating 248–249
silicon 245–248
spectroscopy with 237–241

Jammers 651–652
IR 661–669
laser 669–670
Johnson-noise limitation 490

k.p theory 61–63, 93–105, 307, 344, 493
heterostructures 63–68
Kramers-Krönig transformations 79

L-point band gap 249
Lagrange invariant 605
Large-area smoke screening system 647
Lasers
(AlGaIn)(AsSb)-based diode 136–140
InGaAsSb 122–125
intersubband quantum cascade 20
MIR-OPSL 303–322
MWIR 3–6
PCDFB 190
quantum cascade 93, 323–355
type-I QW 131–157
VCSEL 140, 158–188
VECSEL 164–165
W-laser 41, 70
Laser gas detectors 609–610
Laser jammers 669–670
Laser surgery and tissue ablation 625–627
Layer doping 400–401
Layer thickness 401–402
Lead *see* Pb
LEDs
InAs 121–122
quantum dot 381–386
resonant cavity 377, 378
LEDs, MIR 359–394
device comparison 379–381
high-pressure measurements 374–377
InAs for methane detection 368–370
InAsSb 366–368
InAsSb grown by liquid phase epitaxy 381–386
internal quantum efficiency 362–363
neutral solvent epitaxy 370–373
optical extraction 377–379
purification by Gd gettering 364–365
purification by rare earth gettering and Pb neutral solvent epitaxy 363–364
superluminescence and ring lasers 386–389
LED-photodiode opto-pairs 395–428
active layer properties 400–403
in chemical sensing 419–424
choice of operating mode 408–415
contacts 403–408
device configuration and fabrication 396–408
out-coupling of radiation 415–419
substrate and buffer layers 397–400
LIDAR 131
differential absorption 639
Light-emitting diodes *see* LEDs
Linewidth enhancement factor 7, 79–80
Liquid phase epitaxy 242
p-i-n junctions 566
Long-wavelength III-Sb-based lasers 150–151
Longitudinal optic phonon energy 323, 327
Low-energy Hamiltonian 60
Luttinger parameters 97, 99

Magnetoconcentration 458–462
  InSb and CdHgTe devices 459–461
  principle of operation 458–459
Material pressure coefficients 120–121
Maxion W diode lasers 199–201
MBE *see* molecular beam epitaxy
Metal-organic chemical vapor deposition 488
Metal-organic vapor phase epitaxy 136, 201, 323, 347–351
Micron-gap TPVs 717–720
Microstructured emitters 720
Mid-infrared *see* MIR
Mid-wavelength infrared laser
  active region design 3–4
  heterostructure design 5–6
  QWIP detectors 429–452
Mie scattering 639, 640, 641
MIR advanced chemical laser (MIRACL) 638
MIR material pressure coefficients 120–121
MIR optically pumped semiconductor lasers *see* MIR-OPSLs
MIR optoelectronic devices 237–264
MIR quantum cascade lasers *see* quantum cascade lasers
MIR-OPSLs 303–322
  beam quality 319–321
  dielectric waveguide design 309–315
  ghost modes 313–315
  guidelines for optical pumping 303–306
  spectral, power and loss measurements 315–319
  type II quantum wells 306–309
MIR-VCSELs 265–301
Missile approach warning system (MAWS) 667
Modal waveguide loss coefficient 9
Molecular beam epitaxy 136, 191, 242, 342, 436, 488
  growth of W laser structures 191–195
Monolithic interconnected modules (MIMs) 705
Monte Carlo integration 82, 109
Multi-wavelength operation 325–326
Multiple quantum well structures 243
Multiplication coefficient 553–554
MWIR laser *see* mid-wavelength infrared laser

Nanopillar fabrication 535–543
  dark current measurements 542–543
  electron beam lithography 535
  fabrication process 535–536
  in GaSb material 537–538
  in InAs/GaSb superlattice material 538
  polyimide etchback 540–541
  polyimide planarization 540
  reactive ion etching 538–540
  top contact deposition 541–542
Nanopillar structures 521–526
Naval Research Laboratory
  GaSb-based VCSELs 173–175
  optical modulation technique 455
Near-infrared applications 114–120
  non-radiative recombination 118–120
  radiative recombination 114–117
Negative differential resistance 469
Negative luminescence 453–485
  camouflage 482
  dynamic infrared scene projection 476–477
  IR sources for gas sensing 473–476
  magnetoconcentration effect 458–462
  optical concentrators 456–458
  photodiodes 462–473
  radiative cooling 477–478
  radiometric reference planes for thermal imagers 479–482
  uncooled IR radiation shields for BLIP IR detectors 478–479
Negative luminescent efficiency 455–456, 466

Negative luminescent power 409, 454
Negative luminescent power efficiency 456
Negative luminescent quantum efficiency 456
Nemesis countermeasure system 639, 668
Nernst Glower 599
Neutral solvent epitaxy 363–364, 370–373
Nitric oxide, exhaled 238–240
NL *see* negative luminescence
Noise current 507–508
Noise-equivalent power 508
Nomarski microscope photography 386
Non-equilibrium absorption 70–73
Non-radiative recombination 118–120

Ohmic power density 8
OPIC *see* optical pumping injection cavity
Optical concentrators 456–458
Optical extraction 377–379
Optical fibers, biomedical 622–624
Optical gain 140–143
Optical immersion gas detectors 602–604
Optical interference filters 604
Optical losses 143–146
Optical parametric oscillator 600
Optical properties 77–80
Optical pumping 601
Optical sensing 395–428
Optical turbulence 641–643
Optically pumped lasers 162, 164, 177–184
  antimonide type-II "W" lasers 196–199
  injection cavity 196–197
  MIR *see* MIR-OPSLs
  single frequency operation and frequency tunability 180–184
  VCSEL and VECSEL fabrication 177–180
Out-coupling of radiation 415–419
Output power 9

Pb neutral solvent epitaxy 363–364
Pb salts 99–100
  Augur suppression 110
  VCSELs 272–281, 288–292
  VSCELs 272–281
PbS, Bragg mirrors 270
PbSe
  Bragg mirrors 270
  quantum wells 249–253
  self-assembled IR quantum dot lasers 281–284
  thermal conductivity 248
  thermal expansion coefficient 248
PbSe/PbSrSe VCSELs 280–281
PbTe
  Bragg mirrors 270
  thermal conductivity 248
  thermal expansion coefficient 248
PbTe QW VCSELs, room temperature 277–279
PbTe/EuTe quantum well VCSELs 274–277
  structure and optical characterization 274–275
  temperature dependence of emission 276–277
  threshold and laser emission 275–276
PbTe/PbEuTe quantum well VCSELs 291
PCDFB lasers *see* photonic-crystal distributed-feedback lasers
Peak specific detectivity 509
Phonon bottleneck 497–498, 522
Photoconductive gain 509–510
Photoconductive gas detectors 602
Photocurrent 507
Photodetectors *see* quantum dot photodetectors, MIR
Photodiodes 395
  HgCdTe 467–471, 490–491
  high-speed avalanche 547–592

InSb 464–466
negative luminescence 462–473
principle of operation 462–464
Photoluminescence 243–245, 247, 251–256, 413
spectroscopy 137
Photon over-excitation factor 686
Photon recuperation 686
Photonic crystals 720
distributed-feedback lasers 190, 203–207
Photovoltaic gas detectors 602
Poisson's ratio 104
Polyimide etchback 540–541
Polyimide planarization 540
Power-power curves 316
Pyroelectric detectors 601–602
Pyrotechnic smoke 647

QinetiQ 457
Quantum cascade lasers 93, 201, 248, 323–355
3–5 μm operation 336–347
active region design 326–336
GaAs-based 333–336
general concepts and current status 323–326
InP-based 333–336
metal organic vapor phase epitaxy 347–351
Quantum confinement 30–32
Quantum dots 487, 491–498
bound-state energy levels 492–493
density of states and carrier distribution 494
intraband absorption 495–496
phonon bottleneck and effective carrier lifetime 497–498
Quantum dot LEDs 381–386
Quantum dot photodetectors, MIR 487–513
dark current 501–502, 507
heterostructure designs 500–501
InAs/GaAs vs HgCdTe photodiodes 490–491
infrared detection with quantum dots 491–498
mid-infrared spectral response 502–503
noise current 507–508
normalized spectral response and peak responsivity 508
peak specific detectivity 508–509
performance at high operating temperatures 503–506
photoconductive gain and quantum efficiency 509–510
photocurrent 507
Stranski-Krastanow growth mode 498–500
Quantum dot VCSELs
growth and characterization 284–286
laser emission 286–288
lasing properties 292
Quantum efficiency 509–510
Quantum Hall effect 220
Quantum photovoltaic devices 515–545
cut-off wavelength 8 μm 531–533
material growth and characterization 526–531
nanopillar fabrication 535–543
nanopillar structures 521–526
superlattice Hamiltonian matrix 543–544
type II focal plane arrays 533–535
type-II InAs/GaSb superlattice 516–521
Quantum wells 7, 189
asymmetric 445–446
Augur suppression 112–114
double-barrier 433–436
optical gain and electric loss 140–143
strained layer 103–105
structures 101–102
type II 306–9
Quantum well infrared photodetectors *see* QWIP detectors
Quantum well TPVs 710–712
QW *see* quantum well

QWIP detectors 93, 429–452, 515
comparison with interband detectors 429–431
enhanced performance and functionality 444–448
MWIR transitions 431–436
photovoltaic operation 446–447
strained-balanced 436–444

Radiative cooling 477–478
Radiative recombination 80–81, 114–117
Rapid-bloom obscurant 647
Rare earth gettering 363–364
Rayleigh scattering 639, 641
Reactive ion etching 538–540
BCl_3:Ar and CH_4:H_2:Ar 538–539
cyclic CH_4:H_2:Ar/O_2 539–540
Recombination
Auger 4–5, 11, 13, 27, 29, 50–52, 73–75, 81–82, 106–109
non-radiative 118–120
radiative 80–81, 114–117
shockley-Read-Hall 12, 33, 55
surface 33
Reflection high-energy electron diffraction 245, 499
Refractive index contrast 268
Resonant cavity LEDs 377, 378
Ring lasers 386–389

Sarnof W diode lasers 199–201
Schrodinger equation 101
Secondary ion mass spectroscopy 137
SELD *see* surface emitting diode laser
Self-assembled IR quantum dot lasers 281–288
PbSe 281–284
Self-heating
III-Sb-based lasers 148
IV-VI semiconductors 248–249
Semiconductors
electronic and optical properties 59–75
impact ionization 548–557
interband ionization 548–549
Semiconductors IV−VI 237–64
Semiconductor disk laser *see* vertical external cavity surface emitting laser
Semiconductor heterostructure optical conversion 303
Semiconductor IR gas detection sources 600–601
Sensors
chemical 419–424
gas *see* gas detection; gas sensors
Separate absorption and multiplication region 572–576
SHOC *see* semiconductor heterostructure optical conversion
Shockley-Read defects 192
Shockley-Read-Hall recombination 12, 33, 55, 319, 362, 367, 375, 694
Si
substrates 245–248
thermal conductivity 248
thermal expansion coefficient 248
TPVs 697–699
Si-Ge alloy TPVs 697–699
Side-mode suppression ratio 139
Signal-to-noise ratio 507
Single p-GaInAsSb/p-InAs heterojunction 220–227
interface EL 220–224
suppression of Auger recombination 226–227
tunnelling-injection laser 224–226
Single-QW lasers 142–143
Single-stage active regions 8–14
Single-stage antimonide type-II “W” lasers 199–201
Slope efficiency 16–18
Smokes and obscurants 64
Solid-state biomedical laser 620
Sommerfeld enhancement 72
Spectral responsivity 508
Spectroscopy 628–631
Square QW structures 431–433
(In)GaAs/AlGaAs on GaAs 432–433
InGaAs/AlInAs on InP 433
Star Wars 636, 638

Staring arrays, non-uniformity correction 481–482
Stark effect 445–446
Strain 30, 31
Strain compensated InGaAs-AlInAs-InP 338–340
Strained GaInAsSb layers 132–134
Strained layer epitaxy 498
Strained layer quantum wells 103
Strained layer supperlattice 361
Strained-balanced QWIP detectors 436–444
  materials growth and characterisation 436–437
  measured transitions 442–443
  modeling intersubband transitions 441–442
  predicted temperature performance 443–444
  stoichiometric growth conditions 437–438
  structural and electrical properties 438–441
Stranski-Krastanow growth 489, 498–500
Substrate removal 254–259
Superlattices 41, 64
  empirical pseudopotential model 307–309
  Hamiltonian matrix 518, 543–544
Superluminescence 386–389
Surface emitting diode laser 164
Surface recombination 33

Tandem TPVs 714–717
Theater High Energy Laser (THEL) system 660
Thermal blooming 642
Thermal conductivity 248
Thermal expansion coefficients 248
Thermal imaging, radiometric reference planes 479–482
  DC restoration in scanned thermal imagers 479–481
  non-uniformity correction in staring arrays 481–482
Thermal IR gas detection sources 599–601
Thermophotonics 720–721
Thermophotovoltaic devices 673–738
  basic operation 680–684
  bulk II-V alloy ternary crystal 701–703
  device modeling 689–696, 728–731
  device structure and delineation 690–691
  GaSb 699–701
  InGaAs and InAsP/InP 703–706
  InGaAsSb 723–728
  literature and information sources 679–680
  low-bandgap InAs and InAsSbP 706–710
  micron-gap 717–720
  minority carrier recombination 692–696
  overview 673–679
  quantum well 710–712
  series interconnection, integrated reflectors and wafer-bonding 713–714
  silicon, germanium and Si-Ge alloy 697–699
  spectral control, microstructured emitters, filters and photonic crystals 720
  tandem 714–717
  thermodynamic limits and modeling 684–689
  thermophotonics 720–721
Thermopiles 601–602
Three-well active regions 328–330
Threshold current density 6, 39
Threshold energy 108
Throughput 605
Time-domain Fourier-transform 204
Top contact deposition 541–542
TPVs *see* thermophotovoltaic devices
Transition rates 105–114
  interband optical 106

Transmission electron microscopy
342
Tuning 292
Tunnelling-injection laser 224–226,
230–233
Two-valley model 552–553
Two-well active regions 328–330
Type I QW lasers 131–157
(AlGaIn)(AsSb)-based 136–140
III-Sb-based material system
132–136
structure 134–136
Type II broken-gap heterojunctions
219–236
Type II focal plane arrays 533–535
Type II quantum wells 306–309

Umklapp processes 82
University of Montpellier, GaSb-based VCSELs 176–184
electrically-pumped 176–177
optically-pumped 177–184

Valence force field model 492
VCSELs *see* vertical cavity surface
emitting lasers
VECSELs *see* vertical external cavity
surface emitting laser
Vertical cavity surface emitting lasers
140, 158–188
2–3 μm range 168–184
6–8 μm range 293–298
bulk-like PbTe VCSELs 289–291
CdHgTe-based structures
169–170
concave/plane laser cavity
165–167
conventional microcavity
160–163
electrical pumping 162
external-cavity 163–167
GaSb-based structures 173–184
history 266–267
InP-based structures 171–173
lead salt 272–281, 288–292
mid-infrared 265–301
optical properties 288–289
optical pumping 162, 164
PbTe/PbEuTe quantum well 291
quantum dot 284–286
W type 203–207
Vertical external cavity surface emitting laser 164
concave/plane laser cavity
165–167
fabrication 177–180
optically-pumped 164–165
Voltage factor 686

W-lasers 41
band structures 70
Walter Schottky Institute, InP-based
VCSELs 171–173
antimony type-II 189–217

X-ray photoelectron spectroscopy
247

Young's modulus 104

Zero-dimensional density of states
493
Zinc-blende crystals 96–99
Auger rates in 109–110