

# **Terahertz quantum cascade lasers**

by

Benjamin S. Williams

Submitted to the Department of Electrical Engineering and Computer  
Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2003

© Massachusetts Institute of Technology 2003. All rights reserved.

Author .....  
.....

Department of Electrical Engineering and Computer Science  
August 22, 2003

Certified by .....  
.....

Qing Hu  
Professor  
Thesis Supervisor

Accepted by .....  
.....

Arthur C. Smith  
Chairman, Department Committee on Graduate Students



# Terahertz quantum cascade lasers

by

Benjamin S. Williams

Submitted to the Department of Electrical Engineering and Computer Science  
on August 22, 2003, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

## Abstract

The development of the terahertz frequency range (1–10 THz,  $\lambda \approx 30\text{--}300\ \mu\text{m}$ ) has long been impeded by the relative dearth of compact, coherent radiation sources of reasonable power. This thesis details the development of quantum cascade lasers (QCLs) that operate in the terahertz with photon energies below the semiconductor *Reststrahlen* band. Photons are emitted via electronic intersubband transitions that take place entirely within the conduction band, where the wavelength is chosen by engineering the well and barrier widths in multiple-quantum-well heterostructures. Fabrication of such long wavelength lasers has traditionally been challenging, since it is difficult to obtain a population inversion between such closely spaced energy levels ( $\hbar\omega \approx 4\text{--}40\ \text{meV}$ ), and because traditional dielectric waveguides become extremely lossy due to free carrier absorption.

This thesis reports the development of terahertz QCLs in which the lower radiative state is depopulated via resonant longitudinal-optical phonon scattering. This mechanism is efficient and temperature insensitive, and provides protection from thermal backfilling due to the large energy separation ( $\sim 36\ \text{meV}$ ) between the lower radiative state and the injector. Both properties are important in allowing higher temperature operation at longer wavelengths. Lasers using a surface plasmon based waveguide grown on a semi-insulating (SI) GaAs substrate were demonstrated at 3.4 THz ( $\lambda \approx 88\ \mu\text{m}$ ) in pulsed mode up to 87 K, with peak collected powers of 14 mW at 5 K, and 4 mW at 77 K.

Additionally, the first terahertz QCLs have been demonstrated that use metal-metal waveguides, where the mode is confined between metal layers placed immediately above and below the active region. These devices have confinement factors close to unity, and are expected to be advantageous over SI-surface-plasmon waveguides, especially at long wavelengths. Such a waveguide was used to obtain lasing at 3.8 THz ( $\lambda \approx 79\ \mu\text{m}$ ) in pulsed mode up to a record high temperature of 137 K, whereas similar devices fabricated in SI-surface-plasmon waveguides had lower maximum lasing temperatures ( $\sim 92\ \text{K}$ ) due to the higher losses and lower confinement factors.

This thesis describes the theory, design, fabrication, and testing of terahertz quantum cascade laser devices. A summary of theory relevant to design is presented,

including intersubband radiative transitions and gain, intersubband scattering, and coherent resonant tunneling transport using a tight-binding density matrix model. Analysis of the effects of the complex heterostructure phonon spectra on terahertz QCL design are considered. Calculations of the properties of various terahertz waveguides are presented and compared with experimental results. Various fabrication methods have been developed, including a robust metallic wafer bonding technique used to fabricate metal-metal waveguides. A wide variety of quantum cascade structures, both lasing and non-lasing, have been experimentally characterized, which yield valuable information about the transport and optical properties of terahertz devices. Finally, prospects for higher temperature operation of terahertz QCLs are considered.

Thesis Supervisor: Qing Hu

Title: Professor

## Acknowledgments

I would like to first thank my advisor Qing Hu for giving me the opportunity to work on this project and for teaching and advising me along the way. Although this research often proved quite difficult with no assurance of success, Qing's persistence and relentless determination in the pursuit of terahertz lasing was motivating for my own efforts. I feel lucky to be engaged in such fast-paced research that is at once stimulating and challenging.

A large part of the success of this project is due to the excellent MBE growths obtained first from Michael Melloch at Purdue, and later from John Reno at Sandia.

I also wish to thank my many colleagues and research group-mates. I am particularly indebted to Bin Xu, from whom I inherited this research project. Bin Xu, and Jurgen Smet before him built the foundations of this research in our group. Much of my early training in this field came from Bin as well as Ilya Lyubormirsky. I am also grateful to my current colleagues on this project: Hans Callebaut, Sushil Kumar, and Stephen Kohen. As well pouring many long hours into this research performing experiments and simulations, they have contributed to the friendly and intellectually stimulating atmosphere in the lab. I also am grateful to many other colleagues for their help and useful conversations including Konstantinos Konistis, Noah Zamdmer, Farhan Rana, Mathew Abraham, Gert de Lange, Brian Riely, and Juan Montoya. Erik Duerr in particular has been a good friend to joke with as well as someone who was always willing to answer all my terahertz questions. I appreciate all the help given to me by the staff at MTL, CMSE, and RLE.

I am indebted to Amanda Gruhl, who provided no end of support during this effort, and kept me going throughout it all. She was good enough not to drag me out of lab at every opportunity, but also good enough to drag me out periodically to keep me sane. I promise to repay her patience when it is her turn. I want to thank my brother Peter, for all of his physics "knowledge." I can only hope he has paid similar attention to my take on law. Of course I want to thank my parents for all of their love and support throughout not only this, but two previous theses, many science

reports, and many basketball games. I couldn't imagine having better teachers, role models, and friends. Thanks to all of my family and friends who I haven't mentioned. I couldn't have done it without you.

# Contents

<b>1</b>	<b>Introduction</b>	<b>35</b>
1.1	Terahertz applications . . . . .	36
1.2	Terahertz sources . . . . .	39
1.2.1	Microwave upconversion . . . . .	41
1.2.2	Tubes . . . . .	41
1.2.3	Photomixing and downconversion . . . . .	41
1.2.4	Terahertz lasers . . . . .	42
1.3	Intersubband transitions and quantum cascade lasers . . . . .	44
1.3.1	Mid-infrared quantum cascade lasers . . . . .	45
1.3.2	Terahertz quantum cascade lasers . . . . .	48
1.3.3	Survey of terahertz QCL active regions . . . . .	52
1.4	Thesis overview . . . . .	55
<b>2</b>	<b>Intersubband laser theory and modeling</b>	<b>57</b>
2.1	Introduction . . . . .	57
2.2	Electronic states in heterostructures . . . . .	57
2.3	Intersubband radiative transitions and gain . . . . .	61
2.3.1	Spontaneous emission lifetime . . . . .	63
2.3.2	Stimulated emission . . . . .	65
2.3.3	Intersubband gain . . . . .	66
2.4	Nonradiative inter- and intra-subband transitions . . . . .	67
2.4.1	Polar longitudinal optical (LO) phonon scattering . . . . .	70
2.4.2	Electron-electron scattering . . . . .	74

2.5	Resonant tunneling transport and anticrossings . . . . .	79
2.5.1	Semiclassical “coherent” model . . . . .	80
2.5.2	Tight-binding density matrix approach . . . . .	81
2.6	Summary . . . . .	95
<b>3</b>	<b>Heterostructures and optical-phonons</b>	<b>97</b>
3.1	Introduction . . . . .	97
3.2	Complex phonon calculation . . . . .	101
3.3	Optimized design of three-level structure . . . . .	108
3.4	Conclusion . . . . .	115
<b>4</b>	<b>Waveguide design</b>	<b>117</b>
4.1	Introduction . . . . .	117
4.2	One-dimensional mode solver . . . . .	118
4.3	Free carrier effects . . . . .	120
4.3.1	Drude model . . . . .	123
4.3.2	Propagation in lightly doped regions . . . . .	124
4.3.3	Propagation in MQW active region . . . . .	125
4.4	Phonon coupling effects . . . . .	128
4.5	Terahertz waveguides . . . . .	129
4.5.1	Plasmon waveguide . . . . .	130
4.5.2	Metal-metal waveguide . . . . .	132
4.5.3	Semi-insulating-surface-plasmon waveguide . . . . .	135
4.5.4	Waveguide comparison . . . . .	137
<b>5</b>	<b>Experimental Setup and Fabrication</b>	<b>139</b>
5.1	Experimental setup . . . . .	139
5.2	Fabrication . . . . .	148
5.2.1	Etching . . . . .	149
5.2.2	Non-alloyed ohmic contacts . . . . .	152
5.3	Metal-metal waveguide fabrication . . . . .	153

5.3.1	Wafer bonding . . . . .	155
5.3.2	Substrate removal . . . . .	158
5.3.3	Metal-metal fabrication results . . . . .	160
<b>6</b>	<b>Survey of early designs</b>	<b>167</b>
6.1	Introduction . . . . .	167
6.2	M series . . . . .	170
6.2.1	M100 . . . . .	170
6.3	L series . . . . .	172
6.3.1	L30 . . . . .	172
6.4	T series . . . . .	176
6.4.1	T65 . . . . .	177
6.4.2	T150C . . . . .	186
6.4.3	T150E, T210E . . . . .	188
6.4.4	T150F . . . . .	192
6.4.5	T215G . . . . .	196
6.4.6	T215H . . . . .	197
6.4.7	Peaks of lateral modes . . . . .	199
6.4.8	T-series summary . . . . .	201
6.5	FL series (resonant phonon) . . . . .	203
6.5.1	FL125 . . . . .	203
6.5.2	FL175B . . . . .	209
6.5.3	FL175D . . . . .	213
6.5.4	FL179E, FL162G, and FL137H . . . . .	214
6.5.5	FL148L . . . . .	221
6.6	FFL140 . . . . .	224
<b>7</b>	<b>Terahertz quantum cascade lasers</b>	<b>229</b>
7.1	Introduction . . . . .	229
7.2	FL175C . . . . .	230
7.3	FL152Fmm . . . . .	241

7.4	FL148I-M1	248
7.5	CSL104	252
7.6	FL178C-M1	257
7.7	Parasitic current channels	263
7.8	Intra-miniband oscillator strengths	268
7.9	High temperature operation	272
7.10	Conclusions and Discussion	280
<b>A</b>	<b>Notes on dephasing and homogeneous broadening</b>	<b>285</b>
<b>B</b>	<b>Processing recipes</b>	<b>289</b>
B.1	Dry etching	289
<b>C</b>	<b>QCL rate equations</b>	<b>291</b>
C.1	Above threshold	293
C.2	Subthreshold	295

# List of Figures

1-1	Diagram of terahertz spectrum. . . . .	36
1-2	Radiated energy versus wavelength showing 30-K blackbody, typical interstellar dust, and key molecular line emissions in the terahertz (reprinted from Phillips and Keene [5]). . . . .	38
1-3	Performance of various cw sources in the sub-millimeter wave range. Figure is reprinted from Siegel (2002) [7]. . . . .	40
1-4	(a) Schematic of original quantum cascade laser, including conduction band profile and magnitude squared of wavefunctions. (b) In-plane momentum space ( $k_{\parallel}$ ) diagram of subbands and allowed relaxation paths via LO-phonon and photon emission. Figure is reprinted from Faist <i>et al.</i> [1]. . . . .	46
1-5	(a) Basic transport schematic for generic THz QCL. (b) In-plane momentum space ( $k_{\parallel}$ ) schematic with typical allowed relaxation paths for the $3 \rightarrow 2$ transition. . . . .	48
1-6	(a) Active region design for the chirped superlattice terahertz QCL demonstrated by Köhler <i>et al.</i> . (b) Mode pattern for semi-insulating surface plasmon waveguide. Figure is reprinted from Ref. [64]. . . . .	51
1-7	Schematic diagrams of successful terahertz QCL active regions. . . . .	53
1-8	Schematic diagram of a superlattice and its dispersion relation. . . . .	54
2-1	Important intersubband scattering mechanisms for the cases (a) $E_{fi} < E_{LO}$ and (b) $E_{fi} > E_{LO}$ . . . . .	68

2-2	Important (a) intrasubband and (b) bi-intrasubband scattering mechanisms. . . . .	70
2-3	(a) Schematic illustration of intersubband LO-phonon scattering process. (b) In-plane reciprocal lattice space diagram illustrating the relationship between initial and final electron wavevectors $k_i$ and $k_f$ and in-plane phonon wavevector $q_{\parallel}$ . . . . .	72
2-4	Average intersubband LO-phonon scattering rate for subband separation $E_{fi} = 13.3$ meV. Rate is calculated for $5 \rightarrow 4$ transition for FL175C device as shown in Fig. 7-1. Inset shows the raw scattering rate versus in-plane energy. . . . .	74
2-5	Intersubband electron-electron scattering rate versus upper state population $N_2$ ( $N_1 = 0$ ) calculated for separation $E_{fi} = 13.3$ meV according to (2.53), without screening. $21 \rightarrow 11$ rate is negligible. Rate is calculated for $5 \rightarrow 4$ transition for FL175C device as shown in Fig. 7-1 for $T_e = 77$ K. When calculating $\tau_{2 \rightarrow 1, tot}$ , the rate $\tau_{22 \rightarrow 11}$ is counted twice, since two electrons relax to the lower subband. . . . .	78
2-6	Anticrossing between two single well states. . . . .	80
2-7	(a) Schematic for semiclassical model for resonant tunneling through a barrier. (b) Hamiltonian using delocalized basis states with energies $E_A$ and $E_S$ , relaxation times $\tau_A$ and $\tau_S$ , and anticrossing gap $\Delta_0$ . . . . .	81
2-8	Schematic and Hamiltonian for (a) tight-binding resonant tunneling model with interaction $-\Delta_0/2$ , and (b) the analogous two-level model with dipole interaction $-\mu E(t)$ . The $2 \rightarrow 1$ relaxation time is $\tau$ . . . . .	82
2-9	Effective tunneling transport times $\tau_{eff}$ as calculated according to Eq. (2.64) versus (a) lifetime $\tau$ and (b) anticrossing gap $\Delta_0$ for a pure dephasing time of $T_2^* = 0.33$ ps. . . . .	86
2-10	Damped oscillation frequency $\omega_l$ versus the pure dephasing FWHM energy linewidth as given in Eq. (2.70). . . . .	88

2-11 (a) Schematic for three-level tight-binding resonant tunneling model with two-level interaction $-\Delta_0/2$ and with a third reservoir state $G$ and (b) the analogous three-level optical model with with dipole interaction $-\mu E(t)$ . The system is driven by pumping $P$ , and the population relaxation times into reservoir are $\tau_2$ and $\tau_1$ .	90
3-1 Band structure and wavefunction magnitude squared for a three-well module for (a) structure 1: M100 and (b) structure 2: L30, grown in GaAs/Al <sub>0.3</sub> Ga <sub>0.7</sub> As. The numbers above the layers represent layer thicknesses in nm.	98
3-2 Dispersion relation of interface phonon modes for M100, where $q$ is the in-plane wavevector.	104
3-3 Interface phonon potential for (a) six “GaAs-like” modes and (b) six “AlAs-like” modes plotted for GaAs at $q = 0.3 \text{ nm}^{-1}$ . Each mode is labeled with its energy in meV, and the vertical lines represent the heterointerfaces. The six modes of lowest energy contribute very little to scattering and are not plotted.	105
3-4 Confined “GaAs-like” modes for $m = 1$ (solid) and $m = 2$ (dashed) at $q = 0.3 \text{ nm}^{-1}$ . The vertical lines represent heterointerfaces. Although the confined mode in each layer is calculated separately, they are plotted together for simplicity.	106
3-5 The scattering rate from $n = 2$ to 1 for M100 is calculated as a function of initial in-plane electron energy $E_k$ using both bulk and complex phonon spectra. Solid lines are scattering rates calculated using bulk GaAs and AlAs modes. Scored lines indicate calculations using the phonon spectra of (a) GaAs/Al <sub>x</sub> Ga <sub>1-x</sub> As and (b) Al <sub>y</sub> Ga <sub>1-y</sub> As/AlAs structures for various values of $x$ and $y$ . The line for $x = 0.01$ is nearly indistinguishable from that of bulk GaAs, just as the line for $y = 0.99$ is nearly indistinguishable from that of bulk AlAs.	107

3-6 Scattering rate (a) $W_{21}$ and (b) $W_{31}$ for M100 as a function of $E_k$ for $E_{21} = 36.25$ meV ( $E_{31} = 53$ meV). Rates calculated using bulk modes are shown for comparison. The insets give scattering time (a) $\tau_{21}$ and (b) $\tau_{31}$ as a function of electron temperature for an upper subband population of $10^{10}$ cm $^{-2}$ . . . . .	109
3-7 Scattering rate (a) $W_{21}$ and (b) $W_{31}$ for M100 as a function of $E_k$ for $E_{21} = 47.0$ meV ( $E_{31} = 64$ meV). Rates calculated using bulk modes are shown for comparison. The insets give scattering time (a) $\tau_{21}$ and (b) $\tau_{31}$ as a function of electron temperature for an upper subband population of $10^{10}$ cm $^{-2}$ . . . . .	110
3-8 Maximum scattering rate versus subband separation $E_{21}$ for (a) M100 and (b) L30. Maximum scattering rate calculated with GaAs bulk modes is present for comparison. . . . .	112
3-9 Plot of the quantity $\tau_3(1 - \tau_{21}/\tau_{32})$ , which is proportional to the population inversion (Eq. (3.3)), versus the lifetime $\tau_{32}$ for (a) M100 and (b) L30. . . . .	113
4-1 GaAs calculated electron Hall (dashed curves) and drift (solid curves) mobilities at 77 K as a function of electron concentration with compensation as a parameter ( $N_A/N_D = 0.0, 0.15, 0.30, 0.45, 0.60$ , and $0.75$ ) (after Chin <i>et al.</i> [128]). Experimental Hall mobilities for a variety of grown structures are also shown. Note that the references cited in the figure do not correspond to those of this thesis. . . . .	122
4-2 Bulk plasma frequency $f_p$ in GaAs for $\epsilon/\epsilon_0 = 12.9$ and $m^*/m_0 = 0.067$ . . . . .	123
4-3 Real and imaginary part of the refractive index ( $\sqrt{\epsilon(\omega)} = n + ik$ ) for GaAs for free carrier densities $10^{15}, 10^{16}, 10^{17}, 10^{18}$ cm $^{-3}$ , using $\tau = 2.3, 0.7, 0.2, 0.1$ ps respectively, as given by mobility data in Fig. 4-1. . . . .	124
4-4 GaAs bulk free carrier loss above the plasma frequency for several carrier densities ( $\epsilon_r = 12.9$ ) using 77 K Drude scattering times[128]. . . . .	125

4-5 The relative permittivity $\epsilon_r$ and loss for GaAs in the terahertz due to phonon interactions are shown in (a) and (b) respectively. The solid lines are a dual oscillator fit to experimental data by Moore <i>et al.</i> [134] and the individual points are experimentally measured from several sources collected in Palik [135]. . . . .	128
4-6 Drawing of three terahertz waveguides. . . . .	129
4-7 Calculated loss and confinement of plasmon waveguide versus frequency for structure schematically shown in part (a). (b) Typical mode profile for plasmon waveguide at 5 THz. (c) and (d) Loss and confinement for structures with various lower cladding layer thicknesses both without (c) and with (d) active region doping at $5 \times 10^{15} \text{ cm}^{-3}$ ( $\tau = 1 \text{ ps}$ ). . . . .	131
4-8 Calculated loss and confinement of metal-metal waveguide versus frequency for structure schematically described in part (a). (b) Typical mode profile for metallic waveguide at 5 THz. (c) and (d) Loss and confinement for structures with various contact layer thicknesses both without (c) and with (d) active region doping at $5 \times 10^{15} \text{ cm}^{-3}$ ( $\tau = 1 \text{ ps}$ ). . . . .	133
4-9 Loss and confinement of metal-metal waveguide versus frequency for active regions of various thickness for an (a) undoped active region and (b) with the active region doped at $5 \times 10^{15} \text{ cm}^{-3}$ ( $\tau = 1 \text{ ps}$ ) (b). The model waveguide structure is the same as in Fig. 4-8 with 0.1- $\mu\text{m}$ -thick contact layers. . . . .	134
4-10 (a) Schematic and (b) typical mode pattern at 5 THz for SI-surface-plasmon waveguide. (c) Loss and confinement for structures with various contact layer thicknesses with active region doping at $5 \times 10^{15} \text{ cm}^{-3}$ ( $\tau = 1 \text{ ps}$ ). (d) and (e) Loss divided by confinement both without (d) and including (e) facet loss from an uncoated 2-mm-long Fabry-Pérot cavity. . . . .	136

5-1	Transmission of 0.25-mm-thick polypropylene sheet at room temperature used as window on device dewar (inset zooms in on 1–5 THz region). . . . .	140
5-2	(a) OFHC copper mount used for edge emission measurements to hold cone and chip carrier with device. (b) Diagram of Winston cone used for output coupling. The cone is formed by rotating a parabolic section about an axis (different from the parabolic axis). Radiation originating within the smaller aperture will emerge with an angle of divergence $\theta < \theta_{max}$ . . . . .	141
5-3	Purged blackbody spectra taken using Si bolometer at $4\text{ cm}^{-1}$ resolution with filters 1,2, and 3 (ignore relative intensity). . . . .	144
5-4	300 K blackbody spectra taken with the Ga:Ge photodetector at $4\text{ cm}^{-1}$ resolution before and after purging of the optical path. The sharp edge at $80\text{ cm}^{-1}$ is due to the detector cut-off. . . . .	145
5-5	Experimental setup for emission measurements. . . . .	146
5-6	(a) Calculated beamsplitter efficiency ( $4R \times T$ ) for a 0.83-mm-thick silicon wafer at $45^\circ$ . (b) Blackbody spectra taken with the bolometer at $0.125\text{ cm}^{-1}$ which display dips regularly spaced at $1.8\text{ cm}^{-1}$ due to the beamsplitter. . . . .	147
5-7	Typical SI-surface-plasmon ridge structure wet etched in $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (20:12:480). Device is $12\text{-}\mu\text{m}$ -high FSL120. . . . .	150
5-8	Schematic of band diagram for nonalloyed ohmic contact to <i>n</i> -GaAs. . . . .	152
5-9	Schematic diagram metal-metal bonding process. . . . .	154
5-10	Schematic diagram deposited metal layers for In-Au reactive bonding. . . . .	157
5-11	Gold-indium equilibrium phase diagram [153] . . . . .	158
5-12	SEM micrographs of typical metal-metal waveguide structures fabricated by Au-Au thermocompression bonding and dry etching in $\text{BCl}_3:\text{N}_2$ (15:5 sccm) (ECR-RIE). Active region is $\sim 7\text{ }\mu\text{m}$ thick. Device is T150Emm. . . . .	161

5-13 SEM micrographs of typical metal-metal waveguide structures fabricated by In-Au reactive bonding bonding and wet etching by (a)  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (10:6:480) (device is FL152Fmm) and (b)  $\text{H}_3\text{PO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (1:1:25) (device is FL162Gmm). Active regions are  $10\text{-}\mu\text{m}$  thick. . . 163

5-14 SEM micrographs of typical metal-metal waveguide structures fabricated by In-Au reactive bonding bonding and dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm) (ECR-RIE). Active region is  $10\text{ }\mu\text{m}$  thick. Device is FL137Hmm.163

5-15 (a), (b) SEM micrographs of In multilayer coated receptor wafer prior to bonding. The wafer in (b) has been scratched with tweezers, which apparently reveals smooth In underneath the “rocky” In-Au alloys. . 164

6-1 Schematic diagrams of the various multiple quantum well designs tested in this thesis. . . . . 168

6-2 (a) Self-consistent conduction band profile of M100 calculated using 65% conduction band offset. The device is grown in the  $\text{GaAs}/\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$  material system. The layer thicknesses are given in monolayers. The 7 ML collector barrier was designed to be delta doped at a sheet density of  $8.8 \times 10^{10} \text{ cm}^{-2}$  per module. (b) Anticrossings, energy separations, and oscillator strengths calculated for M100 (without self-consistency). 170

6-3 (a) Current and optical power versus voltage at 5 K for M100, taken at 50% duty cycle with a 40 kHz pulse repetition frequency. Inset gives  $L-I$ . (b) Spectra at 5 K measured using the photodetector with  $32 \text{ cm}^{-1}$  (0.96 THz) resolution. The device is  $200 \times 400 \text{ }\mu\text{m}^2$  with a  $15\text{-}\mu\text{m}$ -period grating with the conductive cap layer etched in between the grating. . . . . 171

6-4 (a) Conduction band profile of L30 calculated using 65% conduction band offset. The device is grown in the GaAs/Al <sub>0.3</sub> Ga <sub>0.7</sub> As material system. The layer thicknesses are given in monolayers. The 16 ML collector barrier was designed to be delta doped at a sheet density of $6.0 \times 10^{10} \text{ cm}^{-2}$ per module, although <i>C-V</i> measurements indicate a doping of $8.4 \times 10^{10} \text{ cm}^{-2}$ per module. (b) Anticrossing plot, energy separations, and oscillator strengths for L30. All data was calculated using 65% offset, no self-consistency. . . . .	173
6-5 (a) Current density, conductance, and optical power versus voltage at 5 K in pulsed mode for L30. Inset gives <i>L-I</i> . (b) Spectra at 5 K and 80 K for a bias of 1.6 V measured using the bolometer with $8 \text{ cm}^{-1}$ resolution. The inset shows a spectrum taken at a 4.0 V bias. The device is $400 \times 600 \mu\text{m}^2$ with a 15-μm-period grating. . . . .	174
6-6 Schematic diagram for the T-series of devices for the case of (a) incoherent sequential or (b) coherent resonant tunneling injection. . . . .	176
6-7 (a) Conduction band diagram and anticrossing diagram of T65 grown in GaAs/Al <sub>0.15</sub> Ga <sub>0.85</sub> As calculated with 65% conduction band offset. Layer thickness are given in monolayers. The 51 ML well was doped at $1.4 \times 10^{16} \text{ cm}^{-3}$ to yield a sheet density of $2.0 \times 10^{10} \text{ cm}^{-2}$ , which is approximately corroborated via <i>C-V</i> measurements. (b) Anticrossing plot and oscillator strengths for T65. . . . .	178
6-8 (a) <i>I-V</i> and (b) <i>L-V</i> and <i>L-I</i> (inset) for T65 at 5 K and 77 K. 715 × 515 μm <sup>2</sup> grating device with 15 μm period was tested using photodetector at 50% duty cycle. . . . .	179
6-9 T65 spectra at 5 K from (a) 715 × 505 μm <sup>2</sup> grating device with 20 μm period and (b) 715×515 μm <sup>2</sup> grating device with 15 μm period collected using photodetector with $4 \text{ cm}^{-1}$ (0.12 THz) resolution. . . . .	180

6-10 (a) Current and conductance versus voltage for several values of applied magnetic field at $T = 4.2$ K. (b) Position of $\Delta\ell = 2$ peaks in $G$ - $B$ plots and associated energy difference near the anticrossing of subbands $n=2$ and 1. The solid lines are the corresponding calculated energy differences $E_{32}$ and $E_{31}$ . . . . .	181
6-11 Current (b) and conductance (c) versus magnetic field ( $\mathbf{B} \parallel \mathbf{J}$ ) for several applied biases. A fan diagram that shows the crossing of Landau levels for an 18.9 meV energy separation is shown in (a). The arrows in (c) indicate the resonance peaks that correspond to the 2-1 anticrossing. . . . .	183
6-12 (a) $I$ - $V$ and $L$ - $V$ plots at 5 K and $L$ - $I$ plots at 5 K and 77 K for T65mm edge emitting device 1.5 mm wide and approximately 70 $\mu\text{m}$ deep. (b) Edge emitting spectra from same device taken at 5 K at 50% duty cycle using the photodetector with a spectral resolution of 16 $\text{cm}^{-1}$ . . . . .	185
6-13 $I$ - $V$ characteristic for T150C $140 \times 140 \mu\text{m}^2$ diode at 5 K, measured both in dc mode and pulsed mode. . . . .	187
6-14 (a) Conduction band diagram and anticrossing diagram of T150E/T210E grown in GaAs/ $\text{Al}_{0.15}\text{Ga}_{0.85}\text{As}$ calculated with 65% conduction band offset. Layer thickness are given in monolayers. The 51 ML well was doped at $1.4 \times 10^{16} \text{ cm}^{-3}$ to yield a sheet density of $2.0 \times 10^{10} \text{ cm}^{-2}$ . (b) Anticrossing plot and oscillator strengths for T150E/T210E. . . . .	187
6-15 (a) $I$ - $V$ and (b) $L$ - $V$ and $L$ - $I$ (inset) for T150E-3 at 5 K and 77 K. 80- $\mu\text{m}$ -wide and 1.9-mm-long ridge structure was tested using photodetector at 50% duty cycle. . . . .	189
6-16 (a) (c) Edge emitting spectra and $L$ - $I$ taken from 80- $\mu\text{m}$ -wide and 1.9-mm-long T150E-3 device at 5 K and 78 K. (b), (d) Edge emitting spectra and $L$ - $I$ taken from 80- $\mu\text{m}$ -wide and 1.9-mm-long T210E device at 5 K and 78 K. Spectra were taken at 50% duty cycle using the photodetector with a spectral resolution of 8 $\text{cm}^{-1}$ . . . . .	191

6-17 (a) $I$ - $V$ and $L$ - $V$ plots at 5 K and $L$ - $I$ plots at 5 K and 77 K for T150E-2mm edge emitting device that is 100- $\mu\text{m}$ wide and 1.875-mm long. (b) Edge emitting spectra from same device taken at 5 K at 50% duty cycle using the photodetector with a spectral resolution of $8 \text{ cm}^{-1}$ .	192
6-18 Edge emitting spectra and $L$ - $I$ from a T150F 145- $\mu\text{m}$ -wide and 0.775-mm-long device taken at 5 K at 50% duty cycle using the photodetector with a spectral resolution of $8 \text{ cm}^{-1}$	194
6-19 (a) $I$ - $V$ , $L$ - $V$ , and $L$ - $I$ plots at 5 K for T150Fmm edge emitting device that was 100- $\mu\text{m}$ wide and 0.975-mm long. (b) Edge emitting spectra from same device taken at 5 K at 50% duty cycle using the photodetector with a spectral resolution of $8 \text{ cm}^{-1}$	195
6-20 (a) $I$ - $V$ , $L$ - $V$ , and $L$ - $I$ plots at 5 K for T215G edge emitting device 150- $\mu\text{m}$ wide and 2.70-mm long. (b) Edge emitting spectra from same device taken at 5 K at 5% duty cycle using the photodetector with a spectral resolution of $8 \text{ cm}^{-1}$	197
6-21 (a) $I$ - $V$ , $L$ - $V$ , and $L$ - $I$ plots at 5 K for T215H edge emitting device 100 $\mu\text{m}$ wide and 2.65 mm long. (b) Edge emitting spectra from same device taken at 5 K at 5% duty cycle using the photodetector with a spectral resolution of $8 \text{ cm}^{-1}$	198
6-22 Conduction band diagram and anticrossing diagram of FL125 grown in GaAs/ $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ calculated with (a), (b) 65% conduction band offset, and (c), (d) 80% offset. Layer thickness are given in monolayers. The 62 ML well was doped at $2.3 \times 10^{16} \text{ cm}^{-3}$ to yield a sheet density of $4.0 \times 10^{10} \text{ cm}^{-2}$ , however $C$ - $V$ measurements indicate a doping of $3.4 \times 10^{10} \text{ cm}^{-2}$ per module.	204
6-23 Spectra and $L$ - $I$ from 145- $\mu\text{m}$ -wide, 260- $\mu\text{m}$ -long edge emitting FL125 device at $T = 5 \text{ K}$ measured with the photodetector at $8 \text{ cm}^{-1}$	205
6-24 (a) $I$ - $V$ and (b) $L$ - $V$ and $L$ - $I$ (inset) for 100- $\mu\text{m}$ -wide, 1.125-mm-long FL125mm edge emitting device with metal-metal waveguide taken with 5% duty cycle	206

6-25 Current and optical power at NDR, and NDR bias, for 170- $\mu\text{m}$ -wide, 0.76-mm-long edge emitting FL125mm device taken at 5% duty cycle versus heat sink temperature. . . . .	207
6-26 Spectra for 100- $\mu\text{m}$ -wide, 1.125-mm-long FL125mm edge emitting device with metal-metal waveguide taken with 50% duty cycle at (a) 5 K and (b) 78 K heat sink temperature. Spectra were taken with photodetector with $4 \text{ cm}^{-1}$ resolution. . . . .	208
6-27 (a) Conduction band diagram of FL175B grown in GaAs/Al <sub>0.3</sub> Ga <sub>0.7</sub> As calculated with 80% offset. Layer thicknesses are given in monolayers. The 60 ML well was doped at $2.3 \times 10^{16} \text{ cm}^{-3}$ to yield a sheet density of $3.9 \times 10^{10} \text{ cm}^{-2}$ per module. (b) Anticrossing diagram and oscillator strengths versus bias. . . . .	210
6-28 (a) <i>I-V</i> and (b) <i>L-V</i> and <i>L-I</i> (inset) for 200- $\mu\text{m}$ -wide, 0.75-mm-long FL175B edge emitting device with SI-surface-plasmon waveguide taken with 50% duty cycle. . . . .	211
6-29 Current and optical power at NDR, and NDR bias, for 150- $\mu\text{m}$ -wide, 2.4-mm-long edge emitting FL175B device taken at 5% duty cycle versus heat sink temperature. . . . .	211
6-30 (a) Spectra from 200- $\mu\text{m}$ -wide, 0.75-mm-long edge emitting FL175B device with 10- $\mu\text{m}$ lateral contact separation and (b) spectra from 150- $\mu\text{m}$ -wide, 2.0-mm-long, rear HR coated edge emitting FL175B device with 50- $\mu\text{m}$ lateral contact separation. Data was at 5 K taken with photodetector at $8 \text{ cm}^{-1}$ resolution. . . . .	212

6-31 (a) Self-consistent conduction band profile of FL175D calculated using 80% conduction band offset. The device is grown in the GaAs/Al <sub>0.15</sub> Ga <sub>0.85</sub> As material system. Beginning with the left injection barrier, the layer thicknesses in Å are 65/86/27/71/38/163/27/104 (as grown 4.7% above design). The 163 Å well is doped at $n = 1.9 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of $3.1 \times 10^{10} \text{ cm}^{-2}$ per module. (b) <i>I-V</i> taken at 5 K and 77 K of a $200 \times 200 \mu\text{m}^2$ FL175D diode taken with 400-ns pulses at a 1-kHz pulse repetition frequency. . . . .	213
6-32 (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL175D. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction. . . . .	214
6-33 (a) Self-consistent conduction band profile of FL179E calculated using 80% conduction band offset. The device is grown in the GaAs/Al <sub>0.15</sub> Ga <sub>0.85</sub> As material system. Beginning with the left injection barrier, the layer thicknesses in Å are 41/85/27/71/33/158/22/98 (as grown 3.4% below design). The 158 Å well is doped at $n = 1.8 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of $2.8 \times 10^{10} \text{ cm}^{-2}$ per module. (b) <i>I-V</i> taken at 5 K and 77 K of FL175E $200 \times 200 \mu\text{m}^2$ diode taken with 400-ns pulses at a 1-kHz pulse repetition frequency. . . . .	216
6-34 (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL179E. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction . . . . .	217

6-35 (a) Self-consistent conduction band profile of FL162G calculated using 80% conduction band offset. The device is grown in the GaAs/Al <sub>0.15</sub> Ga <sub>0.85</sub> As material system. Beginning with the left injection barrier, the layer thicknesses in Å are 41/71/16/63/22/63/35/152/22/101 (as grown 3.7% below design). The 152 Å well is doped at $n = 1.9 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of $2.9 \times 10^{10} \text{ cm}^{-2}$ per module. (b) <i>I-V</i> , <i>L-V</i> , and <i>L-I</i> taken at 5 K at 5% duty cycle with bolometer with filter 2 from a 100-μm-wide, 0.86-mm-long wet-etched FL162Gmm device. . . . .	217
6-36 (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL162G. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction . . . . .	218
6-37 Self-consistent conduction band profile of FL137H calculated using 80% conduction band offset. The device is grown in the GaAs/Al <sub>0.15</sub> Ga <sub>0.85</sub> As material system. Beginning with the left injection barrier, the layer thicknesses in Å are 35/67/16/61/21/61/35/152/16/107/21/85 (as grown 5.6% below design). The 107 Å well is doped at $n = 2.7 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of $2.9 \times 10^{10} \text{ cm}^{-2}$ per module. . . . .	219
6-38 (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL137H. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction . . . . .	219
6-39 (a) <i>I-V</i> , <i>L-V</i> , and <i>L-I</i> taken at 5% duty cycle with photodetector and (b) spectra taken at 5% duty cycle at 5 K from a rear facet HR coated 150-μm-wide, 0.84-mm-long wet-etched FL137Hmm device with photodetector and bolometer at 8 cm <sup>-1</sup> resolution. . . . .	220
6-40 Self-consistent conduction band profile of FL148L calculated using 80% conduction band offset (at design thickness). Layer thicknesses are given in monolayers. The 59 ML well is doped at $n = 1.8 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of $3.0 \times 10^{10} \text{ cm}^{-2}$ per module. . . . .	221

6-41 (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL148L. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as designed . . . . .	222
6-42 (a) $I$ - $V$ , $L$ - $V$ , and $L$ - $I$ taken at 5% duty cycle and (b) spectra taken at 20% duty cycle at 5 K from an uncoated 150- $\mu\text{m}$ -wide, 1.60-mm-long FL148Lmm device. Spectra were taken using photodetector and with 8 $\text{cm}^{-1}$ resolution. . . . .	223
6-43 Self-consistent conduction band profile of FFL140 calculated using 80% conduction band offset (as grown at design thickness). Layer thicknesses are given in monolayers. The 57 ML well is doped at $n = 2.4 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of $3.9 \times 10^{10} \text{ cm}^{-2}$ per module. . . . .	225
6-44 (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FFL140. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction . . . . .	225
6-45 (a) $I$ - $V$ , $L$ - $V$ , and $L$ - $I$ taken at 2.5% duty cycle and (b) spectra taken at 5% duty cycle at 5 K from a rear facet HR coated 150- $\mu\text{m}$ -wide, 1.07-mm-long FFL140 device. Spectra were taken using photodetector and with 8 $\text{cm}^{-1}$ resolution. . . . .	226
7-1 (a) Self consistent conduction band profile of FL175C (calculated using 80% band offset) biased at 64 mV/module, which corresponds to a field of 12.2 kV/cm. The device is grown in the GaAs/Al <sub>0.15</sub> Ga <sub>0.85</sub> As material system. The four well module is outlined. Beginning with the left injection barrier, the layer thickness in Å are 54/78/24/64/38/148/24/94 (as grown 4.2% below design). The 148 Å well is doped at $n=1.9 \times 10^{16} \text{ cm}^{-3}$ , yielding a sheet density of $2.8 \times 10^{10} \text{ cm}^{-2}$ . (b) Optical mode profile of SI-surface-plasmon waveguide. . . . .	231

- 7-2 (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL175C. All data was calculated using an 80% conduction band offset and layer thicknesses as measured by x-ray diffraction. . . . . 232
- 7-3 (a)  $I$ - $V$  and  $L$ - $V$  and (b) subthreshold spectra taken at 5 K and  $L$ - $I$  curves from a uncoated 150- $\mu\text{m}$ -wide, 0.95-mm-long FL175C device at 5% duty cycle. Data ends immediately before onset of negative differential resistance (NDR). Spectra were taken using photodetector and with  $8 \text{ cm}^{-1}$  resolution. . . . . 233
- 7-4 Emission spectrum above threshold for 150- $\mu\text{m}$ -wide, 1.19-mm-long FL175C device with rear HR coating. Device is biased at 1.64 A with 100-ns pulses repeated at 10 kHz, at  $T = 5 \text{ K}$ . Spectra were measured with the Ge:Ga photodetector. The inset shows an expanded view of spectra at various bias points, offset for clarity. The spectra were collected in linear-scan mode, and the measured linewidth is limited by the instrumental resolution of  $0.125 \text{ cm}^{-1}$  (3.75 GHz). . . . . 234
- 7-5 (a) Collected light versus current at various temperatures for 150- $\mu\text{m}$ -wide, 1.19-mm-long FL175C device. Data is taken using 100-ns pulses repeated at 2 kHz with the Ge:Ga photodetector. The inset is a semi-log plot of the threshold current density  $J_{th}$  versus temperature. (b) Applied bias and peak optical power versus current, collected at various duty cycles with a pulse repetition frequency of 1 kHz. The data was taken with heat sink temperatures of 5 K, except at 10% and 50% duty cycle, which had  $T = 8 \text{ K}$  and 15 K, due to the large power dissipation. The pyroelectric detector was used and the collected power level was calibrated using the thermopile. . . . . 236
- 7-6 Edge emission electroluminescence spectrum taken from lateral edge of longitudinally cleaved  $\sim 180\text{-}\mu\text{m}$ -long, 1.03-mm-wide FL175C device at  $\sim 860 \text{ A/cm}^2$ . Lorentzian fit to spectrum is shown. . . . . 237

7-7	<i>L-I</i> taken with 200-ns pulses repeated at 1.5 kHz from 200- $\mu\text{m}$ -wide, 2.56-mm-long FL175C ridge with rear HR coating. The <i>V-I</i> curve is from a different device with a smaller area, and only the current density is directly comparable. The 78 K spectrum was taken with 200-ns pulses repeated at 10 kHz. . . . .	238
7-8	Self-consistent conduction band profile for FL152F at a field of 9.5 kV/cm (60.5 mV/module) calculated using 80% conduction band offset. Beginning with the left injection barrier, the layer thicknesses in Å are 44/77/28/69/36/157/17/102/25/83 (as grown 2.5% below design). The 102 Å well is doped at $n = 2.9 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of $3.0 \times 10^{10} \text{ cm}^{-2}$ per module. . . . .	242
7-9	(a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL152F. All data was calculated using 80% band offset and layer thicknesses as measured by x-ray diffraction. . . . .	243
7-10	<i>L-I</i> and <i>V-I</i> taken in pulsed mode for 150- $\mu\text{m}$ -wide, 2.59-mm-long FL152Fmm ridge with metal-metal waveguide and rear HR coating. Data was measured using 100-ns pulses repeated at 1 kHz. Note that the current-voltage characteristic was measured using a similar, but smaller device, so only the current density scale is applicable. Spectra taken at 5 K and 77 K (with liquid nitrogen cooling) using 100-ns pulses repeated at 10 kHz are also shown. Emitted light was measured using a Ge:Ga photodetector, and the spectra were collected in linear-scan mode with a resolution of $0.125 \text{ cm}^{-1}$ . . . . .	244
7-11	Mode intensities (solid lines) and the real part of the dielectric constant $\epsilon(\omega)$ (dashed lines) for the (a) SI-surface-plasmon waveguide and (b) metal-metal waveguide. The dotted line represents the square of the modal effective index. . . . .	247

- 7-12 Self-consistent conduction band profile for FL148I-M1 calculated using 80% band offset. Beginning with the left injection barrier, the layer thicknesses in Å are 58/78/26/69/40/161/26/106/32/83 (as grown 1.8% above design). The 161/26/106-Å-thick layers are doped at  $n = 1.0 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $2.9 \times 10^{10} \text{ cm}^{-2}$  per module. 249
- 7-13 (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL148I-M1. All data was calculated using an 80% conduction band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction. . . . . 250
- 7-14 (a)  $I-V$ ,  $L-V$ , and conductance (digitally taken  $dI/dV$ ) versus voltage at 5 K in pulsed mode for a 200- $\mu\text{m}$ -wide, 1.47-mm-long FL148I-M1mm ridge with metal-metal waveguide (uncoated). (b)  $L-Is$  versus temperature and  $V-I$  at 5 K for the same device. Data was measured using 200-ns pulses repeated at 10 kHz with the photodetector. An inset shows  $J_{th}$  versus heat sink temperature. A spectrum taken at 5 K for a bias of 1.58 A is also shown. . . . . 250
- 7-15 Self-consistent conduction band profile for CSL104 calculated using 80% band offset. Beginning with the left injection barrier, the layer thicknesses in Å are 43/188/8/158/6/117/25/103/29/102/30/108/33/99. The 102 Å well is doped at  $n = 4 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $4.1 \times 10^{10} \text{ cm}^{-2}$  per module. These layer thicknesses are the design values, as opposed to the actual growth values. . . . . 253
- 7-16 (a) Spectra at 5 K taken from 100- $\mu\text{m}$ -wide, 2.97-mm-long CSL104-3 ridge using 100-ns pulses repeated at 10 kHz ( $L-I$  in inset). (b)  $L-Is$  versus temperature for 200- $\mu\text{m}$ -wide, 2.97-mm-long uncoated ridge. A  $V-I$  at 5 K is plotted from a smaller device (100- $\mu\text{m}$ -wide, 2.97-mm-long), so only the current density should be compared. Data was measured using 90-ns pulses repeated at 1 kHz with the photodetector. An inset shows  $J_{th}$  versus heat sink temperature. . . . . 255

7-17 <i>L</i> - <i>Is</i> at various duty cycles for 200- $\mu\text{m}$ -wide, 2.97-mm-long uncoated ridge, measured at 5 K with pyroelectric detector. The pulse width was varied for a constant pulse repetition frequency of 1 kHz. . . . .	255
7-18 (a) Pulsed <i>I</i> - <i>V</i> characteristic for a $200 \times 200 \mu\text{m}^2$ FL178C-M1mm diode and 100- $\mu\text{m}$ -wide, 1.45-mm-long ridge structure that lases. (b) Typical <i>L</i> - <i>Is</i> and <i>V</i> - <i>I</i> from FL178C-M1mm devices. The <i>L</i> - <i>I</i> in upper panel is from 60- $\mu\text{m}$ -wide, 2.48-mm-long ridge. The <i>V</i> - <i>I</i> was taken using a smaller device (100- $\mu\text{m}$ -wide, 1.45-mm-long), so only the current density axis is valid. The <i>L</i> - <i>Is</i> in the lower panel are from a 150- $\mu\text{m}$ -wide, 2.74-mm-long FL178C-M1mm ridge. Inset gives $J_{th}$ versus temperature. Data was measured using 200-ns pulses repeated at 1 kHz with the photodetector. . . . .	258
7-19 Spectra at various heat sink temperatures from a 100- $\mu\text{m}$ -wide, 2.74-mm-long FL178C-M1mm uncoated ridge. Spectra were taken using 200-ns pulses repeated at 5 kHz with the photodetector, all with approximately the same current densities of 1170–1200 A/cm <sup>2</sup> . . . . .	260
7-20 Conduction band profile for FL175C (80% barrier offset) calculated at the parasitic 1'-3 anticrossing point. . . . .	263
7-21 Actual and predicted parasitic current densities using Eq. (7.3), with $\tau_{\parallel}^{-1} = \tau^{-1} + T_2^{*-1}$ as per Eq. (2.61). . . . .	264
7-22 Magnitude squared wavefunctions for multi-well 30/5 ML superlattices with 15% Al barriers. . . . .	271
7-23 Calculated threshold current density using Eq. (7.14) for various parameters. The radiative energy was $E_{32} = 14$ meV, $\tau_2 = 0.25$ ps, $g_{th} = 30 \text{ cm}^{-1}$ , and $\Delta\nu = 1$ THz. . . . .	275
7-24 Thermal simulations of 150- $\mu\text{m}$ -wide 10- $\mu\text{m}$ -high ridge structure using an active region thermal conductivity of (a) $\kappa_{AR} = 0.1 \text{ W}/(\text{cm K})$ and (b) $\kappa_{AR} = 1.0 \text{ W}/(\text{cm K})$ . A heat sink temperature of 10 K is used, the substrate thickness is 600 $\mu\text{m}$ with $\kappa_{sub} = 10 \text{ W}/(\text{cm K})$ , and the power density is $10^7 \text{ W cm}^{-3}$ in the active region. . . . .	278

C-1 Schematic for coupled carrier and photon population model for quantum cascade laser with non-unity injection efficiency $\eta$ .	292
--	-----



# List of Tables

3.1	Optical phonon scattering times for electrons with sufficient kinetic energy to scatter from $n = 3 \rightarrow 2$ via GaAs-like and AlAs-like phonons.	100
3.2	Average intersubband scattering rates for M100 at T=10 K for different subband separations. . . . .	111
3.3	Average intersubband scattering rates for L30 at T=10 K for different subband separations. . . . .	114
6.1	Calculated waveguide parameters for various growths in the T150E/T210E family. . . . .	188
6.2	Measured spontaneous output powers for edge emitting devices in the T150E family. Power was collected in face-to-face mode using the photodetector at 50% duty cycle at 10 kHz. The units of peak power correspond to the lock-in amplifier voltage (in bandpass mode). Due to optical alignment, there is some uncertainty in the measured power levels. . . . .	193
6.3	Measured output powers for various T150F edge emitting devices collected in face-to-face mode with the photodetector. The units of peak power correspond to the lock-in amplifier voltage (in bandpass mode) for 50% duty cycle at 10 kHz. Due to optical alignment, there is some uncertainty in the measured power levels. . . . .	196

6.4	Measured spontaneous output powers for various T215G edge emitting devices collected in face-to-face mode with the photodetector. Power was measured at 5% duty cycle (50% duty pulse train at 10 kHz, with 332 ns pulses at 300 kHz) but scaled up by a factor of 10 to be comparable with that listed in Tables 6.2 and 6.3. . . . .	197
6.5	Measured spontaneous output powers for various T215H edge emitting devices collected in face-to-face mode with the photodetector. Power was measured at 5% duty cycle but scaled up by a factor of 10 to be comparable with that listed in Tables 6.2 and 6.3. . . . .	199
6.6	Injection barrier thicknesses $t_{inj}$ along with injection anticrossings (calculated for 65% band offset) and peak injection current density $J_0$ for T-series devices. . . . .	202
7.1	Characteristics of FL152Fmm devices that successfully lased. All listed devices had rear HR coatings. Devices marked with “*” were cleaved from edge of wafer piece, where the bonding quality was poor. . . . .	246
7.2	Characteristics of four chirped superlattice (CSL) growths. Only CSL104-3 lased. . . . .	253
7.3	Characteristics of dry-etched, uncoated CSL104-3 lasing devices. The last two devices were longitudinally cleaved for lateral emission. . . . .	256
7.4	Characteristics of FL178C-M1mm devices that successfully lased. Listed devices were uncoated. . . . .	261
7.5	Characteristics of FL178C-M1 (SI-surface-plasmon) devices that successfully lased. Listed devices were rear facet HR coated. . . . .	262
7.6	Current densities $J_0$ and $J_{para}$ at injection point and parasitic point respectively, and calculated injection and parasitic anticrossing gaps $\Delta_0$ and $\Delta_{para}$ for various devices. Devices with no data for $J_0$ had no clearly defined NDR at the design point. The sheet density per module $N_{mod}$ is given, although only for FL125 was the doping measured. All data calculated with 80% conduction band offset. . . . .	265

7.7 Oscillator strengths for various transitions for $n$ -well superlattices. Quantities in parentheses are well/barrier thicknesses for structures in monolayers (1 ML=2.825 Å). 15% aluminum barriers were used. Structures were adjusted to keep the relevant transition energy approximately constant at 14.0 meV, and the lower radiative state 26.8 meV from the well bottom. . . . .	270
7.8 Drude relaxation times and waveguide losses for SI-surface-plasmon waveguide and metal-metal waveguide calculated for FL175C using the Drude times in this chapter, which roughly corresponds to 77 K, and at 300 K. . . . .	276



# Chapter 1

## Introduction

The terahertz (THz) frequency range is an underused region of the electromagnetic spectrum which lies between ranges covered by electronic techniques (RF and microwave), and photonic techniques (the infrared and visible). In this thesis, I will generally use the term terahertz to describe the band of frequencies from 1–10 THz, which corresponds to a wavelength of 30–300  $\mu\text{m}$ . However, terminology varies, as this range is also known as the far-infrared, or sub-millimeter range, and the lower limit is often extended down to 300 GHz ( $\lambda = 1 \text{ mm}$ ) or even 100 GHz ( $\lambda = 3 \text{ mm}$ ). This range remains technologically underdeveloped, mostly due to the lack of economical, compact, and efficient components. However, interest in the terahertz has been increasing dramatically, especially in the last decade, and a variety of new components and applications promise rapid development of this previously underused spectral region.

Perhaps the largest barrier to development in the terahertz has been the lack of compact, coherent, low-cost radiation sources. In this thesis, I describe the successful development of terahertz semiconductor lasers that operate via unipolar intersubband transitions in heterostructure quantum wells. These lasers are essentially quantum cascade lasers (QCLs)[1] that operate below the semiconductor *Reststrahlen* band, which lies at 8–9 THz in GaAs. Operation at such long wavelengths brings unique challenges not present for mid-infrared QCLs, and requires fundamentally different approaches. Terahertz quantum cascade lasers promise to provide powerful new sources

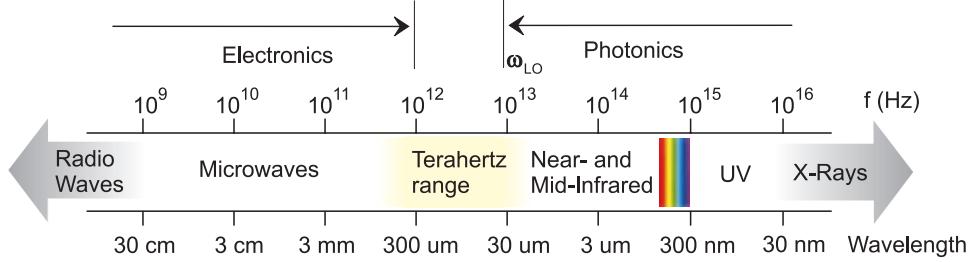


Figure 1-1: Diagram of terahertz spectrum.

of long-wavelength radiation that may impact applications from astronomical spectroscopy to imaging.

## 1.1 Terahertz applications

The primary, and oldest, application involving terahertz radiation is spectroscopy. The far-infrared range has long been of scientific interest, particularly because many chemical species have very strong characteristic rotational and vibrational absorption lines here [2]. Specifically, for gasses the typical absorption strengths are 10<sup>3</sup>–10<sup>6</sup> stronger than in the microwave region. Absorption strengths then fall off exponentially toward the infrared past their peak in the terahertz [3]. Spectroscopy was originally performed with incoherent thermal sources and Fourier-transform spectrometers using cryogenic bolometric detection. However, this becomes difficult below 1 THz, and high-resolution spectroscopy requires inconveniently long mirror travel lengths. Therefore the largest application for continuous-wave (cw) sources has been as local oscillators for heterodyne spectroscopy. In this technique, the signal of interest is mixed with the local oscillator signal to downconvert it to an intermediate frequency where it can be amplified and processed. Thus sensitive, high-resolution spectroscopy can be performed without the need for an illuminating source, especially since the energies of terahertz radiation are often comparable or less than the source temperatures (1 THz  $\sim$  4 meV  $\sim$  48 K). It is also possible to perform active spectroscopy, where tunable narrowband terahertz radiation is detected after transmission or reflection through some medium of interest.

The primary interest in terahertz spectroscopy has come from the astronomical and space science community. It is estimated that one-half of the total luminosity of the galaxy and 98% of the photons emitted since the Big Bang fall into the terahertz range [4]. Much of this radiation is emitted by cool interstellar dust inside our own and in other galaxies, and thus study of the continuum radiation, as well the discrete lines emitted by light molecular species gives insight into star formation and decay [5]. A schematic of some important resonances in interstellar dust as well as comparable blackbody curves is shown in Fig. 1-2. The far-infrared also contains information on the cosmic background, and very distant newly formed galaxies. However, most terahertz astronomy must be performed on satellite platforms or at high altitude due to the strong atmospheric absorption that results from pressure broadened water and oxygen lines. Several far-infrared and sub-millimeter satellites have been launched and more are planned in the future. Examples include the Submillimeter Wave Astronomy Satellite (SWAS) launched 1998, the satellite Herschel (formerly known as the Far InfraRed and Submillimeter space Telescope (FIRST)) scheduled for 2007 launch [6], as well as several other proposed satellites. Herschel includes a heterodyne instrument that provides coverage from 480 GHz to 1.25 THz in five bands and 1.41 THz to 1.91 THz in two bands.

Atmospheric science is another area of interest for terahertz spectroscopy. Thermal emission from gasses in the stratosphere and upper troposphere such as water, oxygen, chlorine, and nitrogen compounds is useful for the study of chemical processes related to ozone depletion, pollution monitoring, and global warming [7, 3]. The Earth observing system Aura (EOS-Aura) is scheduled for launch in 2004, and will include a heterodyne band at 2.5 THz specifically for monitoring atmospheric OH levels [8]. Atmospheric observation of other planets and small bodies such as asteroids, moons, and comets, and potentially even extra-solar planets is also facilitated far-infrared spectroscopy.

Other spectroscopic applications include plasma fusion diagnostics, where temperature profiles of the plasma can be obtained [9]. Efficient gas spectroscopy and identification is also promising, since absorption lines are stronger in the terahertz

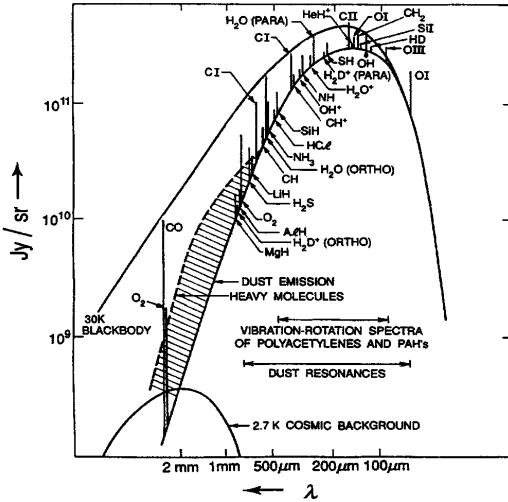


Figure 1-2: Radiated energy versus wavelength showing 30-K blackbody, typical interstellar dust, and key molecular line emissions in the terahertz (reprinted from Phillips and Keene [5]).

than in the microwave or infrared.

Terahertz, or “T-Ray” imaging is a relatively new field, following its development in the mid 1990s at Bell Laboratories [10], but has seen tremendous scientific advancement and commercial interest. Most imaging is performed using ultra-short (broadband) terahertz pulses generated using femtosecond optical pulses incident on a photoconductor or non-linear crystal (see Sec. 1.2.3). Using this system, the transmitted or reflected pulse can be coherently detected using room temperature detection methods, despite the low average power level ( $nW-\mu W$ ) of the pulses. Imaging is performed either by rapidly scanning the sample or via free-space two-dimensional sampling on charge-coupled device (CCD) arrays [11]. Imaging can take place via amplitude and phase transmission or reflection, via time of flight for the pulse, or via coarse time-domain spectral analysis of the received broadband pulse.

The development of imaging using terahertz radiation is of interest specifically because many materials are transmissive in the terahertz that are opaque at visible frequencies and vice versa. Non-invasive medical imaging of teeth or sub-dermal melanoma has been demonstrated. Detection of concealed weapons at airports or currency forgeries is also a commonly cited application. Monitoring of water levels

in plants, fat content in packaged meats, and manufacturing defects in automotive dashboards and high voltage cables have all been performed.

While terahertz imaging is presently dominated by the use of pulsed time domain systems, there is likely still a role for high-power cw sources. Current limiting factors for imaging are the low average power levels of terahertz pulses and the high cost of the femtosecond lasers required to produce them. New terahertz sources would be welcome. Additionally, imaging can be performed using cw sources with direct detection or with coherent detection [12, 13]. Heterodyne detection can also be used to provide imaging along with high resolution spectroscopic capability [14]. The use of cw sources is advantageous when narrowband imaging is desired, since the short pulsed systems typically have frequency content of approximately 0.5–2.5 THz or higher. In all of these cases versatile, powerful THz lasers would be a boon.

Terahertz sources are also sought for additional applications that do not fit neatly into the above categories. Non-contact semiconductor wafer characterization of spatially dependent doping and mobility can be performed using terahertz radiation in a magnetic field [15]. Lasers would also be useful for high-bandwidth wireless communication. Such systems would be useful between satellites where atmospheric absorption is negligible. Alternatively, the high atmospheric absorption on earth can be advantageous, creating the opportunity for short-distance secure point-to-point communication links.

While the terahertz range has traditionally been underused, there is an increasing interest in developing commercial as well as scientific applications. This interest is made possible by the development, especially in the past decade, of novel detectors and sources that operate in this frequency regime.

## 1.2 Terahertz sources

Early far-infrared spectroscopy generally used thermal blackbody sources for generation. Due to the low-power, incoherent, and broadband nature of the radiation available by this method, the development of efficient coherent sources has long been

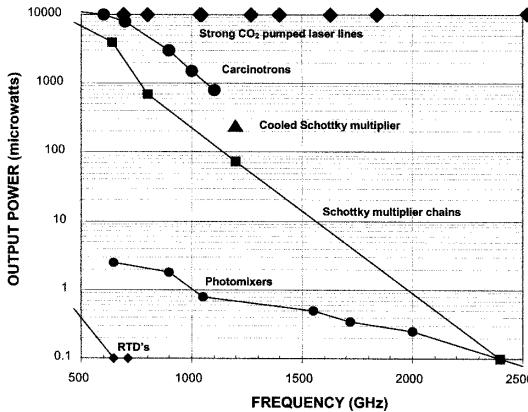


Figure 1-3: Performance of various cw sources in the sub-millimeter wave range. Figure is reprinted from Siegel (2002) [7].

both a priority and limiting factor for applications. Generation of coherent terahertz radiation has traditionally involved either extending electronic techniques to higher frequencies, or extending photonic sources to longer wavelengths. Electronic semiconductor devices have trouble operating far beyond 100 GHz, as resistive and reactive parasitics as well as transit time limitations cause high frequency roll-offs. For laser sources operating in the terahertz the small energy level separations ( $1\text{--}10\text{ THz} \rightarrow 4\text{--}40\text{ meV}$ ) make obtaining population inversion difficult. This problem is especially acute in the effort to obtain solid-state lasers, where the energies are comparable to phonon resonances, which tend to depopulate excited states, especially above cryogenic temperatures. An ideal source would be continuous wave, narrowband, highly tunable, have at least milliwatts of power. Additionally it should be economical, compact, operate at room temperature, and be efficient without high voltage, high power or magnetic field requirements. Unfortunately, no one source fulfills all, or even most, of these requirements. At this time, the most successful sources include upconversion from microwave oscillators, optically pumped gas lasers, as well as various methods for optical downconversion including photomixers. These methods for generation, in addition to a few others, will be briefly described in the section below. The typical power levels from cw devices are shown in Fig. 1-3. Sources such as free-electron lasers and gyrotrons will not be discussed due to their overwhelming size and expense.

### **1.2.1 Microwave upconversion**

The dominant method of obtaining low-frequency terahertz radiation (0.5–2 THz) is nonlinear multiplication of a lower frequency oscillator (100–200 GHz) by chains of Schottky doublers and triplers [7]. Such a device delivers continuous-wave narrowband power with limited tunability ( $\sim 10\%$ ) [16] and is suitable for use as a local oscillator (LO) for mixing. Multiplication is a robust technology and will be used as the LO on Herschel to pump various mixers in bands up to 1.9 GHz [6]. However, the output power falls off rapidly with increasing frequency due to reduced multiplication efficiency, and is sub-milliwatt for  $f > 1$  THz[17] and is microwatt level or below for  $f > 1.6$  THz.

### **1.2.2 Tubes**

Tube sources such as backward-wave-oscillators (carcinotrons) offer mW levels of power at up to 1.2 THz, but require magnetic fields and high voltages to operate [18]. Tubes also suffer a strong roll-off of power with increasing frequency due to physical scaling and metallic losses, and so are unlikely candidates for operation at higher terahertz frequencies. Perhaps more importantly they can now only be obtained from Russia, and suffer from a relatively short operating lifetime [7].

### **1.2.3 Photomixing and downconversion**

Downconversion from optical sources is another method for obtaining terahertz radiation. The technique of photomixing operates by illuminating a fast photoconductive material with two optical lasers detuned by the desired terahertz frequency. The intensity beating of the lasers modulates the conductivity of the photoconductor, and generates terahertz current flow in a dc-biased antenna. This technique produces a narrowband, cw, frequency agile source that operates at room temperature. However, optical conversion efficiencies are low and power drops with increasing frequency due both to the antenna impedance and limited photoconductive response speed [16]. Power levels available are on the order of 1–10  $\mu$ W below 1 THz, and sub-microwatt

above 1 THz.

Difference frequency generation has been used in nonlinear optical materials such as LiNbO<sub>3</sub>, KTP, DAST, GaSe, and ZnGeP<sub>2</sub> to generate pulsed terahertz radiation. The most successful material has been GaSe, where radiation tunable from 0.18–5.27 THz has been obtained in 5 ns pulses with a peak power of 70 W (2.5  $\mu$ W average power) at 1.53 THz with a conversion efficiency of 3.3% [19]. The efficiency improves at higher frequencies because of the Manley-Rowe factor.

While not suitable for use as a local oscillator, it is important to mention broadband terahertz downconversion as a source. A biased photoconductive switch or nonlinear crystal is illuminated with femtosecond optical pulses, which generates an output with the bandwidth determined by the pulse length [20, 21]. The result is a few-cycle coherent terahertz pulse, which is coupled out of the photoconductor by an antenna, as with photomixing described above. This process is effectively difference frequency mixing of the incident pulse with itself. While the average power is extremely low (nW– $\mu$ W), this technique for generation has become technologically important since it can be coherently detected at room temperature, and is the source for most terahertz imaging, as described above.

#### 1.2.4 Terahertz lasers

Along with Schottky multiplier chains, optically pumped gas lasers have been the other dominant source of far-infrared radiation. Low-pressure molecular gasses are pumped by a CO<sub>2</sub> gas laser to produce lasing between rotational levels of excited vibrational states. Although discrete lines can be obtained between 0.1 and 8 THz [22], most strongly pumped lines are below 3 THz. For these strongly pumped lines continuous-wave power of 1–20 mW is typical [7]. Such lasers are readily commercially available, and one is being used as a 2.5 THz local oscillator source in the EOS satellite [23]. However, the selection of laser frequencies is limited by available gasses, and tunability is extremely limited. Additionally, such gas lasers are expensive, bulky, and power hungry, which makes them less than ideal for spaceflight applications.

Semiconductor lasers in the terahertz have traditionally consisted only of the hot-

hole *p*-type Ge laser. In this device, lasing action results from a hole population inversion that is established between the light and heavy hole bands due to a “streaming motion” that takes place in crossed electric and magnetic fields [24]. Several watts of peak power have been obtained in broadband lasing (linewidths of 10–20 cm<sup>−1</sup>) that can be tuned from 1–4 THz. However the need for a magnetic field, high voltage, and cryogenic operation ( $T < 20$  K) limits the utility of this source. Traditionally, due to high power consumption and low efficiency the maximum duty cycle was limited to less than  $10^{-4}$ . However recent improvements have taken place, and duty cycles of up to 5% have been demonstrated [25]. Its broadband gain has proven useful for mode-locked operation though, where pulses with a 100-ps width have been obtained [26].

A recently developed semiconductor laser is the strained *p*-Ge resonant state laser. Stimulated emission was observed in 1992 by Altukhov *et al.* [27] and cw lasing that is tunable with pressure from 2.5 to 10 THz was demonstrated by Gousev *et al.* in 1999 [28]. Power levels of tens of microwatts were observed and operation takes place at liquid helium temperatures. In this device, application of strain lifts the degeneracy of the light-hole and heavy-hole band such that the  $1s$  impurity state of the heavy-hole band is brought into resonance with the light-hole band. This leads to a population inversion between the heavy hole  $1s$  state and the light-hole impurity states, which are depopulated by electric field ionization. No magnetic field is necessary. A similar laser was demonstrated in 2000 in SiGe/Si quantum wells, where the mechanism of lasing is the same but the strain is provided instead by the epitaxial mismatch [29, 30]. This is a promising method which eliminates the need for externally applied strain, but the power level is still quite low and high voltage (300–1500 V) is required.

The most recent development has been the extension of quantum cascade laser operation from the mid-infrared to the terahertz. The development of quantum cascade lasers that operate below the semiconductor *Reststrahlen* band of the semiconductor is the topic of this thesis, and the background on this subject is given in more detail below.

## 1.3 Intersubband transitions and quantum cascade lasers

In a traditional bipolar semiconductor laser, the emitted wavelength is determined by the material bandgap, although some flexibility can be gained via use of quantum wells or strained layers. The extension of bipolar laser technology to the far-infrared is impractical, as materials with bandgaps less than 40 meV would be required. One solution, that is the focus of this thesis, is for the radiative transition to take place entirely within the conduction band between quantized states in heterostructure quantum wells. When thin ( $\leq$  hundreds of Angstroms) semiconductor layers of differing composition are sequentially grown, discontinuities are introduced in the band edges, which leads to the quantum confinement of carriers in the growth direction. As a result, the band breaks up into “subbands,” where the energy is quantized in the growth direction, and parabolic free carrier dispersion relations are retained in the in-plane directions. The photon energy that results from an intersubband transition can be chosen by tailoring the thicknesses of the coupled wells and barriers, which makes such structures ideal for the generation of long-wavelength radiation.

The use of intersubband transitions for radiation amplification was first proposed in 1971 by Kazarinov and Suris [31] in a superlattice structure. A superlattice, first described by Esaki and Tsu in 1970 [32], is a periodic repetition of two material layers of differing composition, i. e. a repeated quantum well and barrier. The enabling technology for creating superlattice and quantum well semiconductor structures was and is molecular beam epitaxy (MBE), whereby layers as thin as several monolayers can be grown with atomic precision. Milestones in intersubband and superlattice experiments included the first observation of intersubband absorption in a quantum well in 1985 [33] and the first observation of sequential resonant tunneling in a superlattice in 1986 [34]. Helm *et al.* were the first to observe intersubband emission in a superlattice, initially pumped by thermal excitation [35] and then by resonant tunneling [36]. In fact, this first intersubband emission was observed in the terahertz frequency range (2.2 THz), perhaps belying the difficulty of establishing a far-infrared intersub-

band laser. As a result, it was initially thought that development of intersubband lasers where the radiative energy is less than the optical phonon energy would be easier, since non-radiative relaxation via longitudinal-optical (LO) phonon emission would be energetically forbidden. While this is technically true, it has proven much easier to operate intersubband lasers in the mid-infrared, where the large subband energy separations made establishing population inversion less difficult, and where free carrier absorption in the waveguide was much lower.

### 1.3.1 Mid-infrared quantum cascade lasers

In the 1980s and early 1990s a number of proposals for intersubband lasers were put forward, and are described in Refs. [37, 38]. The basic superlattice structure turned out to be unsuitable for lasing due to the difficulty of obtaining population inversion as well as its tendency to break up into high-field domains. The breakthrough came in 1994 at Bell Labs in the group of Federico Capasso, where Jerome Faist *et al.* developed the quantum cascade laser (QCL) [1]. The first device was grown in the InGaAs/AlInAs/InP material system and operated at a wavelength of  $4.2\text{ }\mu\text{m}$ . Lasing took place only in pulsed mode at cryogenic temperatures, and required a high threshold current density ( $\sim 14\text{ kA/cm}^2$ ). The schematic for this first laser is shown in Fig. 1-4(a). Operation is based on a radiative transition that is diagonal in real space. Depopulation is achieved by setting the subband separation between levels 2 and 1 to the LO-phonon resonance energy, which causes electrons to quickly relax from level 2 via electron-phonon scattering. LO-phonon scattering is also the dominant non-radiative relaxation mechanism from level 3, but is reduced due to the large in-plane momentum exchange necessary (Fig. 1-4(b)). The digitally graded alloy is essentially a doped quasi-classical multi-well superlattice region, which collects electrons from the lower states 2 and 1 and injects them into the excited radiative state in the next stage. At the same time the superlattice acts as a Bragg reflector at higher energies, which prevents electrons from tunneling out of the excited state 3 into the continuum. A key feature of QCLs is the ability to cascade  $N$  modules together, so that a single injected electron can emit many photons, which allows for

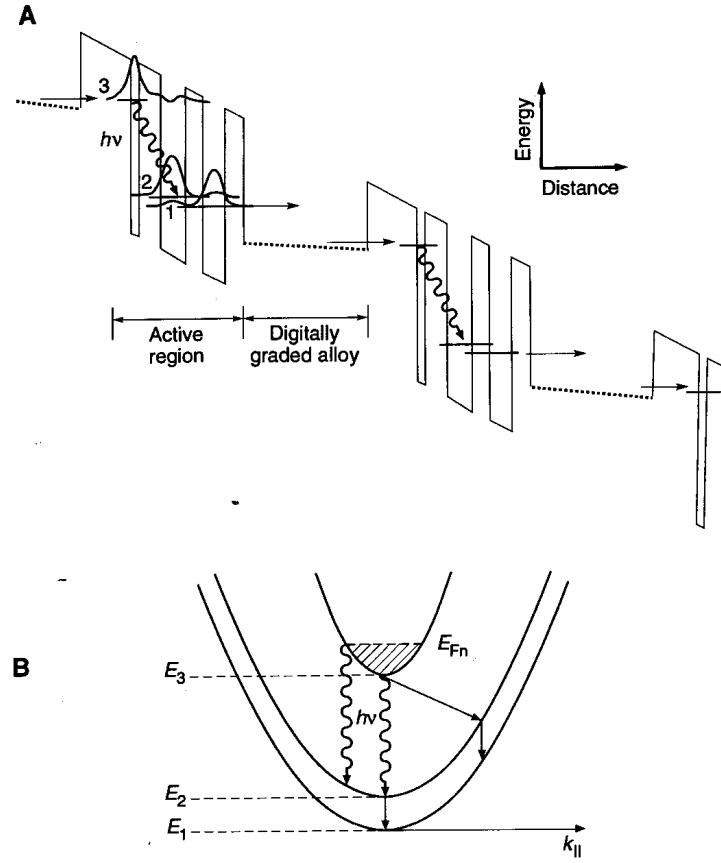


Figure 1-4: (a) Schematic of original quantum cascade laser, including conduction band profile and magnitude squared of wavefunctions. (b) In-plane momentum space ( $k_{\parallel}$ ) diagram of subbands and allowed relaxation paths via LO-phonon and photon emission. Figure is reprinted from Faist *et al.* [1].

differential quantum efficiencies greater than unity.

Quantum cascade lasers have since achieved dramatic performance improvements since their invention and are poised to become the dominant laser sources in the mid-infrared spectral range. An overview of QCL development history as well descriptions of state of the art designs and applications are given in the review papers by Capasso *et al.* [39] and Faist *et al.* [40]. Lasing has been obtained at wavelengths of 3.4–24  $\mu\text{m}$  in the mid-infrared [41, 42], and has recently been extended to far-infrared wavelengths of 60–130  $\mu\text{m}$  (see Sec. 1.3.2). Lasers at wavelengths as long as 16  $\mu\text{m}$  have been demonstrated in pulsed mode at room temperature, often with watt-level

peak powers. The highest reported operating temperature is 470 K in pulsed mode at  $\lambda = 5.5 \mu\text{m}$  [43]. Continuous-wave operation is available at most wavelengths at cryogenic temperatures and has recently been achieved at room temperature where up to 90 mW at a  $9 \mu\text{m}$  wavelength has been observed [44, 45]. Single-mode, temperature tunable distributed-feedback QCLs have been demonstrated from 4.5–16.5  $\mu\text{m}$ , and typically display 30 dB of side-mode suppression and 0.2–1.0 nm/K temperature tunability [46]. Such lasers have found use as sensitive chemical sensors in gas spectroscopy systems [47]. Because the flexible nature of intersubband transitions, quantum cascade lasers are not tied to one particular material system, although the conduction band offset between layers limits the shortest attainable wavelength. Most QCLs have been grown in the InGaAs/AlInAs/InP material system, although many devices have been demonstrated in the GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As material system [48]. However, performance in GaAs QCLs has lagged somewhat behind that of the InGaAs QCLs, partially because of the lower conduction band offsets in the GaAs/AlGaAs material system. A laser has recently been demonstrated at  $\lambda \approx 10 \mu\text{m}$  in the InAs/AlSb material system. Efforts are currently underway to develop QCLs based on valence-band intersubband transitions in SiGe/Si quantum wells [49, 50]. There is also strong interest in developing a quantum cascade laser that operates at telecom wavelengths. This is motivated by the fact that intersubband lasers are expected to have very small linewidth enhancement factors and do not suffer from relaxation oscillations, since the system is over-damped due to very fast non-radiative relaxation rates. Such a short wavelength ( $\sim 1.5 \mu\text{m}$ ) requires a material system with a large band offset, and preliminary work on intersubband transitions has begun in the AlGaN/GaN system [51, 52, 53]. Since its invention nine years ago, the quantum cascade laser has seen remarkable improvement which is by no means over. There remains much work to be done, including obtaining cw performance at room temperature over a greater wavelength range, raising power levels, improving single mode performance and tunability, and extending the operational wavelength range.

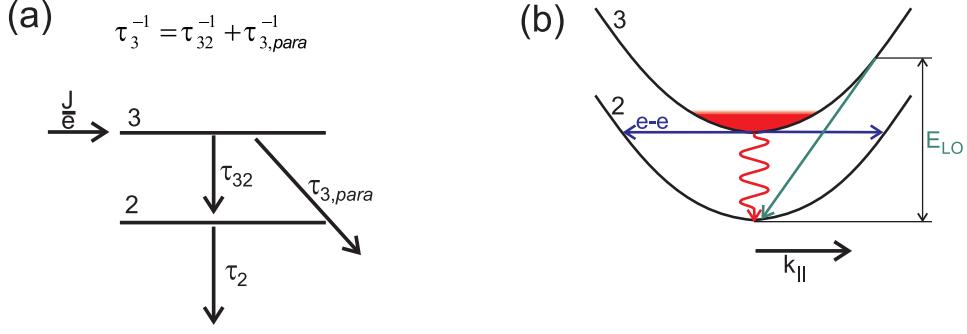


Figure 1-5: (a) Basic transport schematic for generic THz QCL. (b) In-plane momentum space ( $k_{\parallel}$ ) schematic with typical allowed relaxation paths for the  $3 \rightarrow 2$  transition.

### 1.3.2 Terahertz quantum cascade lasers

Despite the fact that the first observation of intersubband electroluminescence took place in the terahertz, the development of a terahertz intersubband laser proved to be much more difficult than the development of a mid-infrared QCL. There are two primary reasons for this: the difficulty of achieving a population inversion for such a small subband separation and the challenge of obtaining a low-loss waveguide for such long wavelengths.

The simplest model for transport in a QCL is shown in Fig. 1-5(a) (see App. C for a more sophisticated model). Electrons are injected into level 3 with a current density  $J$  where they either relax to level 2 with time  $\tau_{32}$ , or parasitically scatter out with time  $\tau_{3,para}$ , yielding a total lifetime of  $\tau_3$ . The depopulation rate of the lower level is  $\tau_2$ . Note that at terahertz frequencies the spontaneous emission lifetime is typically much longer (microseconds) compared to the non-radiative lifetimes (picoseconds), so radiative relaxation plays no role in transport below threshold. Solving for the rate equations we obtain the following for the 2D population inversion:

$$\Delta N = \frac{J}{e} \tau_3 \left( 1 - \frac{\tau_2}{\tau_{32}} \right). \quad (1.1)$$

As is intuitively obvious, to obtain a large population inversion, we desire  $\tau_{32} \gg \tau_2$  and  $\tau_3$  to be as long as possible. It turns out to be difficult to engineer a system

where the lower level is depopulated without also depopulating the upper state as well. In a terahertz QCL, the radiative state separation is less than the longitudinal-optical-phonon energy  $E_{LO}$ , and LO-phonon scattering is nominally forbidden at low temperatures (Fig. 1-5(b)). Only the high energy tail of the thermal electron distribution can emit LO-phonons. As a result, at low temperatures the lifetime is determined by a combination of electron-electron (e-e) scattering (Sec. 2.4.2), electron-impurity scattering, interface roughness scattering, and thermally activated LO-phonon scattering (Sec. 2.4.1). The e-e scattering rate roughly varies as  $N_3/E_{32}$ , where  $N_3$  is the subband population and  $E_{32}$  is the subband separation. While a long upper state lifetime is possible at low temperatures and populations, both scattering mechanisms are highly dependent on the electron distribution, which causes  $\tau_{32}$  to drop with increasing injection, and limits the maximum upper state population. Designing a system to selectively depopulate the lower state while still maintaining a large upper state population is a major challenge for terahertz QCL design. Additionally, although not illustrated in Fig. 1-5, for small energy separations  $E_{32}$ , it becomes difficult to selectively inject electrons into level 3 and not level 2.

The second challenge is designing an appropriate low-loss waveguide for mode confinement. For most mid-infrared QCLs, as in traditional semiconductor diode lasers, confinement is provided by dielectric cladding with a lower refractive index than the core. Because of the long wavelength at terahertz frequencies, such cladding would need to be prohibitively thick (tens of microns). Furthermore, loss due to free carrier absorption scales as  $\lambda^2$ , so any modal overlap with doped regions must be minimized, which is problematic for electrically pumped semiconductor lasers.

Efforts to obtain terahertz QCLs proceeded simultaneously with efforts to develop mid-IR QCLs, and continued after the demonstration of mid-IR QCLs in 1994. In addition to the work described above in Sec 1.3, early proposals for terahertz intersubband lasers proposed obtaining a population inversion through resonant tunneling selective injection and removal [54, 55, 56, 57]. For long wavelength devices, confinement between heavily doped “metallic” semiconductor contact layers was the preferred waveguiding solution [56]. Although such double well structures were tested

by Smet [37], no intersubband emission was observed. A detailed analysis of intersubband scattering mechanisms was however undertaken by Smet *et al.* [58, 37], which provided substantially more rigor to the existing feasibility arguments. Following the success of the mid-infrared QCL in 1994 various cascade designs that featured LO-phonon scattering for depopulation were considered by Bin Xu [59], and far-infrared intersubband emission was observed in such a structure based around a diagonal radiative transition in 1997 [60]. However, due to the interwell nature of the transition, the oscillator strength was small and emitted linewidth was large (see Sec. 6.2.1), and no evidence of gain was seen. Bin Xu also developed an improved terahertz waveguide, where the radiation was confined between two metal layers instead of a metal layer and an  $n^+$  semiconductor layer. Such a metal-metal waveguide has the advantage of low losses and a high confinement factor, and is extensively discussed in Sec. 4.5.2 and Sec. 5.3. Narrow linewidth terahertz spontaneous emission was observed via an intrawell transition by both Rochat *et al.* [61] in 1998 and Ulrich *et al.* [62] in 1999 where depopulation was dependent solely upon resonant tunneling in a narrow-band superlattice, rather than by LO-phonon depopulation. The struggle to obtain gain persisted, as narrow-linewidth vertical interwell spontaneous emission was observed by Williams *et al.* in 1999 [63], but the LO-phonon depopulation remained too slow (see Sec. 6.3.1).

The breakthrough finally came in October 2001 from Alessandro Tredicucci's group at Pisa, when Köhler *et al.* demonstrated the first terahertz quantum cascade laser, which operated at 4.4 THz ( $\lambda = 68 \mu\text{m}$ ). Lasing was observed up to 40 K, with a maximum duty cycle of 10%; 2.5 mW of peak power was observed in pulsed mode at 8 K. The winning structure, shown in Fig. 1-6, was composed of a new chirped superlattice active region placed in a new type of low-loss waveguide. The chirped superlattice design yielded a large inter-miniband dipole matrix element ( $z = 7.8 \text{ nm}$ ,  $f = 2.7$ ) for the radiative transition. Depopulation occurred via electron-electron scattering and resonant tunneling inside a wide (17 meV) miniband. The semi-insulating surface plasmon waveguide, first demonstrated by Ulrich *et al.* [62], replaced the thick heavily doped lower cladding layer with a thin (0.8- $\mu\text{m}$ )  $n^+$

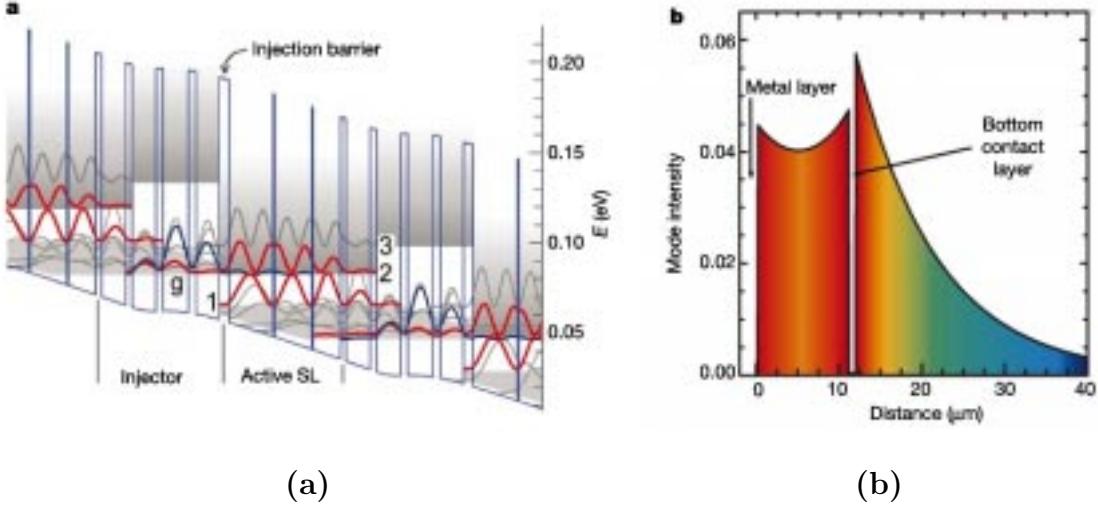


Figure 1-6: (a) Active region design for the chirped superlattice terahertz QCL demonstrated by Köhler *et al.*. (b) Mode pattern for semi-insulating surface plasmon waveguide. Figure is reprinted from Ref. [64].

GaAs contact layer grown on a semi-insulating (SI) GaAs substrate. The mode is composed of surface plasmons bound to the upper metallic contact and the lower contact layer; although the mode extends substantially into the substrate, the overlap with  $n$ -doped regions is small so the free carrier loss is minimized.

Soon afterward, in January 2002 Jerome Faist's group at Neuchâtel reported lasing in a similar structure at 4.5 THz ( $\lambda = 66 \mu\text{m}$ ) [65]. Most recently, in November 2002 a terahertz quantum cascade laser operating at 3.4 THz ( $\lambda = 87 \mu\text{m}$ ) was demonstrated by Williams (the author) and coworkers [66]. Approximately 14 mW of peak power was obtained in pulsed mode and operation was observed up to 87 K [67]. The active region of this laser operated on a different principle, one that used resonant phonon scattering to achieve depopulation. The development of this laser is the culmination of the work in this thesis, and is described extensively in Chap. 7. While less than two years have passed since the advent of the terahertz quantum cascade laser, much improvement has taken place. Continuous-wave operation has been demonstrated [68, 69] and lasing out to a wavelength of 130  $\mu\text{m}$  has been demonstrated [45]. Terahertz QCLs have been demonstrated using alternative low loss metal-metal waveguides to

obtain unity confinement factors, which should allow lower frequency operation (see Sec. 7.3)[70]. The field is rapidly progressing, and terahertz lasers promise to soon make an impact on scientific and hopefully even commercial applications.

### 1.3.3 Survey of terahertz QCL active regions

The design goal of the multiple quantum well active region is usually to provide the maximum gain, while minimizing any accompanying losses caused by free carrier or intersubband cross absorption. The peak bulk gain  $g_0$  (see Eq. (2.37)) provided by intersubband transitions can be described by three adjustable parameters, the oscillator strength  $f$ , the population inversion  $\Delta N$ , and the transition linewidth  $\Delta\nu$ , so that

$$g \propto \frac{\Delta N f}{\Delta\nu}. \quad (1.2)$$

The challenge for obtaining terahertz QCLs is the difficulty of obtaining a population inversion, i.e. selectively depopulating the lower state, without unduly sacrificing oscillator strength. The oscillator strength is proportional to the square of the dipole matrix element (Eq. (2.25)), which is essentially a wavefunction overlap integral. Because of this, an increase in  $f$  tends to be accompanied by an increase in non-radiative scattering, which reduces  $\Delta N$ . While the transition linewidth  $\Delta\nu$  cannot be directly engineered, it can be minimized by reducing homogeneous broadening mechanisms such as impurity and interface roughness scattering. The choice of a design may also affect the importance of inhomogeneous broadening due to module non-uniformity and dependence of the transition energy on applied electric field.

At this time, three active region types have successfully been used in terahertz QCLs (Fig. 1-7): chirped superlattice, bound-to-continuum, and resonant phonon. The first successful devices were based on chirped superlattice designs [64, 65], which have been previously used for mid-infrared QCLs [71]. Superlattice QCLs are based upon inter-miniband radiative transitions at the miniband edge. To illustrate this, the schematic energy diagram of a superlattice and its dispersion relation is shown in Fig. 1-8. Essentially, in a superlattice the single well states couple together to form

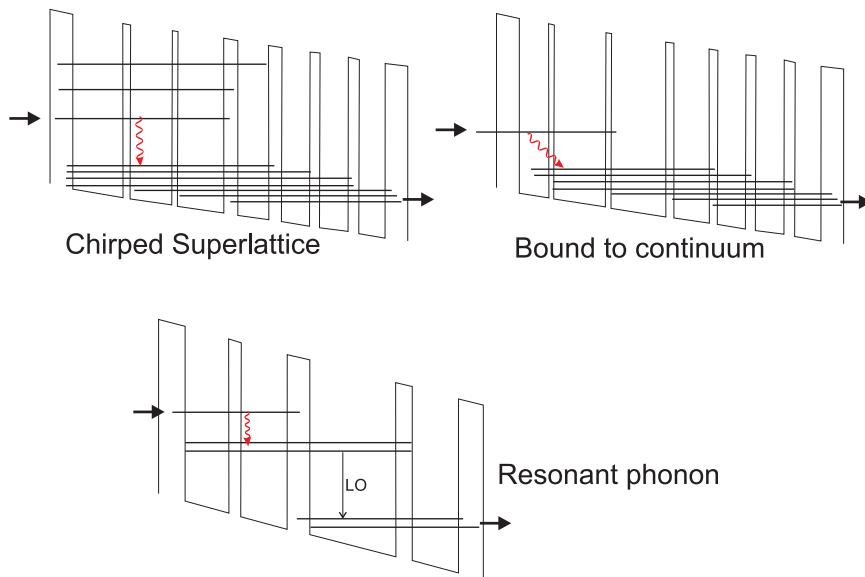


Figure 1-7: Schematic diagrams of successful terahertz QCL active regions.

“minibands” separated by a “minigap.” To first order, the minigap is determined by the well width, and the miniband width is determined by the barrier thickness. In reality, QCL structures do not use true infinite superlattices, but are in fact composed of several (4–8) wells. Therefore, one can model the active regions either as true superlattices, or as several tightly coupled individual subbands. Superlattice QCL designs tend to have very large oscillator strengths due to the large spatial extent of the wavefunctions.

In a superlattice QCL, a population inversion is established on the principle that intra-miniband scattering occurs much faster than inter-miniband scattering, so that carriers will thermalize and collect in the bottom of each miniband. In other words, the lower radiative subband is more tightly coupled to the levels close in energy, so out-scattering (depopulation) processes see a larger phase space than do in-scattering processes. In terahertz QCLs, the miniband width is typically less than an LO-phonon energy, so intersubband scattering inside the miniband takes place largely through electron-electron scattering, electron-impurity scattering, and interface roughness scattering, while optical-phonon emission cools the distribution as a whole. Transport through the superlattice is also affected by the miniband width

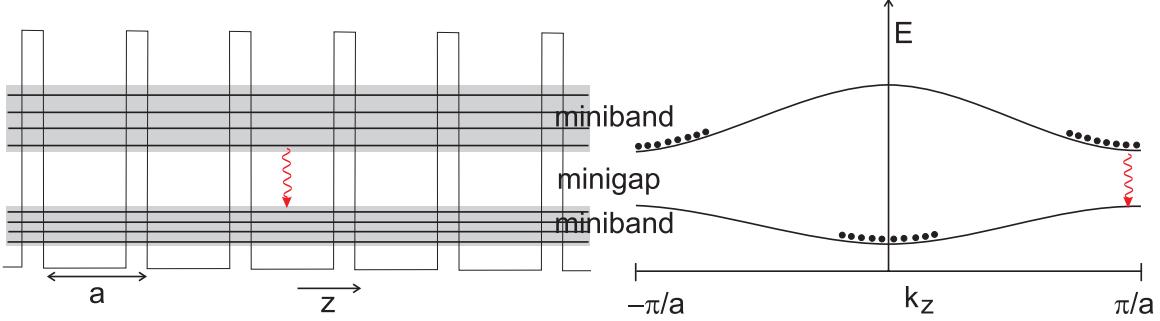


Figure 1-8: Schematic diagram of a superlattice and its dispersion relation.

which affects the miniband curvature and hence the effective mass of a wavepacket. This is essentially equivalent to controlling the rate of coherent resonant tunneling by adjusting the barrier thickness to change the subband energy separations. In practice, calculations of intra-miniband transport and scattering are extremely difficult to perform with much precision, and most design is qualitative in nature. However, the success of designs that use miniband final states indicate that the depopulation process can be quite efficient.

The bound-to-continuum design was recently demonstrated by Scalari *et al.* [72] and has displayed improved performance over the original chirped superlattice design. Notably, pulsed operation up to 100 K at 3.5 THz has been observed, with CW powers up to 50 mW at 10 K. Like the chirped superlattice, this design also uses a miniband for depopulation of the lower radiative state. However, instead of an inter-miniband transition, the upper radiative state is a somewhat localized state that sits in the middle of the minigap. This leads to a more diagonal radiative transition, and a slightly reduced oscillator strength. However, this is accompanied by reduced nonradiative scattering of the upper state into the miniband, especially at higher temperatures where the upper state lifetime is decreased by thermally activated LO-phonon scattering. The diagonal nature also helps to reduce coupling from the injector into the lower radiative state, which improves injection efficiency.

Contemporaneously with the development of the bound-to-continuum design, the resonant phonon design for terahertz QCLs was recently demonstrated by the author and coworkers [66]. This design is fundamentally different than the two previous ter-

ahertz QCL designs, as it does not rely on a superlattice for depopulation, but rather uses combination of resonant tunneling and direct electron-LO-phonon scattering. In this way, this design has more in common with the conventional mid-infrared QCLs. The lower radiative state is anticrossed with the excited state in the wide quantum well, which enables fast selective depopulation via LO-phonon scattering. Because the upper radiative state has a reduced overlap with the injector states, parasitic nonradiative relaxation is minimized. The use of direct LO-phonon depopulation has two potential advantages. First, depopulation via LO-phonon emission is a fast, robust mechanism which doesn't depend very much on temperature or the electron distribution. Second, the large energy separation between the injector states and the lower radiative state ( $> E_{LO} = 36$  meV) provides intrinsic protection against thermal backfilling of the lower radiative state. Both properties are important in allowing higher temperature operation of lasers at longer wavelengths. The first design of this type demonstrated lasing at 3.5 THz in pulsed mode up to 87 K (Sec. 7.2).

## 1.4 Thesis overview

This thesis describes the design theory, fabrication techniques, and finally experimental characterization of operational terahertz quantum cascade lasers. In Chapter 2, the theoretical models used to design the multiple quantum well active region are discussed, including calculation of radiative and non-radiative intersubband transition rates, and modeling of resonant tunneling transport. In Chapter 3, the effect of the heterostructure on the phonon spectra and how it changes QCL design rules is considered. In Chapter 4, optical loss mechanisms and waveguide designs at terahertz frequencies are considered. Chapter 5 reviews experimental methods and fabrication techniques, including fabrication of the metal-metal waveguide. Chapter 6 is an extensive experimental survey of many of the preliminary designs that failed to lase. Finally, Chapter 7 describes the successful design of terahertz quantum cascade lasers. The five successful structures in which we have obtained lasing are reviewed, as well as prospects for improvements.



# Chapter 2

## Intersubband laser theory and modeling

### 2.1 Introduction

In this chapter I will review the theory and modeling necessary for the design of terahertz quantum cascade lasers. While most of the concepts are quite general, my focus will be on the GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As material system. As a result, only intersubband transitions in the conduction band are considered, thus avoiding all of the complexity of the valence band. This material system is lattice matched, therefore the effect of strain is neglected. Also, as mentioned below, the effects of nonparabolicity are quite small in this system for the small energies involved in terahertz QCLs, and are generally neglected.

### 2.2 Electronic states in heterostructures

Intersubband lasers are made possible by the growth of multiple heterostructures, which are atomically abrupt layers composed of materials with different bandgaps. At the heterojunction between two such materials, there is almost always a sharp discontinuity in the band energies. When the layers are made thin enough, i.e. on the order of a DeBroglie wavelength, electron motion is restricted in the growth

direction  $\hat{\mathbf{z}}$  and its energy is quantized.

In this treatment, the effective mass theorem in the envelope function approximation is used to obtain the electronic wavefunctions, and the  $\Gamma$ -point effective mass is used to describe the conduction band curvature ( $m^* = 0.067m_0$  in GaAs). In general, the wavefunction for the electron is given by

$$\Psi(\mathbf{r}) = F(\mathbf{r})U_{n,0}(\mathbf{r}), \quad (2.1)$$

where  $U_{n,0}(\mathbf{r})$  is the Bloch state wavefunction at the band minimum and  $F(\mathbf{r})$  is the envelope function and satisfies the effective mass equation. The varying material composition is represented by a spatially varying effective mass  $m^*(z)$  and the potential  $E_c(z)$ , which represents the conduction band edge profile, including any externally applied field and local variations due to space charge. The effective mass equation then becomes

$$\left[ -\frac{\hbar^2 \nabla_{||}^2}{2m^*(z)} - \frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{m^*(z)} \frac{\partial}{\partial z} + E_c(z) \right] F(\mathbf{r}) = EF(\mathbf{r}), \quad (2.2)$$

where  $\nabla_{||}$  is the in-plane differential operator [73]. The solution for the envelope function is given by

$$F(\mathbf{r}) = \frac{1}{\sqrt{S_{||}}} e^{i\mathbf{k}_{||} \cdot \mathbf{r}_{||}} \psi_n(\mathbf{k}_{||}, z), \quad (2.3)$$

where  $\psi_n(\mathbf{k}_{||}, z)$  satisfies

$$\left[ -\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{m^*(z)} \frac{d}{dz} + E_c(z) + \frac{\hbar^2 \mathbf{k}_{||}^2}{2m^*(z)} \right] \psi_n(\mathbf{k}_{||}, z) = E_n(\mathbf{k}_{||}) \psi_n(\mathbf{k}_{||}, z), \quad (2.4)$$

and  $\mathbf{k}_{||}$  is the in-plane wavevector,  $n$  is the subband index, and  $S_{||}$  is the normalization area. The spatially dependent effective mass introduces a coupling between the in-plane and  $z$  directions. This coupling is usually neglected and (2.4) becomes the one-dimensional Schrödinger equation

$$\left[ -\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{m^*(z)} \frac{d}{dz} + E_c(z) \right] \psi_n(z) = E_n \psi_n(z), \quad (2.5)$$

where the total energy is given by

$$E_n(\mathbf{k}_\parallel) = E_n + \frac{\hbar^2 k_\parallel^2}{2m^*}, \quad (2.6)$$

where  $m^*$  is taken to be the well material effective mass. This is the sum of the  $z$ -direction energy and the in-plane free particle kinetic energy. We can justify the use of Eq. (2.5) where the in-plane/ $z$ -direction coupling is neglected by considering that inclusion of this coupling in Eq. (2.5) would effectively change the barrier height  $E_c(z)$  by the energy

$$\frac{\hbar^2 k_\parallel^2}{2m^*} \left( \frac{m^*}{m^*(z)} - 1 \right), \quad (2.7)$$

as can be derived from Eq. (2.4), assuming the form of Eq. (refeq:schren) for  $E_n(\mathbf{k}_\parallel)$ . As long as the in-plane kinetic energy is modest compared to the barrier height, and the barrier and well effective masses do not differ too much, it is reasonable to neglect this coupling. While it cannot be rigorously derived, the form of the kinetic energy operator is chosen to preserve continuity of the envelope functions and current density across material interfaces. In this work, Eq. (2.5) is solved using a finite-difference method, where the structure of interest is divided into nodes with constant material parameters. Although the states calculated are typically assumed to be stationary states, this is not strictly true. Under the application of an electric field, the states become metastable, i.e. they may escape into the continuum. However, usually only a limited spatial extent of the quantum-well structure is simulated, and the states are assumed to be bound. Escape to continuum is considered via an escape time, and is usually much longer than intersubband scattering times.

Population of the electronic states by electrons introduces space charge which consequently affects the conduction band profile  $E_c(z)$ . For this reason, it is sometimes necessary to solve the Poisson equation

$$\frac{d}{dz} \left[ \epsilon(z) \frac{d}{dz} \Phi(z) \right] = -\rho(z), \quad (2.8)$$

where  $\Phi(z)$  is the electrostatic potential,  $\epsilon(z)$  is the spatially varying permittivity,

and  $\rho(z)$  is the charge density. This gives  $E_c(z) = E_{c,0}(z) - e\Phi(z)$ , where  $E_{c,0}(z)$  is the intrinsic conduction band profile. The Poisson and Schrödinger equations are iteratively solved to obtain a self-consistent solution. However, for many designs discussed in this thesis, the electron density is sufficiently low that the  $E_c(z)$  is not significantly perturbed, and no self-consistent solution is necessary.

Band non-parabolicity is neglected, since for terahertz lasers most energies of interest are close to the band edge with accordingly small  $k$  values. To test this assumption, an empirical two-band model for non-parabolicity was implemented as described in Ref. [74, 75]. In this treatment, an energy dependent effective mass was used in conjunction with the Schrödinger equation to obtain the subband energies. For typical designs discussed, the subband energy shifts due to the inclusion of non-parabolicity are very small, typically less than 1 meV.

The model described above is among the simplest used to describe electronic states in semiconductors. One might expect the effective mass model to break down for wells or barriers that are only several monolayers thick, due to the sharp discontinuity in potentials. In such cases, using more sophisticated tight-binding approaches might be justified. However, this turns out not to be the case. Burt theoretically compared the effective mass envelope function solution with the exact atomic wavefunctions in quantum well systems [76]. He found that the effective mass solution provides an excellent approximation, even for very abrupt structures, provided that the envelope functions themselves are slowly varying. This is consistent with the slowly varying envelope function approximation which restricts the range of wavevectors about some value. Additionally, the use of effective mass theory is well justified by the excellent experimental confirmation that has been provided by an extremely wide variety of structures, including QCL devices, some of which have layers as thin as two monolayers.

## 2.3 Intersubband radiative transitions and gain

The optical gain for a QCL is provided by stimulated emission of photons from electron transitions between subbands in the MQW structure. Neglecting non-parabolicity, the initial and final subbands have the same curvature, and hence the joint density of states for an intersubband transition is a Dirac delta function at the subband separation  $\hbar\omega = E_i - E_f$ . The following section calculates the spontaneous and stimulated transition rates used to find the bulk gain of our structure. The treatment is standard, and is based on that in Yariv [77] and Smet [37].

Transitions between conduction subbands in quantum wells occur at a rate given by Fermi's golden rule:

$$W_{i \rightarrow f} = \frac{2\pi}{\hbar} \left| \langle f, n_{\mathbf{q},\sigma} | H' | i, m_{\mathbf{q},\sigma} \rangle \right|^2 \delta(E_f(\mathbf{k}_f) - E_i(\mathbf{k}_i) \pm \hbar\omega_{\mathbf{q}}), \quad (2.9)$$

where

$$H' = -\frac{e}{m^*} \mathbf{A} \cdot \mathbf{p} \quad (2.10)$$

is the interaction Hamiltonian. The initial and final states  $|i, n_{\mathbf{q},\sigma}\rangle$  and  $|f, m_{\mathbf{q},\sigma}\rangle$  are product states of the electron conduction band envelope function eigenstates  $i, f$  and the photon eigenstates with  $n$  and  $m$  photons in each mode given by the photon wavevector  $\mathbf{q}$  at frequency  $\omega_{\mathbf{q}}$ , and the polarization state described by  $\sigma = 1, 2$ . We take  $m^*$  to be the effective mass of the well material since the the electron spends the majority of its time there, and we ignore the barrier material. We also neglect any effects of non-parabolicity of the band structure, as this will modify the oscillator strengths and selection rules. More complete expressions that include the spatial dependence of the effective mass and non-parabolicity are discussed in Ref. [78, 79].

The Lorentz-gauge vector potential  $\mathbf{A}$  for a harmonic interaction can be written in term of the raising and lowering operators  $a_{\mathbf{q},\sigma}^\dagger$  and  $a_{\mathbf{q},\sigma}$ :

$$\mathbf{A} = \sqrt{\frac{\hbar}{2\epsilon\omega_{\mathbf{q}}V}} \hat{\mathbf{e}}_{\mathbf{q},\sigma} \left[ a_{\mathbf{q},\sigma} e^{i\mathbf{q} \cdot \mathbf{r}} + a_{\mathbf{q},\sigma}^\dagger e^{-i\mathbf{q} \cdot \mathbf{r}} \right]. \quad (2.11)$$

In this expression,  $\epsilon$  is the permittivity,  $V$  is the volume of the cavity, and  $\hat{\mathbf{e}}_{\mathbf{q},\sigma}$  is the polarization vector. Application of the raising and lower operators yield this expression for the matrix element:

$$|\langle H'_{i \rightarrow f} \rangle|^2/C = m_{\mathbf{q},\sigma} \delta_{m_{\mathbf{q},\sigma}-1,n_{\mathbf{q},\sigma}} \left| \left\langle f \left| e^{i\mathbf{q} \cdot \mathbf{r}} \hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{p} \right| i \right\rangle \right|^2 + (m_{\mathbf{q},\sigma} + 1) \delta_{m_{\mathbf{q},\sigma}+1,n_{\mathbf{q},\sigma}} \left| \left\langle f \left| e^{-i\mathbf{q} \cdot \mathbf{r}} \hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{p} \right| i \right\rangle \right|^2; \quad (2.12)$$

$$C = \frac{e^2 \hbar}{2m^{*2} \epsilon \omega_{\mathbf{q}} V}. \quad (2.13)$$

The first term corresponds to absorption of a photon, and the second term emission. The terms proportional to the number of photons in the mode  $m_{\mathbf{q},\sigma}$  correspond to stimulated processes, whereas the field independent term corresponds to spontaneous emission noise. Multi-photon processes are not included.

The initial and final electron states can be projected onto real space and written as the product of a transverse and longitudinal envelope function as in (2.3), where the in-plane coupling has been neglected ( $\psi_n(\mathbf{k}_{||}, z) = \psi_n(z)$ ).

$$\langle \mathbf{r} | i \rangle = \frac{1}{\sqrt{S_{||}}} e^{i\mathbf{k}_{||,i} \cdot \mathbf{r}_{||}} \psi_i(z) \quad (2.14)$$

$$\langle \mathbf{r} | f \rangle = \frac{1}{\sqrt{S_{||}}} e^{i\mathbf{k}_{||,f} \cdot \mathbf{r}_{||}} \psi_f(z), \quad (2.15)$$

where  $S_{||}$  is the normalization area.

Since we are considering terahertz radiation ( $\lambda > 10 \mu\text{m}$ ), and  $r$  is on the order of the quantum well width (100 Å), we can adopt the electric dipole convention  $e^{i\mathbf{q} \cdot \mathbf{r}} \cong 1$ . We obtain

$$\langle f \left| e^{\pm i\mathbf{q} \cdot \mathbf{r}} \hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{p} \right| i \rangle \cong \langle f \left| \hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{p} \right| i \rangle, \quad (2.16)$$

$$\begin{aligned} &= [\hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{k}_{||,i}] \frac{\hbar}{S_{xy}} \int d\mathbf{r}_{||} e^{i(\mathbf{k}_{||,i} - \mathbf{k}_{||,f}) \cdot \mathbf{r}_{||}} \int dz \psi_f^*(\mathbf{k}_{||,f}, z) \psi_i(\mathbf{k}_{||,i}, z) \\ &\quad - \hat{\mathbf{z}} \frac{i\hbar}{S_{xy}} \int d\mathbf{r}_{||} e^{i(\mathbf{k}_{||,i} - \mathbf{k}_{||,f}) \cdot \mathbf{r}_{||}} \int dz \psi_f^*(\mathbf{k}_{||,f}, z) \frac{\partial \psi_i(\mathbf{k}_{||,i}, z)}{\partial z} \end{aligned} \quad (2.17)$$

$$= [\hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \hat{\mathbf{z}}] \delta_{\mathbf{k}_{||,f}, \mathbf{k}_{||,i}} \langle \psi_f | p_z | \psi_i \rangle. \quad (2.18)$$

The first term of (2.17) vanishes due to the orthonormality of the envelope functions. This is the well known intersubband selection rule: only transitions with the **E** field polarized along the growth axis  $\hat{\mathbf{z}}$  are permitted. The delta function in (2.18) ensures conservation of in-plane momentum over a transition.

Equation (2.18) is more commonly written in terms of the dipole interaction  $e\mathbf{r}\cdot\mathbf{E}$ . We can rewrite the matrix elements by making use of the commutation relation for the unperturbed Hamiltonian  $H_0$

$$\frac{i}{\hbar}[H_0, z] = \frac{p_z}{m^*}. \quad (2.19)$$

We then obtain

$$\langle f | p_z | i \rangle = \frac{im^*}{\hbar}(E_f - E_i) \langle \psi_f | z | \psi_i \rangle. \quad (2.20)$$

where the integral  $\langle \psi_f | z | \psi_i \rangle$  is the dipole matrix element  $z_{i\rightarrow f}$ . Inserting this matrix element into (2.13) and (2.9) give the following spontaneous and stimulated transition rates:

$$W_{i\rightarrow f/mode}^{(sp)} = \frac{\pi e^2 \omega_{\mathbf{q}}}{\epsilon V} |\hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \hat{\mathbf{z}}|^2 |z_{i\rightarrow f}|^2 \delta(E_f - E_i + \hbar\omega_{\mathbf{q}}), \quad (2.21)$$

$$W_{i\rightarrow f/mode}^{(st)} = \frac{\pi e^2 \omega_{\mathbf{q}}}{\epsilon V} |\hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \hat{\mathbf{z}}|^2 |z_{i\rightarrow f}|^2 \delta(E_f - E_i + \hbar\omega_{\mathbf{q}}) m_{\mathbf{q},\sigma}, \quad (2.22)$$

where the Kronecker delta functions conserving in-plane momentum and photon number have been dropped to simplify notation.

### 2.3.1 Spontaneous emission lifetime

To obtain the total spontaneous emission rate, we must sum (2.21) over all of the photon modes and polarizations in the cavity. The number of electromagnetic modes in a differential volume  $d^3\mathbf{q}$  in  $\mathbf{q}$  space is

$$\rho(\mathbf{q}) d^3\mathbf{q} = \frac{d^3\mathbf{q}}{8\pi^3/V} = \frac{q^2 dq \sin \theta d\theta d\phi V}{8\pi^3}, \quad (2.23)$$

provided the cavity dimensions are much larger than the wavelength.

If we choose the polarization direction such that  $\hat{\mathbf{e}}_{\mathbf{q},\sigma=1}$  lies in the plane defined by  $\hat{\mathbf{z}}$  and  $\mathbf{q}$ ,  $|\hat{\mathbf{e}}_{\mathbf{q},\sigma=2} \cdot \hat{\mathbf{z}}|^2 = 0$  and  $|\hat{\mathbf{e}}_{\mathbf{q},\sigma=1} \cdot \hat{\mathbf{z}}|^2 = \sin^2 \theta$ . One need only sum over one set of polarizations. We obtain

$$\begin{aligned} W_{i \rightarrow f}^{(sp)} &= \frac{e^2}{8\pi^2\epsilon\hbar} |z_{i \rightarrow f}|^2 \int_0^\infty \int_0^\pi \int_0^{2\pi} \left(\frac{n\omega_{\mathbf{q}}}{c}\right)^3 \sin^3 \theta \delta(E_f - E_i + \hbar\omega_{\mathbf{q}}) d\hbar\omega d\theta d\phi \\ &= \frac{e^2 n \omega_0^3}{3\pi\epsilon_0\hbar c^3} |z_{i \rightarrow f}|^2 = \frac{e^2 n \omega_0^2}{6\pi m^* \epsilon_0 c^3} f_{i \rightarrow f}, \end{aligned} \quad (2.24)$$

where  $n$  is the index of refraction at the frequency  $\omega_0 = (E_i - E_f)/\hbar$ , and  $c$  is the speed of light in a vacuum. As shown above, the spontaneous emission rate can be expressed in terms of the dipole matrix element, or the scaled oscillator strength

$$f_{i \rightarrow f} = \frac{m^*}{m_0} f_{i \rightarrow f, unscaled} = \frac{2m^*(E_f - E_i)|z_{i \rightarrow f}|^2}{\hbar^2}. \quad (2.25)$$

The scaled oscillator strength is  $f \equiv 1$  for a classical Hertzian dipole. We find it convenient to use the scaled oscillator strength  $f_{i \rightarrow f}$  since it is proportional to the gain, and is a measure of the strength of a transition. Furthermore, we can use the commutation rules  $[z, p_z] = i\hbar$  and  $im^*[z, H] = \hbar p_z$  to obtain the Thomas-Reiche-Kuhn sum rule for the scaled oscillator strength:

$$\sum_{f \neq i} f_{i \rightarrow f} = 1. \quad (2.26)$$

Note that these commutation relations strictly hold only in the case of a single effective mass  $m^*$ , rather than a spatially dependent one. Modifications to this sum rule due to a spatially dependent effective mass and non-parabolicity are discussed in Sirtori *et al.* [78]. However, these modifications are not significant in the terahertz range due to the small energies involved as well as the generally low barrier heights (small  $x$  for  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  barriers).

A typical terahertz spontaneous emission time is longer than a microsecond; for example  $\tau_{i \rightarrow f}^{(sp)} \approx 12 \mu\text{s}$  for  $f_{i \rightarrow f} = 1$  at 3.0 THz in GaAs, and  $\tau_{i \rightarrow f}^{(sp)}$  scales as  $\omega^{-2}$ . Therefore radiative transitions are expected to play no role in subthreshold trans-

port, since nonradiative lifetimes are on the order of picoseconds. Eq. (2.24) gives the spontaneous emission rate for an intersubband transition taking place in a three dimensional cavity. Transitions occurring in a waveguide microcavity will see a different wavevector mode distribution and will consequently have a higher transition rate [80]. For example, in a 2D microcavity formed between two conducting layers separated by distance  $t$ , the spontaneous emission rate is enhanced by the factor  $3\lambda/4t$ , where  $\lambda$  is the wavelength inside the semiconductor [37]. Our waveguides are typically about 10- $\mu\text{m}$  thick, so some enhancement might be expected, however in practice it is difficult to distinguish any microcavity power enhancement from variations in waveguide losses.

### 2.3.2 Stimulated emission

The stimulated emission rate can be found by using the single mode transition rate (2.22) to examine the interaction of the dipole with a monochromatic incident wave at frequency  $\nu$ , mode  $\ell$ . Since each energy level has a certain finite width due to its finite lifetime, it is useful to replace the Dirac delta-function with a normalized lineshape function  $\gamma(\nu)$  for the transition. A homogeneously broadened transition will have the Lorentzian lineshape

$$\gamma(\nu) = \frac{(\Delta\nu/2\pi)}{(\nu - \nu_0)^2 + (\Delta\nu/2)^2}, \quad (2.27)$$

where

$$\Delta\nu = \frac{1}{\pi T_2} = \frac{1}{\pi} \left( \frac{1}{2\tau_i} + \frac{1}{2\tau_f} + \frac{1}{T_2^*} \right) \quad (2.28)$$

is the full-width half maximum linewidth of the transition centered about  $\nu_0$ .  $T_2$  is the total phase breaking time,  $\tau_i$  and  $\tau_f$  are the initial and final state lifetimes (also known as the  $T_1$  time), and  $T_2^*$  is the pure dephasing time. Note that lifetime broadening contributes to the homogeneous linewidth half as much as does pure dephasing (see Appendix A).

The stimulated emission rate is given by

$$W_{i \rightarrow f}^{(st)} = \frac{\eta \lambda^2 I_\nu}{8\pi h\nu n^2 \tau_{i \rightarrow f}^{(sp)}} \gamma(\nu), \quad (2.29)$$

where we have substituted in the incident wave intensity in the mode

$$I_{\nu, \ell} = \frac{cm_\ell h\nu_\ell}{nV}, \quad (2.30)$$

with  $m_\ell$  photons in the mode of frequency  $\nu_\ell$ . The interaction of the incident field with the atom-like dipoles is described by the parameter

$$\eta = \frac{3|z_{i \rightarrow f}|^2}{|\mathbf{r}_{i \rightarrow f}|^2}. \quad (2.31)$$

As long as the incident radiation is  $\hat{\mathbf{z}}$  polarized, as is the case for intersubband spontaneous noise input,  $\eta = 3$  for intersubband stimulated transitions. This is because all of the MQW “dipoles” are oriented in the same direction and not randomly oriented as in atomic media.

### 2.3.3 Intersubband gain

We can now use the expressions for spontaneous and stimulated emission derived above to consider the optical gain in our laser medium. Given a two level system with population densities  $N_1$  and  $N_2$  in subbands with energies  $E_1$  and  $E_2$ , in the presence of a field at frequency  $\nu_\ell$  in mode  $\ell$ , there will be  $N_2 W_{21}$  induced transitions from  $2 \rightarrow 1$ , and  $N_1 W_{12}$  induced transitions from  $1 \rightarrow 2$ . This leads to an induced power of

$$\frac{\text{Power}}{\text{Volume}} = [N_2 W_{21} - N_1 W_{12}]h\nu, \quad (2.32)$$

where  $W_{12}^{(st)} = W_{21}^{(st)}$  is proportional to the field intensity  $I_\nu$  as seen in (2.29). Neglecting non-parabolicity, both the initial and final subbands have the same density of states, thus we can neglect differences in level degeneracy. If  $N_2 > N_1$ , the induced power will be added to the inducing wave  $I_\nu$  so that it will grow as it propagates

according to

$$\frac{dI_\nu}{dz} = g(\nu)I_\nu. \quad (2.33)$$

We can insert the stimulated transition rate (2.29) into (2.32) to obtain the gain for intersubband transitions. The small signal bulk gain per unit length for a transition with a population inversion per unit volume  $\Delta N$  is

$$g(\nu) = \frac{\Delta Ne^2\pi\nu_0|z_{i\rightarrow f}|^2}{\hbar cn\epsilon_0}\gamma(\nu), \quad (2.34)$$

$$= \frac{\Delta Ne^2f_{i\rightarrow f}}{4m^*cn\epsilon_0}\gamma(\nu). \quad (2.35)$$

The three-dimensional population inversion density is written as  $\Delta N = \Delta N^{(2D)}/L_{mod}$ , where  $\Delta N^{(2D)}$  is the two-dimensional population inversion density, and  $L_{mod}$  is the length of a module. The peak gain at frequency  $\nu_0$  is therefore

$$g(\nu_0) = \frac{2\Delta Ne^2\nu_0|z_{i\rightarrow f}|^2}{\hbar cn\epsilon_0\Delta\nu}, \quad (2.36)$$

$$= \frac{e^2}{2\pi m^*cn\epsilon_0} \frac{\Delta N f_{i\rightarrow f}}{\Delta\nu}, \quad (\text{SI units}) \quad (2.37)$$

$$\simeq 70 \times \left( \frac{\Delta N}{10^{15} \text{ cm}^{-3}} \right) \left( \frac{f_{i\rightarrow f}}{\Delta\nu/\text{THz}} \right) \text{ cm}^{-1}. \quad (2.38)$$

In Eq. (2.38),  $\Delta N$  is in units of  $\text{cm}^{-3}$ , and  $\Delta\nu$  is in units of THz. Note that inside a waveguide structure, the  $n$  should represent the effect the effective index of the mode of interest.

## 2.4 Nonradiative inter- and intra-subband transitions

In a quantum cascade laser, the design of energy levels and wavefunctions allows scattering rates to be engineered to provide a population inversion. A proper understanding of inter- (and intra-) subband scattering is essential for the design and understanding of quantum cascade structures. The important intersubband scatter-

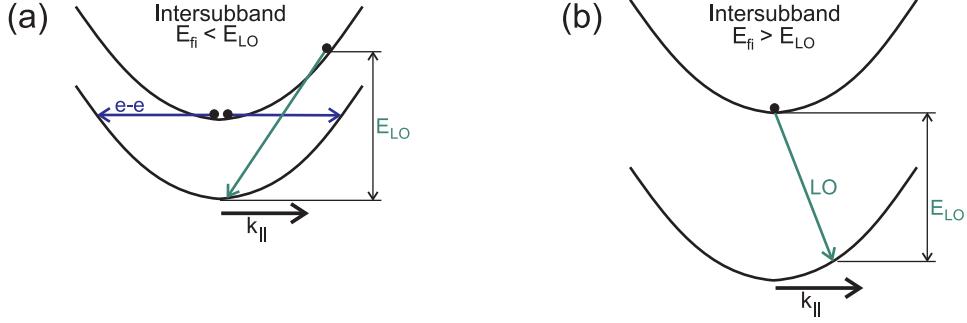


Figure 2-1: Important intersubband scattering mechanisms for the cases (a)  $E_{fi} < E_{LO}$  and (b)  $E_{fi} > E_{LO}$ .

ing channels are displayed in Fig. 2-1. It has been well established both theoretically and experimentally that electron-LO-phonon scattering is the dominant intersubband scattering mechanism for subband separations greater than the LO-phonon energy  $E_{LO}$  [81, 82]. For intrawell transitions, these scattering times tend to be less than 1 ps. The details of calculating LO-phonon scattering rates using bulk phonon modes are given in Sec. 2.4.1; a more sophisticated model that takes account of the heterostructure is presented in Chap. 3.

For intersubband transitions where the separation is less than  $E_{LO}$ , emission of LO-phonons is energetically forbidden at low temperatures. Nonradiative relaxation is therefore dominated by a combination of electron-electron (e-e) scattering, electron-impurity scattering, and LO-phonon scattering of the high energy tail of the subband electron distribution. However, even when devices are operated in liquid helium, one cannot assume that the electron gas temperature is also low. Examination of energy loss rates for cooling in single subband suggest that the electron temperature may be anywhere from 50–100 K higher than the lattice temperature [83]. Because emission of LO-phonons is the only efficient way to cool the electron gas, during operation the carrier ensemble will heat up until sufficient phonon emission occurs to reach steady state. For example, Monte Carlo simulations of the FL175C device (Sec. 7.2) performed by Hans Callebaut indicate electron temperatures of at least 100 K [84]. Calculating intersubband scattering rates for small subband separations is therefore difficult, as both depend sensitively on the electron distribution. In particular, accu-

rate calculation of e-e scattering is difficult and computationally intensive (Sec. 2.4.2).

Ongoing work by Hans Callebaut suggests that electron scattering with ionized impurities may play a more important role than previously thought in intersubband transport, and in some cases may be stronger than e-e scattering. The scattering matrix element for impurity scattering is a factor of four larger than for e-e scattering. A factor of two results from the fact that exchange effects are absent, and thus electrons are not restricted to interacting only with scatterers of opposite spin. The other factor of two results from the fact that electrons scatter with their full effective mass, not a reduced effective mass  $m^*/2$  as for e-e scattering. QCL designs are generally modulation doped away from the radiative states in order to limit the effects of impurity scattering on radiative lifetimes and linewidths. Even so, calculations of impurity scattering in the FL175C structure yielded a total scattering time of approximately 20 ps from level  $n = 5$  to  $n = 4$  and 3, where the doping was  $3 \times 10^{10} \text{ cm}^{-2}$  per module. In addition to directly contributing to transport, the additional elastic scattering will further heat the electron gas, thus increasing thermally activated phonon scattering as well. Further Monte Carlo simulations are underway to assess the importance of this transport channel more quantitatively.

Other relaxation paths turn out to be less important. Acoustic phonon scattering in particular is relatively inefficient, especially at low-temperatures where the phonon population is small. Typical acoustic phonon intersubband scattering times are in the 100 ps range [37, 58]. Furthermore, the scattering is quasi-elastic, as roughly 1 meV is carried away by the phonon [59], which makes it an inefficient relaxation and cooling mechanism at all but the lowest injection levels.

Interface roughness scattering is also unlikely to play a significant role, at least for intersubband scattering between the radiative levels. Calculation of this mechanism is difficult since it requires detailed knowledge of the microscopic growth characteristics. Smet performed calculations of interface roughness scattering and found typical intersubband scattering times greater than 100 ps for defect concentrations of  $10^{10} \text{ cm}^{-2}$  for interwell and intrawell transitions [37]. However, the importance of this mechanism strongly depends on the overlap of the wavefunctions with interfaces, as

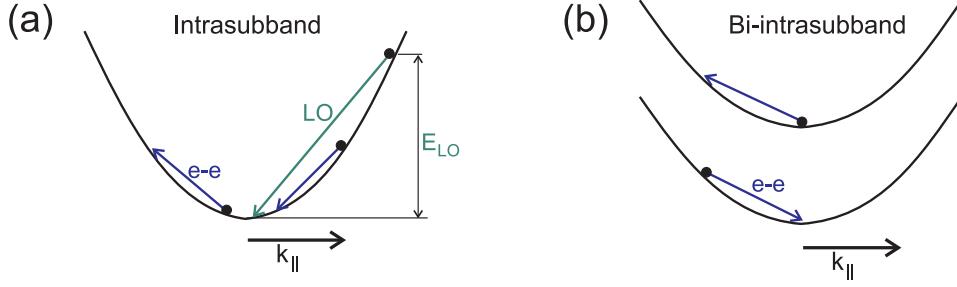


Figure 2-2: Important (a) intrasubband and (b) bi-intrasubband scattering mechanisms.

well as on growth characteristics, so this mechanism cannot be ruled out. Interface roughness may play more of a role in intrasubband scattering where it can contribute to radiative linewidth broadening.

Intrasubband transitions, as shown in Fig. 2-2, are also important in predicting device performance. Intrasubband LO-phonon scattering is an important process in cooling the subband electron gas. The effect of intrasubband e-e scattering is to thermalize the electron distribution inside a particular subband. Bi-intrasubband e-e scattering involves an electron scattering with another electron in a different subband without either electron leaving its subband. Energy is exchanged without affecting particle number. This process determines the rate at which electrons in different subbands reach the same temperature.

#### 2.4.1 Polar longitudinal optical (LO) phonon scattering

In this section, a brief derivation of electron-LO-phonon scattering rates in polar semiconductor heterostructures is given. The formalism used below to describe this scattering is borrowed from Smet[37, 58]. In the derivation, the phonon spectrum is taken as the equilibrium bulk spectrum, with modifications due to the heterostructure ignored. Effects of confined phonons and interface phonons in GaAs/AlGaAs quantum wells have been studied using macroscopic slab models [85]. These results compare favorably with a fully microscopic treatment of LO-phonon scattering in a 56 Å GaAs/AlAs quantum well performed by Rücker *et al.* [86]. They found that

for subband separations greater than the largest LO-phonon energy in the material, the LO-phonon emission rates were between the bulk rates for the well and barrier materials, tending toward the bulk values of GaAs and AlAs for the limiting cases of wide wells and narrow wells respectively. For subband separations in the range of the LO-phonon resonances, depending on the details of the structure, the effect of the complex phonon modes may be important. A detailed look at the inclusion of interface and confined phonon modes in calculations, and how it affects the design of terahertz quantum cascade lasers is given in Chap. 3.

In many calculations the use of bulk phonon modes for GaAs is reasonably accurate, and is sufficient for understanding device transport. Screening by the electron gas is neglected, although an account is given by Smet in Ref. [37] The optical phonon branch is taken as dispersionless at the  $\Gamma$ -point frequency  $\omega_{LO}$  ( $E_{LO} = 36.25$  meV in GaAs). The scattering rate for an electron initially in state  $|i, \mathbf{k}_i\rangle$  (subband  $i$ , in-plane wavevector  $\mathbf{k}_i$ ) to the final state  $|f, \mathbf{k}_f\rangle$  through an interaction potential  $H'$  is evaluated using Fermi's golden rule

$$W_{i \rightarrow f}(\mathbf{k}_i, \mathbf{k}_f) = \frac{2\pi}{\hbar} \left| \langle f, \mathbf{k}_f | H' | i, \mathbf{k}_i \rangle \right|^2 \delta(E_f(\mathbf{k}_f) - E_i(\mathbf{k}_i) \pm \hbar\omega_{LO}). \quad (2.39)$$

The envelope wavefunctions  $\psi_i(z)$  for each subband are obtained as described in Sec. 2.2. The electron-phonon interaction Hamiltonian takes the form of

$$H' = \sum_{\mathbf{q}} \left[ \alpha(\mathbf{q}) \left( e^{i\mathbf{q} \cdot \mathbf{r}} b_{\mathbf{q}} + e^{-i\mathbf{q} \cdot \mathbf{r}} b_{\mathbf{q}}^\dagger \right) \right], \quad (2.40)$$

where  $\alpha(\mathbf{q})$  is the electron-phonon interaction and  $b_{\mathbf{q}}$  and  $b_{\mathbf{q}}^\dagger$  are the creation and annihilation operators for a phonon in mode  $\mathbf{q}$ . The Fröhlich interaction strength for electron-polar-optical-phonon scattering is given in SI units by

$$|\alpha(\mathbf{q})|^2 = \frac{\hbar\omega_{LO}}{2} \frac{e^2}{q^2} \left( \frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_{dc}} \right), \quad (2.41)$$

where  $\epsilon_{dc}$  and  $\epsilon_\infty$  are the static and high frequency permittivities. The  $q^{-2}$  dependence of this term reduces scattering rates when large in-plane momentum transfers are

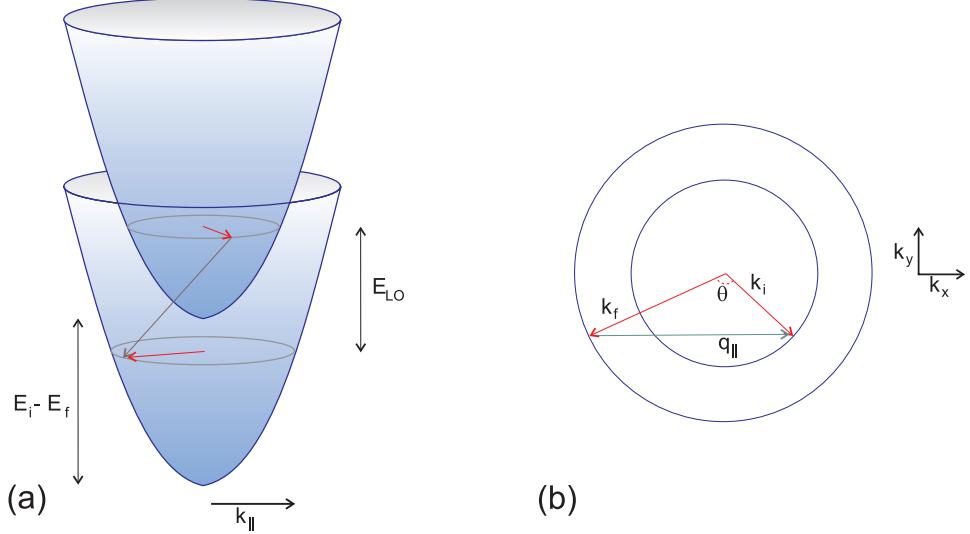


Figure 2-3: (a) Schematic illustration of intersubband LO-phonon scattering process. (b) In-plane reciprocal lattice space diagram illustrating the relationship between initial and final electron wavevectors  $k_i$  and  $k_f$  and in-plane phonon wavevector  $q_{\parallel}$ .

required. This reduces scattering between subbands with a large energy separation.

The matrix element is given by

$$\begin{aligned} \left| \langle f, \mathbf{k}_f | H' | i, \mathbf{k}_i \rangle \right|^2 &= \frac{e^2 \hbar \omega_{LO}}{2V} \left( \frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon_{dc}} \right) \frac{1}{q_z^2 + q_{\parallel}^2} |A_{i \rightarrow f}(q_z)|^2 \\ &\times \delta_{\mathbf{k}_i, \mathbf{k}_f \mp \mathbf{q}_{\parallel}} (n_{\omega_{LO}} + 1/2 \mp 1/2), \end{aligned} \quad (2.42)$$

where  $\mathbf{q}_{\parallel}$  and  $\mathbf{q}_z$  are the components of the phonon wavevector that are perpendicular (in-plane) and parallel to the growth axis ( $\hat{\mathbf{z}}$ ) respectively,  $n_{\omega_{LO}}$  is the Bose-Einstein occupation, and the upper and lower signs correspond to phonon absorption and emission respectively. The delta function ensures in-plane momentum conservation, and the form factor

$$A_{i \rightarrow f}(q_z) = \int_{-\infty}^{\infty} dz \psi_f^*(z) \psi_i^*(z) e^{\pm q_z z} \quad (2.43)$$

is related to the  $q_z$  momentum uncertainty due to the spatially localized envelope wavefunctions  $\psi_i(z)$  and  $\psi_f(z)$ .

This expression can then be integrated over the phonon modes  $\mathbf{q}$  and final states  $\mathbf{k}_f$  to yield the scattering rate from an initial wavevector  $W(\mathbf{k}_i)$ . Assuming parabolic

subband dispersion, the final states lie on a circle with radius  $k_f$  determined by conservation of energy.

$$k_f^2 = k_i^2 + \frac{2m^*(E_f(0) - E_i(0) \mp \hbar\omega_{LO})}{\hbar^2}. \quad (2.44)$$

Energy conservation and the in-plane momentum conservation rule allows us to write the phonon wavevector  $\mathbf{q}_{\parallel}$  in terms of  $\mathbf{k}_i$  and  $\mathbf{k}_f$ :

$$q_{\parallel}^2 = |\mathbf{k}_i - \mathbf{k}_f|^2 = k_i^2 + k_f^2 - 2k_i k_f \cos \theta, \quad (2.45)$$

where the angle  $\theta$  is the angle between the in-plane wavevectors  $\mathbf{k}_i$  and  $\mathbf{k}_f$ . This is illustrated schematically in Fig. 2-3. After summation over the phonon modes  $\mathbf{q}$ , Eq. (2.39) can be integrated over these final states  $\mathbf{k}_f$  to yield the total scattering rate from an initial wavevector:

$$\begin{aligned} W_{i \rightarrow f}^{abs}(\mathbf{k}_i) &= \frac{m^* e^2 \omega_{LO}}{8\pi \hbar^2} \left( \frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon_{dc}} \right) n_{\omega_{LO}} \int_0^{2\pi} d\theta B_{i \rightarrow f}(q_{\parallel}), \\ W_{i \rightarrow f}^{em}(\mathbf{k}_i) &= \frac{m^* e^2 \omega_{LO}}{8\pi \hbar^2} \left( \frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon_{dc}} \right) (n_{\omega_{LO}} + 1) \int_0^{2\pi} d\theta B_{i \rightarrow f}(q_{\parallel}), \end{aligned} \quad (2.46)$$

where  $B_{i \rightarrow f}$  is given by

$$B_{i \rightarrow f} = \int_{\infty}^{\infty} dz \int_{\infty}^{\infty} dz' \psi_f^*(z) \psi_i(z) \psi_i^*(z') \psi_f(z') \frac{1}{q_{\parallel}} e^{-q_{\parallel}|z-z'|}. \quad (2.47)$$

Effects of state blocking are neglected in Eq. (2.46). Intrasubband scattering times can also be calculated using this formalism, simply by setting  $i = f$ ,  $\psi_i(z) = \psi_f(z)$ , etc.

The total scattering time between subbands  $\tau_{i \rightarrow f}$  can then be obtained by averaging over all possible initial states in the subband:

$$\frac{1}{\tau_{i \rightarrow f}} = \frac{\int_0^{\infty} dE_k f_i(E_k) W_{i \rightarrow f}(E_k)}{\int_0^{\infty} dE_k f(E_k)}, \quad (2.48)$$

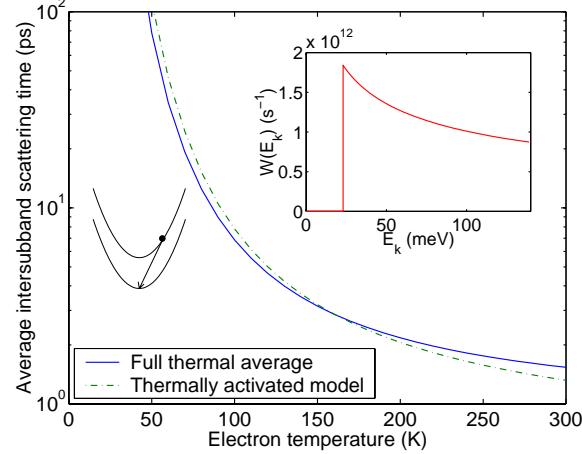


Figure 2-4: Average intersubband LO-phonon scattering rate for subband separation  $E_{fi} = 13.3$  meV. Rate is calculated for  $5 \rightarrow 4$  transition for FL175C device as shown in Fig. 7-1. Inset shows the raw scattering rate versus in-plane energy.

where  $E_k = \hbar^2 k_{\parallel}^2 / 2m^*$  is the in-plane kinetic energy in the initial subband, and  $f_i[E_k]$  is the corresponding Fermi function. Non-equilibrium subband distributions can be dealt with by assigning different quasi-Fermi levels and electron temperatures for the initial and final Fermi distributions. For most of our intersubband scattering calculations, the final states were assumed to be empty.

When the transition energy  $E_{fi} < E_{LO}$ , rather than perform the full average described in Eq. (2.48), one can use the thermally activated expression

$$\frac{1}{\tau_{i \rightarrow f}} \approx W_{i \rightarrow f}^{(hot)} \exp \left( \frac{E_{fi} - E_{LO}}{k_B T} \right), \quad (2.49)$$

where  $W_{i \rightarrow f}^{(hot)}$  is the scattering rate at the lowest energy  $E_k$  in the subband where LO-phonon scattering is energetically allowed. Typical scattering rates, calculated using both the proper thermal average (2.48) and the approximation (2.49), are shown in Fig. 2-4 for an energy separation of  $E_{fi} = 13.3$  meV.

## 2.4.2 Electron-electron scattering

The importance of the electron-electron (e-e) interaction mechanism in terahertz QCLs was recognized as early as 1989 by Borenstein and Katz [87]. While an e-e

scattering event can change the energy of an individual electron, no energy is removed from the global electron system. Rather, the overall distribution is thermalized. Electron-electron scattering is expected to be important for intersubband transitions between closely spaced subbands at low temperatures where LO-phonon scattering is suppressed. It also plays a role in miniband transport by promoting scattering between miniband states.

Experimentally, it has proven difficult to measure intersubband relaxation times for subband separations  $E_{21} < E_{LO}$  and unambiguously attribute those lifetimes to e-e intersubband scattering. In fact, a wide variety of relaxation times has been reported in the literature [88, 89, 90, 91, 92] and there is still large uncertainty in the overall role e-e scattering plays. Examples include the steady state absorption and photovoltage measurements made by Heyman *et al.* using a far-infrared laser for a system with  $E_{21} = 11$  meV. Depending on the laser pump intensity, measured lifetimes at 10 K varied from 15 ps to 1 ns; however it is difficult to correlate these numbers to absolute subband populations. Much shorter lifetimes were measured by Hartig *et al.* [91] in observation of time resolved photoluminescence in a structure with  $E_{21} = 28$  meV in single and coupled quantum well structures. Scattering times on the order  $\sim 2$  ps were observed for excitation densities of  $5 \times 10^{10} \text{ cm}^{-2}$ . In a similar experiment, Hartig *et al.* [92] measured a decrease of the lifetime in a coupled quantum well from 40 ps down to 5 ps as the carrier density was increased from  $10^8 \text{ cm}^{-2}$  to  $10^{11} \text{ cm}^{-2}$ . However, despite the authors' claims to the contrary, it is not clear that these lifetimes are due to intersubband e-e scattering and not thermally activated LO-phonon scattering. This is suggested by the close proximity of the subband separation to the  $E_{LO} = 36$  meV, and the fact that when using interband pumping to create a cold electron distribution in level 2, a hot distribution of similar size has been pumped into level 1. Monte-Carlo simulations by Dür also suggest this conclusion [93]. The presence of electron-hole scattering in interband pumped experiments further complicates matters.

Experimental confirmation of the fast intrasubband e-e scattering predicted [58] has proven more successful. For moderate electron sheet densities ( $10^{10}\text{--}10^{11} \text{ cm}^{-2}$ ),

energy exchange within a given subband is extremely fast, and happens on a subpicosecond time scale [82]. Energy exchange between subbands takes somewhat longer. For example, photoluminescence measurements by Hartig *et al.* found that subbands separated by about 80 meV took from 1–10 ps to equilibrate temperature for densities from  $10^{10}$ – $10^{11}$  cm $^{-2}$ . Since these times may be on the same order as the intersubband lifetimes, different subbands may maintain different temperatures in the steady state. This is consistent with Monte Carlo simulations performed in our group by Hans Callebaut and others [84, 94], and is contrary to the conclusion by Harrison [95]. Energy is exchanged between subbands through bi-intrasubband scattering events, such as pictured in Fig. 2-2(b). While these events occur on the same subpicosecond time scale as intrasubband scattering, very little energy is exchanged ( $\ll 1$  meV). Therefore, many such scattering events are required and thermalization might take several picoseconds, depending on the carrier densities [94].

Accurate calculation of e-e scattering rates in quantum well systems is a theoretically complicated many-body problem, and a full theoretical analysis is beyond the scope of this thesis. A brief summation of some relevant results follow. Electron-electron interaction is mediated by the coulomb potential, which is screened by the surrounding electron system. In low density systems where the single particle picture is applicable, Fermi's Golden Rule is the most common method for calculating scattering rates. Because of the mathematical complexity of describing screening in a multi-subband system at finite temperature, simplified models are typically used.

Smet *et al.* analyzed scattering using this method using a single subband static screening model. The abbreviated results of this derivation are described below. For a full description, see Refs. [37, 58].<sup>1</sup> Consider an electron in subband  $i$  with wavevector  $\mathbf{k}_i$  scattering with another electron in subband  $j$  with wavevector  $\mathbf{k}_j$ , where the electrons end up in subbands  $f$  and  $g$  with wavevectors  $\mathbf{k}_f$  and  $\mathbf{k}_g$  respectively.

---

<sup>1</sup>Some derivations of e-e scattering rates in the literature are too fast by a factor of four (see Ref. [96]). For example, in Ref. [58], the right side of the equation (51) should be reduced by a factor of four, and the plotted scattering times should be scaled accordingly. The same is true for rates calculated by the e-e scattering code in Harrison [97].

The unscreened matrix element for this event is

$$H'_{i,j \rightarrow f,g}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_f, \mathbf{k}_g) = \frac{2\pi e^2}{S_{\parallel}\epsilon} A_{i,j \rightarrow f,g}(q_{\parallel}) \delta(\mathbf{k}_f + \mathbf{k}_g - \mathbf{k}_i - \mathbf{k}_j), \quad (2.50)$$

where

$$\mathbf{q}_{\parallel} = |\mathbf{k}_i - \mathbf{k}_f|, \quad (2.51)$$

and the form factor is

$$A_{i,j \rightarrow f,g}(q_{\parallel}) = \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} dz' \psi_i(z) \psi_f^*(z) \psi_j(z') \psi_g^* e^{-q_{\parallel}|z-z'|}. \quad (2.52)$$

Only scattering with anti-parallel spin electron is considered due to exchange effects.

The total scattering out rate for an electron in state  $|i, \mathbf{k}_i\rangle$  is given by

$$W_{i,j \rightarrow f,g}(\mathbf{k}_i) = \frac{e^4}{2\pi\hbar\epsilon^2} \int d^2\mathbf{k}_j \int d^2\mathbf{k}_f \int d^2\mathbf{k}_g \frac{|A_{i,j \rightarrow f,g}(q_{\parallel})|^2}{\epsilon_{sc}^2(q_{\parallel}, T) q_{\parallel}^2} f_{j,\mathbf{k}_j} (1 - f_{g,\mathbf{k}_g})(1 - f_{f,\mathbf{k}_f}) \times \delta(E_f(\mathbf{k}_f) + E_g(\mathbf{k}_g) - E_i(\mathbf{k}_i) - E_j(\mathbf{k}_j)) \delta(\mathbf{k}_f + \mathbf{k}_g - \mathbf{k}_i - \mathbf{k}_j), \quad (2.53)$$

where  $\epsilon_{sc}^2(q_{\parallel}, T)$  represents the correction due to the dielectric constant resulting from screening.

Analysis of various form factors reveals that  $22 \rightarrow 11$  scattering is typically the most efficient, especially since  $A_{22 \rightarrow 21}$  vanishes completely for symmetric structures [58]. This scattering is doubly efficient since both electrons end up in the lower subband. Intrasubband events such as  $11 \rightarrow 11$  and bi-intrasubband events such as  $21 \rightarrow 21$  tend to have rates at least an order of magnitude faster than intersubband rates. As an example, e-e scattering times for the device FL175C for the  $5 \rightarrow 4$  transition with  $E_{54} = 13.3$  meV are shown in Fig. 2-5. No screening is used and the lower population is assumed to be empty. These e-e scattering times are somewhat longer than is suggested by the experimental device behavior and Monte Carlo simulations [84].

Recent theoretical work by Lee and Galbraith has suggested that the use of the single-subband screening model substantially underestimates both intersubband and

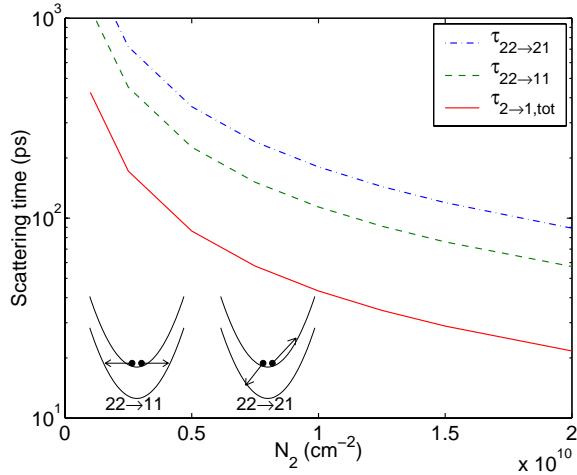


Figure 2-5: Intersubband electron-electron scattering rate versus upper state population  $N_2$  ( $N_1 = 0$ ) calculated for separation  $E_{fi} = 13.3$  meV according to (2.53), without screening.  $21 \rightarrow 11$  rate is negligible. Rate is calculated for  $5 \rightarrow 4$  transition for FL175C device as shown in Fig. 7-1 for  $T_e = 77$  K. When calculating  $\tau_{2 \rightarrow 1, \text{tot}}$ , the rate  $\tau_{22 \rightarrow 11}$  is counted twice, since two electrons relax to the lower subband.

intrasubband e-e scattering rates. They model the screening using the full dynamic, finite-temperature, multi-subband dielectric function derived in the random-phase approximation (RPA). For the intersubband case, the difference between the static and dynamic models is small; it is the inclusion of the multi-subband aspect that reduces the screening effectiveness and increases the scattering rate. In fact calculations indicate that for moderate densities ( $< 5 \times 10^{11} \text{ cm}^{-2}$ ), there is little difference between intersubband rates calculated with and without screening [98]. The dynamic model allows for the treatment of plasmon emission, which may significantly impact intra-subband scattering rates as well. Also, by including the dynamic lattice permittivity, plasmon-phonon coupling was included, which further increased the scattering rates.

Additionally, even more sophisticated theoretical approaches are available beyond Fermi's Golden Rule. For example, both Tavares *et al.* [99] and Kempa *et al.* [100] have calculated e-e interaction by obtaining the imaginary part of the self-energy matrix which includes a fully dynamic multi-subband RPA dielectric treatment. However, the computational expense may outweigh the benefits, and may not be necessary for QCL design.

Due to the complicated nature of the e-e scattering calculation, and the lack of definitive experimental confirmation of calculated intersubband rates, when designing terahertz QCLs it is best not to rely on absolute numbers, but rather on phenomenological guidelines. Hyldgaard and Wilkins proposed a simplified model that uses a fixed screening length [101]. Assuming an empty lower state, they obtained the scattering rate

$$W_{e-e} \propto \frac{N_2}{E_{21}} |U(q_{\parallel})|^2, \quad (2.54)$$

where  $N_2$  is the upper state population,  $E_{21}$  is the subband separation, and  $U(q_{\parallel})$  is a form factor similar to (2.52) that includes the fixed screening. While this model is perhaps too simple for accurate rate calculations, it provides a qualitative understanding of the essential scattering behavior at low electron densities. That is, the scattering rate increases with electron density and is inversely proportional to subband separation. This model will of course break down at higher densities (perhaps  $> 10^{11} \text{ cm}^{-2}$ ), as screening and band filling effects become increasingly important.

## 2.5 Resonant tunneling transport and anticrossings

Resonant tunneling is perhaps the most critical transport mechanism in quantum cascade structures. While a qualitative understanding of this process is straightforward and has successfully been used for some time to design QCLs, a full theoretical understanding that is applicable to practical structure design is still somewhat elusive. The role of coherent and incoherent transport in quantum cascade structures has become an active experimental and theoretical area of research [102, 103], and a full analysis is beyond the scope of this thesis. Nevertheless, in the section below I present some basic methods for the analysis of tunneling that are useful for QCL design.

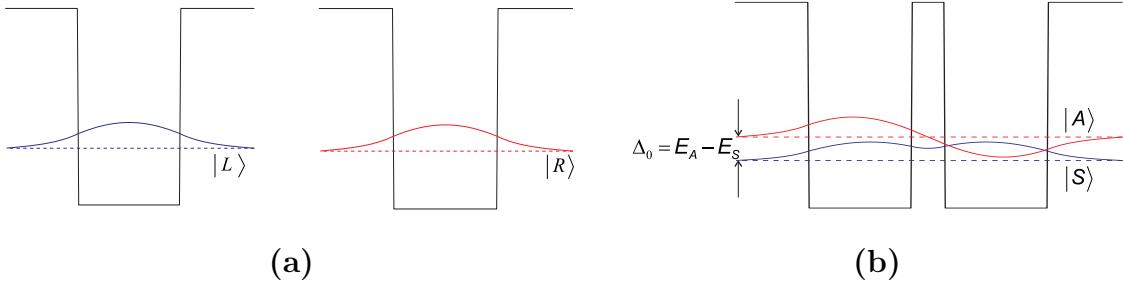


Figure 2-6: Anticrossing between two single well states.

### 2.5.1 Semiclassical “coherent” model

The primary model used to describe transport in our structures is a semiclassical Boltzmann-like model, where transport is described via in- and out-scattering from stationary states that are solutions to the Schrödinger equation as described in Eq. (2.5). Since the Hamiltonian for the system includes the well and barrier structure, interactions and tunneling are already included in the eigenstates. In this model, transport occurs via scattering between these states (labeled by subband index and in-plane momentum) as described by Fermi’s Golden Rule. No coherent interactions between states are considered, although individual wavefunctions are sometimes referred to as “coherent” in that they are spatially extended across the calculated structure. Limits to the wavefunction delocalization which might be caused by dephasing are decided upon in an ad-hoc manner, according to the qualitative understanding of the structure in question.

For example, in this picture “tunneling” between two quantum well states  $|L\rangle$  and  $|R\rangle$  would be described by solving the coupled well system and obtaining a new set of delocalized eigenstates (see Fig. 2-6). The minimum separation between the symmetric and antisymmetric states is known as the anticrossing gap  $\Delta_0 = |E_A - E_S|$ . When using such sets of delocalized states as basis functions with the semiclassical model, there is no interwell tunneling time as such; transport through the barrier effectively takes place via intersubband scattering into and out of  $|A\rangle$  and  $|S\rangle$  from other subbands. This is shown schematically in Fig. 2-7, assuming a total intrawell relaxation time  $\tau$ . For example, assuming a sheet density of  $N$  per well, one would

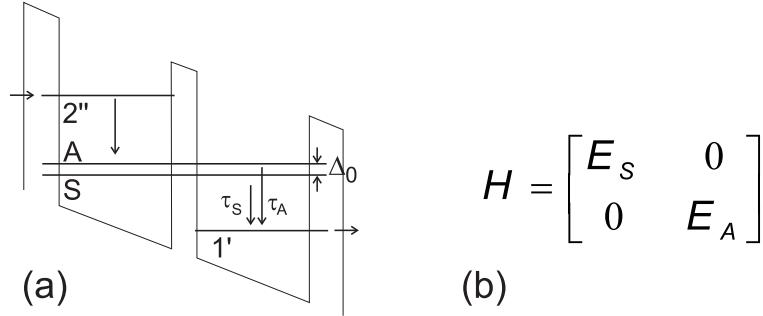


Figure 2-7: (a) Schematic for semiclassical model for resonant tunneling through a barrier. (b) Hamiltonian using delocalized basis states with energies  $E_A$  and  $E_S$ , relaxation times  $\tau_A$  and  $\tau_S$ , and anticrossing gap  $\Delta_0$ .

calculate the current density to be

$$J = \frac{eN_A}{\tau_A} + \frac{eN_S}{\tau_S} \approx \frac{eN}{2\tau}, \quad (2.55)$$

where the last term is equivalent if  $N_A \approx N_S = N/2$  and  $\tau^{-1} = \tau_A^{-1} + \tau_S^{-1}$ .

While this picture is accurate for strong coupling (large  $\Delta_0$ ), for weaker interactions neglecting the tunneling time may overestimate the current density through the barrier. For example, this model would predict the peak current density in QCL structures to be independent of the injection barrier thickness, which is experimentally known to be false (see for example Sec. 6.4). Since a localized wavefunction can be prepared via a coherent superposition of the delocalized states  $|A\rangle$  and  $|S\rangle$ , it is common to define the tunneling time as the time it takes for such a localized state to move through the barrier, i.e. half of the Rabi period  $T_{Rabi} = \hbar/\Delta_0$ . In order to account for the effects of dephasing on transport between two weakly-coupled states, a density matrix approach can be used.

### 2.5.2 Tight-binding density matrix approach

The semiclassical approach described above does not readily lend itself to describing coherent evolution of wavefunctions, nor does it allow treatment of phase breaking processes. Ensemble density matrix formalism addresses both issues. While such a

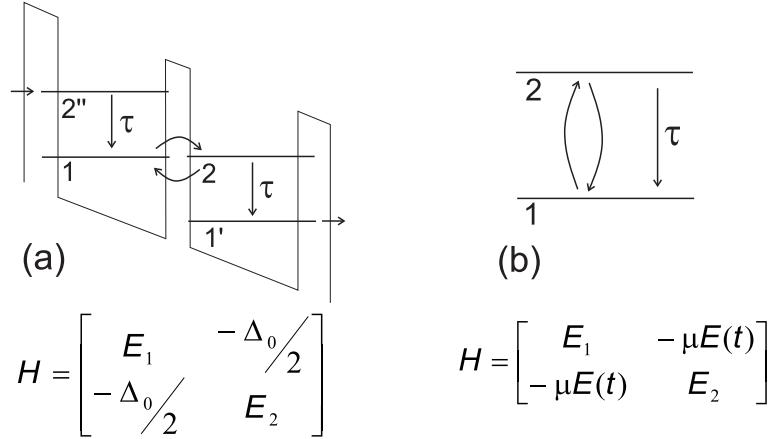


Figure 2-8: Schematic and Hamiltonian for (a) tight-binding resonant tunneling model with interaction  $-\Delta_0/2$ , and (b) the analogous two-level model with dipole interaction  $-\mu E(t)$ . The  $2 \rightarrow 1$  relaxation time is  $\tau$ .

formalism is difficult to apply to the entire active region of a QCL, it is instructive to use this model to describe situations where resonant tunneling might be expected to be a transport bottleneck. Tunneling through the injection barrier in a quantum cascade structure is the most appropriate case, since coupling is relatively weak and efficient injection of electrons into a single state is desired. This model would also be useful for the analysis of the resonant phonon depopulation mechanism used in the lasers described in Chap. 7. In this section, two tight-binding density matrix models for resonant tunneling are described: a two-level system in which carriers tunnel between states and then relax back into the initial state, and a three-level system, in which a reservoir level is added to decouple the out-scattering and in-scattering processes.

### Two-level system

Consider resonant tunneling from state  $|1\rangle$  into  $|2\rangle$  with tight-binding coupling  $-\Delta_0/2$  (Fig. 2-8(a)). These localized states are eigenstates of single quantum wells only, and are not eigenstates of the coupled well system. If we were to diagonalize the Hamiltonian, we would obtain symmetric and antisymmetric eigenstates separated by an anticrossing gap  $\Delta_0$ , similar to what is shown in Fig. 2-6(b). However, the use

of localized states as our basis provides a physical picture where dephasing processes interrupt coherent tunneling, and effectively localize the wavefunctions. The structure is periodic, so electrons in  $|2\rangle$  relax back into  $|1\rangle$  with time  $\tau$ . As discussed at the end of this section, this model for resonant tunneling is almost directly analogous to the optical Bloch equations for a two-level system in the presence of an electromagnetic field. The schematic for such an analogous two-level system is shown in Fig. 2-8(b).

Using standard density matrix formalism, the wavefunction for an electron in the ensemble is given by

$$|\psi(t)\rangle = \sum c_i(t)|i\rangle, \quad (2.56)$$

and the density matrix  $\rho$  for a mixed state is defined by the ensemble average

$$\rho_{ij} = \langle c_i c_j^* \rangle. \quad (2.57)$$

The diagonal elements  $\rho_{ii}$  correspond to the state populations, and the off-diagonal elements represent the average degree of coherence between basis states. The time evolution of  $\rho$  is described by the equation of motion

$$\frac{d}{dt}\rho = -\frac{i}{\hbar}[H, \rho], \quad (2.58)$$

where  $H$  is the Hamiltonian of the total system, as shown in Fig. 2-8. We obtain the following set of coupled equations that describes both the population difference  $\rho_{11} - \rho_{22}$  and the off-diagonal element  $\rho_{21}$  (note  $\rho_{21} = \rho_{12}^*$ ):

$$\frac{d}{dt}(\rho_{11} - \rho_{22}) = \frac{i\Delta_0}{\hbar}(\rho_{21} - \rho_{21}^*) - \frac{(\rho_{11} - \rho_{22}) - (\rho_{11} - \rho_{22})_0}{\tau}, \quad (2.59)$$

$$\frac{d}{dt}\rho_{21} = \frac{i\Delta_0}{2\hbar}(\rho_{11} - \rho_{22}) - \frac{iE_{21}}{\hbar}\rho_{21} - \frac{\rho_{21}}{\tau_\parallel}. \quad (2.60)$$

In this equation  $E_{21} = E_2 - E_1$  is the detuning from resonance, and  $(\rho_{11} - \rho_{22})_0 = 1$  is the population difference at equilibrium. Note that two phenomenological terms have been added. In (2.59), the last term allows the population to relax to its equilibrium value with time  $\tau$ ; in (2.60), the last term represents the relaxation of phase coherence

with time  $\tau_{\parallel}$  due to various scattering mechanisms. In analogy to the case of homogeneous broadening of a radiative transition (see Appendix A), the dephasing time  $\tau_{\parallel}$  is defined to include both the effects of population relaxation and pure dephasing according to

$$\frac{1}{\tau_{\parallel}} = \frac{1}{2\tau_1} + \frac{1}{2\tau_2} + \frac{1}{T_2^*} = \frac{1}{\tau} + \frac{1}{T_2^*}, \quad (2.61)$$

where  $T_2^*$  is the pure dephasing time and  $\tau_1$  and  $\tau_2$  are the lifetimes of each level. While in fact state  $|1\rangle$  doesn't have a relaxation time *per se*, electrons scatter out of  $|2\rangle$  with time  $\tau$  and into  $|1\rangle$  with the same time  $\tau$ , so we consider  $\tau_1 = \tau_2 = \tau$  for this two-level system. The justification for this definition is discussed more completely in the next section on the three-level model.

In steady state, we can solve for the various elements of the density matrix and obtain an expression for the current density through the barrier:

$$J = \frac{eN\rho_{22}}{\tau} = \frac{eN}{2} \frac{(\frac{\Delta_0}{\hbar})^2\tau_{\parallel}}{1 + (\frac{E_{21}}{\hbar})^2\tau_{\parallel}^2 + (\frac{\Delta_0}{\hbar})^2\tau\tau_{\parallel}}, \quad (2.62)$$

where  $N$  is the total electron sheet density of the two levels. This expression, which is equivalent to the one reported in Kazarinov and Suris [31] and Sirtori *et al.* [104], describes the resonant current versus detuning bias  $E_{21}$  as a Lorentzian with a full-width half-max of

$$\Delta E_{21,FWHM} = \frac{2\hbar}{\tau_{\parallel}} \left[ 1 + \left( \frac{\Delta_0}{\hbar} \right)^2 \tau\tau_{\parallel} \right]^{1/2}. \quad (2.63)$$

Maximum current is obtained at resonance for  $E_{21} = 0$  where

$$J_{max} = \frac{eN}{2} \frac{(\frac{\Delta_0}{\hbar})^2\tau_{\parallel}}{1 + (\frac{\Delta_0}{\hbar})^2\tau\tau_{\parallel}}. \quad (2.64)$$

The tight-binding model provides a picture where the localized wavepacket oscillates between wells with a frequency  $\omega_{Rabi} = \Delta_0/\hbar$  with a damping rate of approximately  $\tau_{\parallel}^{-1}$ , where the wavepacket is subject to scattering with a rate  $\tau^{-1}$  every time it is in state  $|2\rangle$ . Examination of the expression (2.64) reveals that there are two regimes of operation. In the limit  $(\Delta_0/\hbar)^2\tau\tau_{\parallel} \gg 1$ , we have "coherent" tunneling, where we

obtain

$$J = \frac{eN}{2\tau}. \quad (2.65)$$

This describes a situation where the a wavepacket is able to oscillate between states many times before relaxing, and thus the relaxation rate  $\tau^{-1}$  and not  $\omega_{Rabi}$  is the transport bottleneck. Therefore, in the steady state the population will be distributed equally between  $|1\rangle$  and  $|2\rangle$ . This expression for current is essentially equivalent to (2.55), which was derived using the semiclassical model.

The “incoherent” limit is reached for  $(\Delta_0/\hbar)^2\tau\tau_{\parallel} \ll 1$ , which gives a current density of

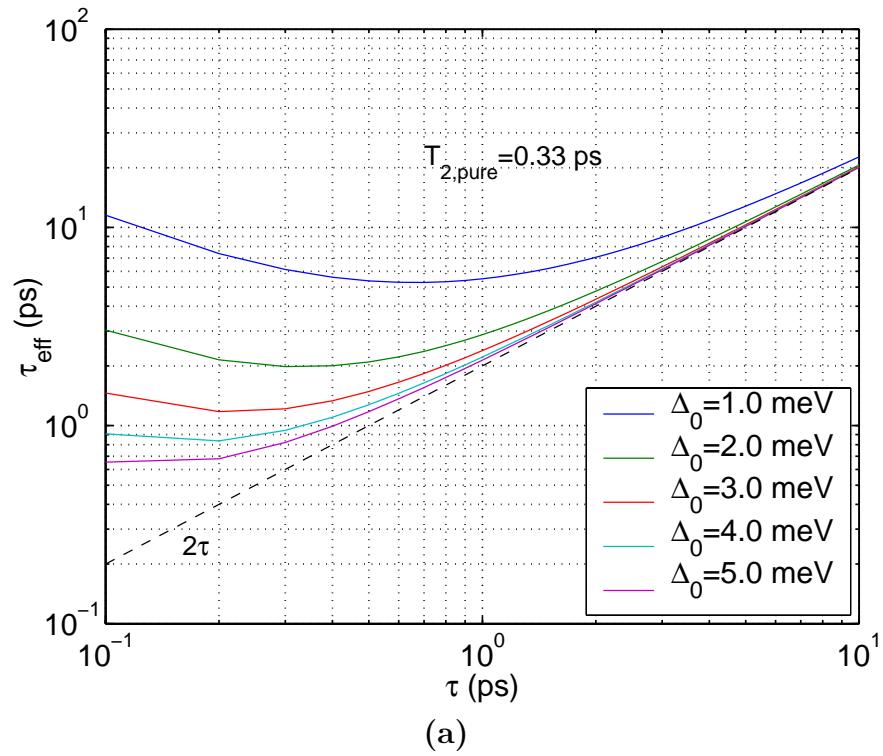
$$J = \frac{eN}{2} \left( \frac{\Delta_0}{\hbar} \right)^2 \tau_{\parallel} = \frac{eN}{2} \omega_{Rabi} (\omega_{Rabi}\tau_{\parallel}). \quad (2.66)$$

For the incoherent case, the current is limited by the  $\omega_{Rabi}$  and the dephasing rate  $\tau_{\parallel}^{-1}$ . In other words, only a small fraction of the electrons in the ensemble successfully tunnel through the barrier without interruption, and those that do quickly scatter before tunneling back. Thus tunneling transport can be considered one-way—from  $|1\rangle \rightarrow |2\rangle$ —and a steady state population difference will develop between the two states. From an energy viewpoint, the level broadening due to scattering is larger than the anticrossing gap, and no distinct doublet will be observed. Thus the states can be considered localized.

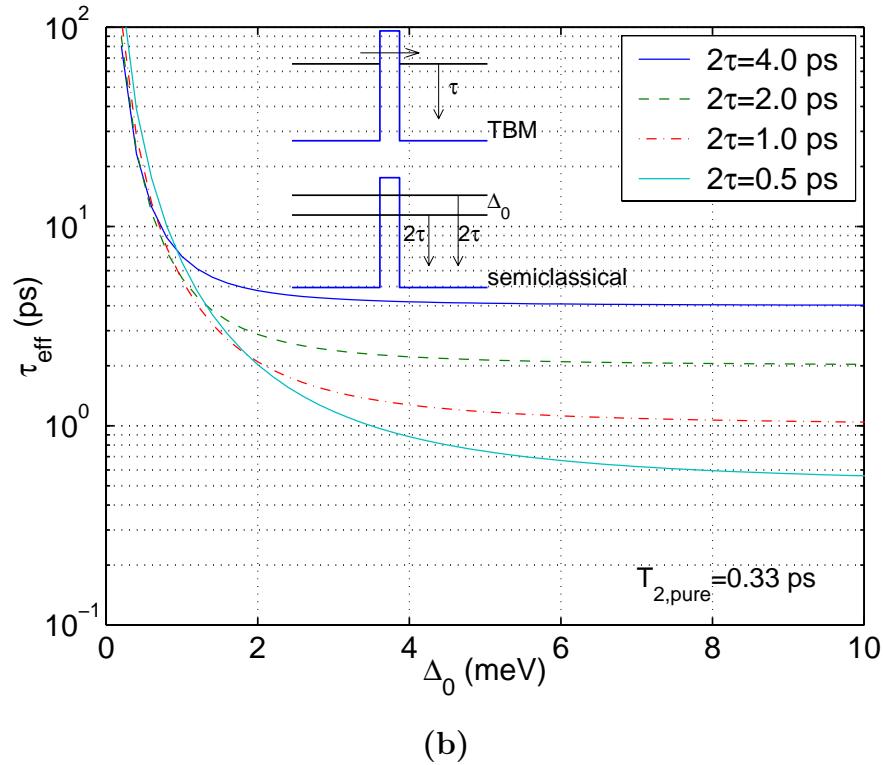
The effect of dephasing on resonant tunneling transport is illustrated in Fig. 2-9 using an effective transport time  $\tau_{eff}$  defined as

$$\frac{1}{\tau_{eff}} = \frac{J_{max}}{eN} = \frac{(\frac{\Delta_0}{\hbar})^2\tau_{\parallel}}{2(1 + (\frac{\Delta_0}{\hbar})^2\tau\tau_{\parallel})} \quad (2.67)$$

is plotted versus lifetime  $\tau$  and anticrossing gap  $\Delta_0$  for various typical parameters. Although the value of this parameter is highly uncertain, a pure dephasing time of  $T_2^* = 0.33$  ps is used, which corresponds to a FWHM linewidth of  $(\pi T_2^*)^{-1} \simeq 1$  THz ( $\simeq 4$  meV) when lifetime broadening is neglected. This number is chosen to correspond with the measured spontaneous emission linewidth from our FL175C device (Sec. 7.2). Although this is measured for a radiative transition and not resonant



(a)



(b)

Figure 2-9: Effective tunneling transport times  $\tau_{\text{eff}}$  as calculated according to Eq. (2.64) versus (a) lifetime  $\tau$  and (b) anticrossing gap  $\Delta_0$  for a pure dephasing time of  $T_2^* = 0.33$  ps.

tunneling, the dephasing times should be similar. This value for  $T_2^*$  is also consistent with experimentally measured dephasing times of  $T_2 \sim 0.5$  ps for resonant tunneling injection in a quantum cascade structure [103]. These figures illustrate the transition from incoherent to coherent transport, which occurs for large  $\Delta_0$  and  $\tau$ , and is characterized by  $\tau_{eff} \approx 2\tau$ .

In order to obtain a clearer idea of the effect of dephasing on the interaction, we revisit the coupled equations (2.59) and (2.60) to obtain a second-order differential equation for the population difference:

$$\frac{d^2}{dt^2}(\rho_{11} - \rho_{22}) = -\left(\frac{\Delta_0^2}{\hbar^2} + \frac{1}{\tau\tau_{||}}\right)(\rho_{11} - \rho_{22}) - \left(\frac{1}{\tau} + \frac{1}{\tau_{||}}\right)\frac{d}{dt}(\rho_{11} - \rho_{22}) + \frac{1}{\tau\tau_{||}}. \quad (2.68)$$

Examination of this equation reveals that it has the form of a damped harmonic oscillator with constant driving. The driving term  $(\tau\tau_{||})^{-1}$  is present because the relaxation is asymmetric, i.e. electrons can relax from  $|2\rangle$  to  $|1\rangle$  but not vice versa.

We define the parameters

$$\omega_0^2 = \left(\frac{\Delta_0^2}{\hbar^2} + \frac{1}{\tau\tau_{||}}\right), \text{ and } \gamma = \frac{1}{2}\left(\frac{1}{\tau} + \frac{1}{\tau_{||}}\right), \quad (2.69)$$

where  $\omega_0$  is the fundamental frequency, and  $\gamma$  is the damping factor. The damped frequency  $\omega_\ell$  is given by

$$\omega_\ell = \sqrt{\omega_0^2 - \gamma^2} = \left[\frac{\Delta_0^2}{\hbar^2} - \frac{1}{4}\left(\frac{1}{\tau_{||}} - \frac{1}{\tau}\right)^2\right]^{1/2} = \left[\frac{\Delta_0^2}{\hbar^2} - \frac{1}{4}\left(\frac{1}{T_2^*}\right)^2\right]^{1/2} \quad (2.70)$$

The right side of (2.70) was obtained by substituting the definition of  $\tau_{||}$  given in Eq. (2.61). This reduction of the oscillation frequency due to damping is the result of a reduction of the effective anticrossing gap due to pure dephasing  $T_2^*$ . The degree of this reduction combined with dephasing state broadening will determine whether any energy splitting is experimentally observable. Critical damping occurs when  $\omega_0^2 = \gamma^2$

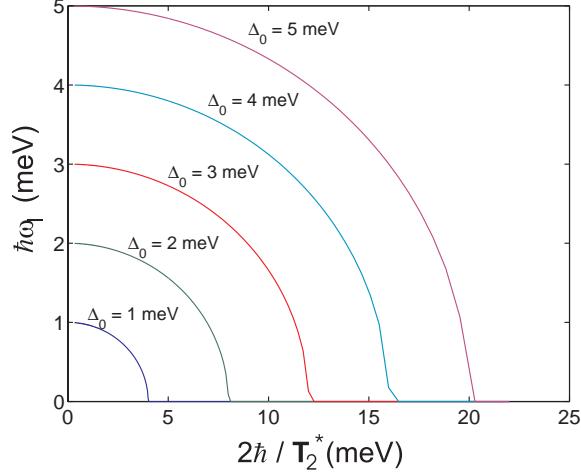


Figure 2-10: Damped oscillation frequency  $\omega_l$  versus the pure dephasing FWHM energy linewidth as given in Eq. (2.70).

which corresponds to

$$\frac{\Delta_0}{\hbar} = \frac{1}{2T_2^*}, \quad \text{or} \quad f_{Rabi} = \frac{\Delta_0}{\hbar} = \frac{1}{4} \frac{1}{\pi T_2^*} = \frac{1}{4} \text{FWHM} \quad (2.71)$$

i.e. when the anticrossing gap is equal to one fourth of the pure dephasing linewidth broadening. For less severe damping, the population difference will oscillate coherently before settling to its steady state value. In the over-damped case ( $\omega_0^2 < \gamma^2$ ), wavepacket transport through the barrier is severely impeded by dephasing. For most electrons in the ensemble, the coherent interaction is interrupted before a significant fraction of the wavepacket is transferred across the barrier. Rather than oscillating back and forth, it only traverses in one direction before being scattered, which yields the classical picture of tunneling based on barrier transparency.

The tight-binding density matrix model can be a useful tool in evaluating the effect of dephasing on transport due to resonant tunneling, which is beyond the reach of the traditional semiclassical model. Still, at this point this model is best introduced phenomenologically to describe tunneling at potential transport bottlenecks in the structure. The complexity and uncertainty in determining appropriate dephasing times limit somewhat its predictive value. Also, the use of localized wavefunctions as

basis states instead of the eigenstates of the full conduction band profile Hamiltonian limits the ability to compute scattering rates and obtain physical understanding of a structure. Since a structure is often modeled in terms of coherent, extended states, the calculation of quantities such as dipole matrix elements that depend strongly on wavefunction symmetry becomes more difficult and less intuitive when fully localized basis functions are used. The alternative is to use the delocalized eigenstates of the full multiple quantum well Hamiltonian as basis states. However, in this picture, dephasing would cause the electron ensemble to simply collapse into the spatially extended basis states. Such a model would fail to model the effect of dephasing on transport due to wavefunction localization.

### Three-level model and dephasing

In the section above, since the lifetime  $\tau$  and total dephasing time  $\tau_{\parallel}$  were added phenomenologically to the density matrix equation of motion Eqs. (2.59) and (2.60), it is worthwhile to consider their relationship. In other words, how does the lifetime  $\tau$  enter into the dephasing time  $\tau_{\parallel}$ , if at all? In Eq. (2.61) we defined the dephasing time of  $\rho_{21}$  as

$$\frac{1}{\tau_{\parallel}} = \frac{1}{2\tau_1} + \frac{1}{2\tau_2} + \frac{1}{T_2^*}, \quad (2.72)$$

where  $T_2^*$  was the pure dephasing time of the ensemble, not associated with any state transitions. This definition can be justified classically; it is analogous to the description of homogeneous broadening given in Appendix A. Pure dephasing  $T_2^*$  reflects the decay of an amplitude while  $\tau_2$  and  $\tau_1$  are energy or population decay rates, hence pure dephasing contributes doubly to energy decay. In this section, I give a quantum justification for this definition of dephasing. This is difficult to do in a two-level system, since it is impossible to decouple the out-scattering from the in-scattering (pumping) events.

Therefore, the two-level model is modified to include a third state  $|G\rangle$ , which acts as a reservoir, thus decoupling population relaxation and pumping. For the optical coupling analogy, an additional level below the two radiative states stores population.

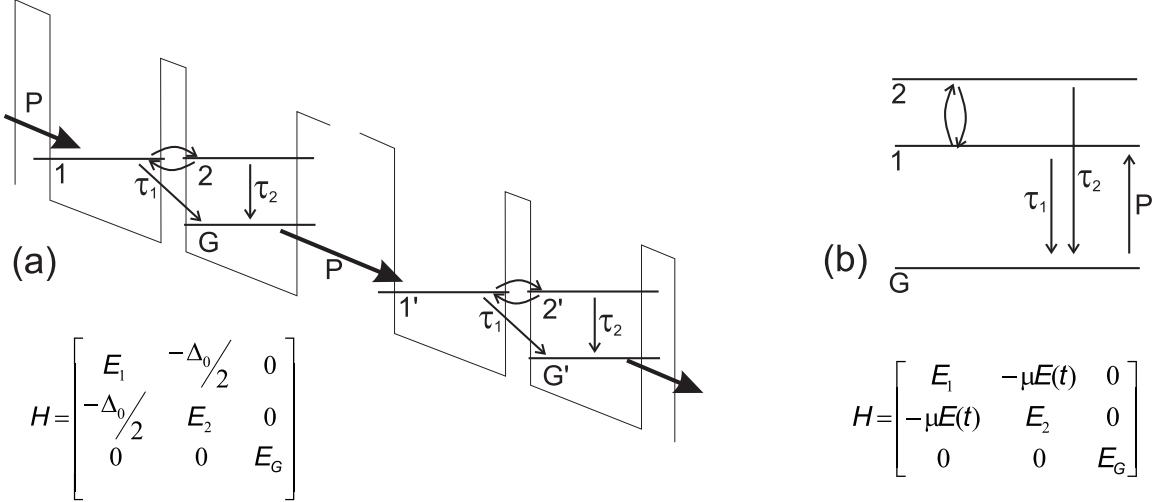


Figure 2-11: (a) Schematic for three-level tight-binding resonant tunneling model with two-level interaction  $-\Delta_0/2$  and with a third reservoir state  $G$  and (b) the analogous three-level optical model with dipole interaction  $-\mu E(t)$ . The system is driven by pumping  $P$ , and the population relaxation times into reservoir are  $\tau_2$  and  $\tau_1$ .

Schematics and Hamiltonians for this model considered with tight-binding resonant tunneling coupling and electric dipole coupling are shown in Fig. 2-11. Note that  $|G\rangle$  only acts to store population, and there is no coherent interaction with other levels ( $\rho_{32} = \rho_{31} = 0$ ). Although not present in the previous two-level model, a relaxation channel  $\tau_1$  from  $|1\rangle$  to  $|G\rangle$  has been included for completeness. In actual resonant tunneling structures, this relaxation channel will likely be much slower than from  $|2\rangle \rightarrow |G\rangle$  ( $\tau_2 \ll \tau_1$ ). Injection into  $|1\rangle$  only is independently controlled via a pumping rate  $P$ , defined as

$$P = \frac{J}{eN}, \quad (2.73)$$

where  $N$  is once again the total population of the three levels.

The equations of motion for the relevant density matrix elements are

$$\frac{d}{dt}\rho_{11} = \frac{i\Delta_0}{\hbar}(\rho_{21} - \rho_{21}^*) - \frac{\rho_{11}}{\tau_1} + P, \quad (2.74)$$

$$\frac{d}{dt}\rho_{22} = -\frac{i\Delta_0}{\hbar}(\rho_{21} - \rho_{21}^*) - \frac{\rho_{22}}{\tau_2}, \quad (2.75)$$

$$\frac{d}{dt}\rho_{33} = \frac{\rho_{11}}{\tau_1} + \frac{\rho_{22}}{\tau_2} - P, \quad (2.76)$$

$$\frac{d}{dt}\rho_{21} = \frac{i\Delta_0}{2\hbar}(\rho_{11} - \rho_{22}) - \frac{\rho_{21}}{\tau_{||}}. \quad (2.77)$$

In the steady state, one can derive the following relations for state populations:

$$\rho_{22} = \rho_{11} \frac{(\frac{\Delta_0}{\hbar})^2 \tau_{||} \tau_2 / 2}{1 + (\frac{E_{21}}{\hbar})^2 \tau_{||}^2 + (\frac{\Delta_0}{\hbar})^2 \tau_{||} \tau_2 / 2}, \quad (2.78)$$

where the pumping current is

$$P = \frac{\rho_{22}}{\tau_2} + \frac{\rho_{11}}{\tau_1}. \quad (2.79)$$

For this system, the current density is now an adjustable parameter rather than a result. However, Eq. (2.78) is equally valid for the two-level system. In fact, in the steady state, there is no difference between the description of transport in the two-level and three-level scenarios, except for the relative differences in populations. This is expected, since current continuity requires the re-injection of electrons from  $|2\rangle$  back into  $|1\rangle$ , whether or not they flow through a reservoir first.

The real value of including a third reservoir level is for analysis of dynamic situations and dephasing, so that the pumping can be independently turned off. Unlike for the two-level system, it is impossible to reduce the equations of motion to a damped harmonic oscillator for  $(\rho_{11} - \rho_{22})$  for arbitrary  $\tau_1$  and  $\tau_2$ . The expression

$$\begin{aligned} \frac{d^2}{dt^2}(\rho_{11} - \rho_{22}) &= -\frac{\Delta_0^2}{\hbar^2}(\rho_{11} - \rho_{22}) - \frac{1}{\tau_{||}} \left( \frac{\rho_{11}}{\tau_1} - \frac{\rho_{22}}{\tau_2} \right) \\ &\quad - \frac{1}{\tau_{||}} \frac{d}{dt}(\rho_{11} - \rho_{22}) - \frac{d}{dt} \left( \frac{\rho_{11}}{\tau_1} - \frac{\rho_{22}}{\tau_2} \right) + \frac{P}{\tau_{||}}. \end{aligned} \quad (2.80)$$

is as close as we can get. However, if driving is turned off ( $P = 0$ ), and  $\tau_1 = \tau_2 \equiv \tau$ , then Eq. (2.68) can be recovered if its driving term  $(\tau\tau_{||})^{-1}$  is likewise neglected. While this special three-level case is not particularly physically applicable in and of itself, the equivalence of the two equations (2.68) and (2.80) gives valuable insight into the the dephasing term  $\tau_{||}$  defined in Eq. (2.61) for the two-level system. Namely, the scattering into  $|1\rangle$  with time  $\tau$  that occurs in a two-level system is dynamically equivalent to a scattering out of  $|1\rangle$  with time  $\tau_1$  that occurs in the three-level system.

In other words, when counting population relaxation events for the two-level system,  $\tau$  should be counted twice, once for each state involved. Hence, the equality in the right hand side of (2.61) is justified, which is critical in the interpretation of Eq. (2.70). However, it should be kept in mind that the two-level system is a special case, where relaxation is inextricably tied to the pumping, and in most cases the use of a generalized three-level system is more relevant for dynamic analysis.

Below, I provide further quantum justification for using this definition of  $\tau_{\parallel}$  in the density matrix formalism used to describe resonant tunneling. The treatment is similar to that in Appendix 6A of Mukamel [105]. Consider first the difference between the density matrices for pure and mixed quantum states, for which

$$\rho_{ij}^{(pure)} = c_i c_j^*, \quad \text{and} \quad \rho_{ij}^{(mixed)} = \langle c_i c_j^* \rangle, \quad (2.81)$$

where the mixed state represents an ensemble of many pure states, which may or may not be in phase. For a pure state we have

$$|\rho_{ij}|^2 = \rho_{ii}\rho_{jj}, \quad (2.82)$$

and for a mixed state we can obtain the relation

$$|\rho_{ij}|^2 \leq \rho_{ii}\rho_{jj}, \quad (2.83)$$

which is equivalent to the Schwartz inequality [105].

Let us consider the decay of a mixed state using the model described above at resonance ( $E_{21} = 0$ ) with no pumping ( $P = 0$ ). Assume that it is coherently prepared as a pure state, so that condition (2.82) initially holds at time  $t = 0$ . For the condition (2.83) to be true we require that

$$\frac{d|\rho_{21}|^2}{dt} \Big|_{t=0} \leq \frac{d(\rho_{22}\rho_{11})}{dt} \Big|_{t=0} \quad (2.84)$$

$$(\rho_{21}\dot{\rho}_{21}^* + \dot{\rho}_{21}\rho_{21}^*) \Big|_{t=0} \leq (\rho_{11}\dot{\rho}_{22} + \dot{\rho}_{11}\rho_{22}) \Big|_{t=0}. \quad (2.85)$$

Following some algebra we then obtain the expression

$$\frac{2|\rho_{21}|^2}{\tau_{\parallel}} \geq \rho_{11}\rho_{22} \left( \frac{1}{\tau_1} + \frac{1}{\tau_2} \right) \quad (2.86)$$

Substituting (2.82) for an initial pure state, we can obtain

$$\frac{1}{\tau_{\parallel}} \geq \frac{1}{2\tau_1} + \frac{1}{2\tau_2}, \quad (2.87)$$

where the greater-than condition allows for the addition of pure dephasing  $T_2^*$ . This result is what we expected, and is consistent with the classical analogy given in Appendix A. We can conclude the following. First, the presence of  $\tau_1$  and  $\tau_2$  in Eq. (2.87) clearly indicates that the dephasing time  $\tau_{\parallel}$  must account for dephasing due to population relaxation. Generally speaking, if we allow  $\rho_{11}$  and/or  $\rho_{22}$  to decay without letting  $\rho_{21}$  also decay, we will violate the inequality of Eq. (2.83). Second, the factors of two in the denominators of the right-hand side of (2.87) correspond to what we expected, i.e. the lifetime broadening contributes half much as pure dephasing. Thus we justify the definition of  $\tau_{\parallel}$  given in Eq. 2.61.

### **Analogy with two-level optically coupled system**

As mentioned above, the density matrix model for resonant tunneling is directly analogous to the optical Bloch equations for a two-level system in the presence of an electromagnetic field. The schematic for such comparable optical systems are shown in Figs. 2-8(b) and 2-11(b). For optical coupling, the two states are electronic eigenstates separated by  $\hbar\omega_0$  which are coupled by the dipole interaction  $-\mu E_0 e^{-i\omega t}$ , instead of the tight-binding coupling  $-\Delta_0/2$ . In both cases, the interaction causes oscillation of populations  $\rho_{22}$  and  $\rho_{11}$  between the two states. In the optical case though, the average dipole moment is of particular interest, which is given by  $\langle \mu \rangle = \mu(\rho_{21} + \rho_{21}^*)$ , and is used to obtain the complex susceptibility. The rapidly varying component  $e^{-i\omega_0 t}$  of the density matrix terms is usually removed, leaving terms that depend only on the detuning from resonance ( $\omega - \omega_0$ ). However, the definitions of

dephasing are equivalent, and the homogeneously broadened Lorentzian lineshape for a radiative transition is equivalent to the Lorentzian current lineshape associated with the detuning  $E_{21}$  from tunneling resonance. The following analogies can be made for these two two-level systems:

$$\begin{aligned}
-\frac{\Delta_0}{2} &\rightarrow -\mu E_0 e^{-i\omega t} \\
\frac{E_{21}}{\hbar} &\rightarrow \omega - \omega_0 \\
\tau &\rightarrow \tau \\
\tau_{\parallel} &\rightarrow T_2 \\
\frac{\Delta_0^2}{\hbar^2} \tau \tau_{\parallel} > 1 \text{ (coherent)} &\rightarrow 4\Omega^2 \tau T_2 > 1 \text{ (saturation)} \\
\frac{\Delta_0^2}{\hbar^2} \tau \tau_{\parallel} < 1 \text{ (incoherent)} &\rightarrow 4\Omega^2 \tau T_2 < 1 \text{ (small signal)} \\
\Delta E_{21} = \frac{2\hbar}{\tau_{\parallel}} \left[ 1 + \left( \frac{\Delta_0}{\hbar} \right)^2 \tau \tau_{\parallel} \right]^{1/2} &\rightarrow \Delta\nu = \frac{1}{\pi \tau_{\parallel}} \left[ 1 + 4\Omega^2 \tau T_2 \right]^{1/2}
\end{aligned}$$

where the “precession frequency”  $\Omega$  is defined by  $\Omega \equiv \mu E_0 / 2\hbar$ . As implied above, the coherent regime in resonant tunneling is analogous to the saturation regime for the optical system. In both cases, the coupling is strong enough to significantly affect the population difference and the system tends toward  $\rho_{11} \approx \rho_{22}$ . This is accompanied by saturation broadening of the linewidth  $\Delta\nu$ , which is analogous to the change in transport linewidth between coherent and incoherent tunneling. The treatment of a two-level optical system is standard, and a full derivation system is given in Yariv [77]. A more sophisticated derivation that encompasses relaxation and dephasing superoperators is given in Mukamel [105].

Aside from offering a pleasing symmetry, the analogy between these two systems suggests a method for obtaining transport information from optical measurements. The pure dephasing  $T_2^*$  is critical for calculating tunneling currents, but is difficult to obtain through either calculation or direct measurement. Measurement of homogeneously broadened spontaneous emission linewidths in our structures offers insight

into dephasing via the relation

$$\Delta\nu = \frac{1}{\pi T_2} = \frac{1}{\pi} \left( \frac{1}{2\tau_1} + \frac{1}{2\tau_2} + \frac{1}{T_2^*} \right). \quad (2.88)$$

Although the dephasing time associated with transitions between two radiative states may be different than the dephasing between two resonant tunneling states, it can be expected to be of similar magnitude. Thus the measured electroluminescence linewidth of 1 THz from FL175C was used to obtain the time  $T_2^* = 0.33$  ps that was used in Fig. 2-9.

## 2.6 Summary

In this chapter, I presented the basic theoretical tools for modeling and designing terahertz quantum cascade lasers. Essentially, transport and gain is modeled by finding the bound states of the one-dimensional Schrödinger equation and calculating the associated intersubband radiative and nonradiative transition rates. Scattering is dominated by LO-phonon scattering for subband separations larger than  $E_{LO}$  and by a combination of electron-electron scattering and thermally activated LO-phonon scattering for subband separations smaller than  $E_{LO}$ . The effect of dephasing on resonant tunneling transport is considered phenomenologically at certain transport bottlenecks.



# Chapter 3

## Heterostructures and optical-phonons

### 3.1 Introduction

Inter- and intra-subband electron relaxation via emission of optical phonons is a critical process in terahertz as well as in mid-infrared QCLs. Qualitative as well as quantitative knowledge of scattering rates is especially important in structures where direct LO-phonon scattering is used for depopulation of the lower radiative state, such as for the FL175C laser described in Chapter 7.2. Traditionally in QCL design, the effect of the heterostructure on the phonon spectrum has been ignored, and scattering rates were calculated using the bulk phonon spectra of the well material. The validity of this assumption depends sensitively on the details of the structure and the material system. In this chapter, I will examine the effect of the complex optical-phonon spectra on intersubband scattering and the associated design rules for terahertz QCLs that use direct LO-phonon depopulation. I will focus on the GaAs/ $\text{Al}_x\text{Ga}_{1-x}\text{As}$  material system.

One method for obtaining a population inversion between subbands is to use resonant LO-phonon scattering in a three- (or more) level system. This type of depopulation is usually accomplished by placing a third subband (or group of subbands) one LO-phonon energy below the lower radiative state. The conduction band diagrams

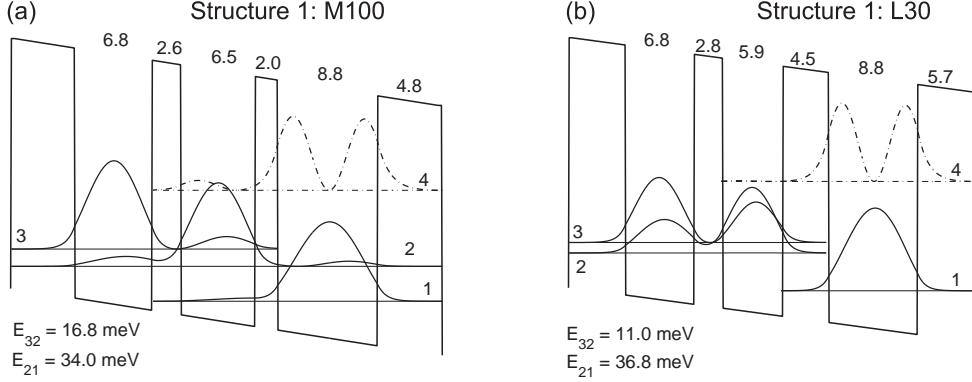


Figure 3-1: Band structure and wavefunction magnitude squared for a three-well module for (a) structure 1: M100 and (b) structure 2: L30, grown in GaAs/Al<sub>0.3</sub>Ga<sub>0.7</sub>As. The numbers above the layers represent layer thicknesses in nm.

for two typical GaAs/Al<sub>0.3</sub>Ga<sub>0.7</sub>As triple well module devices are shown in Fig. 3-1<sup>1</sup>. Measurements and analysis of the electronic properties of structure 1, known as M100, are reported in Refs. [59, 60] and in Sec. 6.2.1. Structure 2 is known as L30, and is discussed in Sec. 6.3.1 and Ref. [63]. Many of these modules are cascade connected, and when an appropriate electric field is applied, electrons are injected from the previous module into the excited state  $n = 3$  via resonant tunneling. After transit through the module, the electrons collect in state  $n = 1$  and are injected into the excited state of the next module. The structure is designed so that THz emission occurs between levels  $n = 3$  and 2, ( $E_{32} \approx 10 - 20$  meV), and rapid LO-phonon scattering occurs between levels  $n = 2$  and 1 ( $E_{21} \approx E_{\text{LO}}$ ). Since  $E_{32} < E_{\text{LO}}$  ( $E_{\text{LO}} = 36.25$  meV in GaAs), LO-phonon scattering is nominally suppressed if the electrons are cold, and one can obtain a small lifetime ratio  $\tau_{21}/\tau_{32}$ . Ideally, this leads to a population inversion  $n_3 - n_2 > 0$ . While lasing was not observed in either of these structures, they nonetheless serve as appropriate model systems to examine the issue of complex phonon modes, and many of our conclusions can be applied to other devices.

<sup>1</sup>Note that these band structures were calculated using a conduction band offset of 65%, and not 80% as is used for some other calculations. While this uncertainty might affect some of the magnitudes of the calculated scattering rates in this chapter, the qualitative trends and conclusions drawn are still valid.

These designs have set  $E_{21} \approx 36$  meV to take advantage of the bulk GaAs phonon resonance. However, for many situations, especially those where subband separation is close to the optical phonon energy, the effect of the heterostructure on the phonon spectra cannot be ignored [106, 107, 108]. The overall result is that scattering takes place at energies clustered around the optical phonon energies of both the well and barrier material. For GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As structures, this leads to scattering by “AlAs-like” modes that occurs around 47 meV (near  $E_{\text{LO,AlAs}} \approx 50$  meV), in addition to that which occurs around  $E_{\text{LO,GaAs}} \approx 36$  meV. It has been suggested that raising  $E_{21}$  to 47 meV would increase the population inversion in two ways. First, all of the phonon modes would be energetically available for scattering, giving a faster depopulation time  $\tau_{21}$  [108]. Second, raising  $E_{21}$  would push  $E_{31}$  away from resonance with the higher energy phonon modes, and thus decrease parasitic scattering from  $n = 3$  to 1 [109].

In this chapter, I examine the issue of the optimal energy separation  $E_{21}$  for achieving population inversion in three-level structures. The total phonon scattering rates are calculated for two model structures that have been investigated experimentally. The devices M100 and L30 are similar, except for two principal differences. M100 is designed so that states  $n = 3$  and 2 are beyond the anticrossing point, so that the optical transition is spatially diagonal, and has a dipole matrix element of  $z_{32} = 30.7$  Å ( $f_{32} = 0.28$ ). L30 is designed around an anticrossed transition that is primarily vertical in nature, which yields a stronger transition strength ( $z_{32} = 44.0$  Å,  $f_{32} = 0.37$ ). While the vertical transition leads to shorter radiative lifetimes, and consequently higher gain, the nonradiative scattering processes are stronger as well. However, vertical transition structures also have the advantage of a narrower spontaneous emission linewidth due to reduced interface roughness scattering. The second major difference is that L30 has much thicker barriers than does M100, which greatly reduces intersubband scattering due to optical phonons into state  $n = 1$ . This makes achieving a population inversion more difficult.

When energetically permitted, scattering due to electron-LO-phonon interaction is generally the dominant intersubband relaxation mechanism. Hence, the scatter-

Table 3.1: Optical phonon scattering times for electrons with sufficient kinetic energy to scatter from  $n = 3 \rightarrow 2$  via GaAs-like and AlAs-like phonons.

	$\tau_{32,\text{LO(G)}}^{(\text{hot})}$	$\tau_{32,\text{LO(A)}}^{(\text{hot})}$
M100 ( $E_{32} = 16.8$ meV)	0.4 ps	1.1 ps
L30 ( $E_{32} = 11.0$ meV)	0.8 ps	2.6 ps

ing times  $\tau_{21}$  and  $\tau_{31}$  are given by LO-phonon scattering. Since the radiative relaxation time for THz intersubband transitions is on the order of microseconds, electron transport from  $n = 3$  to 2 is dominated by nonradiative transitions. Of these, at low temperatures electron-electron scattering is the principal  $3 \rightarrow 2$  relaxation mechanism, since acoustic phonon scattering times are generally greater than 100 ps [58]. Electron-electron scattering rates scale with the upper subband population and thereby place an upper limit on the populations achievable in the upper state. Calculations and experiments indicate that electron-electron scattering times for THz transitions might range anywhere from 1–50 ps for the range of densities used ( $\sim 10^{10}$  cm $^{-2}$ ) (see Sec. 2.4.2). When a current is applied during operation, the in-plane electron temperature  $T_e$  will rise above the lattice temperature, and electrons can gain enough energy to emit an LO-phonon (i.e.  $\hbar^2 k^2 / 2m^* + E_{32} > \hbar\omega_{\text{LO}}$ ). The scattering of hot electrons by LO-phonons can be approximately described by the rate

$$W_{32,\text{LO}} \cong W_{32,\text{LO(G)}}^{(\text{hot})} \exp\left(\frac{E_{32} - E_{\text{LO(G)}}}{k_B T_e}\right) + W_{32,\text{LO(A)}}^{(\text{hot})} \exp\left(\frac{E_{32} - E_{\text{LO(A)}}}{k_B T_e}\right), \quad (3.1)$$

where  $W_{32,\text{LO(G)}}^{(\text{hot})}$  ( $W_{32,\text{LO(A)}}^{(\text{hot})}$ ) is the scattering rate of electrons with sufficient in-plane kinetic energy to scatter via emission of a GaAs-like (AlAs-like) LO-phonon, seen in Table 3.1. Thus the total scattering time from  $n = 3$  to 2 can be written

$$\tau_{32}^{-1} = \tau_{32,e-e}^{-1} + \tau_{32,\text{LO}}^{-1}. \quad (3.2)$$

Since both rates depend sensitively on the injection level  $J/e$ ,  $\tau_{32}$  is highly uncertain.

Solution of steady-state rate equations for this three-level system with unity injection efficiency yields the following expression for the population difference:

$$\Delta n_{32} = n_3 - n_2 = n_3 \left(1 - \frac{\tau_{21}}{\tau_{32}}\right) = \frac{J}{e} \tau_3 \left(1 - \frac{\tau_{21}}{\tau_{32}}\right), \quad (3.3)$$

where  $J$  is the injection current density. The lifetime of the excited state  $\tau_3$  is given by  $\tau_3^{-1} = \tau_{32}^{-1} + \tau_{31}^{-1}$ . Clearly the existence of a population inversion requires only  $\tau_{32} > \tau_{21}$ , but the magnitude of  $\Delta n_{32}$  depends on  $\tau_3$  as well. Parasitic scattering via LO-phonons from  $n = 3$  to 1 is undesirable, as it decreases  $\tau_3$  and therefore  $\Delta n_{32}$ . In this analysis based on a three-level system, the maximum population inversion is achieved when  $\tau_{31} \gg \tau_{32} \gg \tau_{21}$ .

## 3.2 Complex phonon calculation

There are several macroscopic models used to account for the effects of heterostructures on the phonon spectra. In this treatment, the dielectric continuum model, also known as the “slab model”, is used, and the phonon modes are described by the electrostatic potential that results from the polarization field created by atomic displacements in a polar semiconductor [110, 111, 112, 113]. This model satisfies the electrostatic boundary conditions, but not the mechanical boundary conditions. Although not used here, there are two other commonly used macroscopic phonon models that deserve mention. The guided-mode model, also known as the “mechanical model”, satisfies the mechanical but not the electrostatic boundary conditions, but is generally considered inappropriate for use [114]. The third model is known as the Huang-Zhu model, and is essentially the dielectric continuum model modified to take account of phonon dispersion and both boundary conditions [115]. In fact the interface modes for both the dielectric continuum and the Huang-Zhu models are identical, and only the expressions for the confined modes differ.

The dielectric continuum model is used here because it is the most straightforward to implement, and gives good agreement with microscopic calculations of

electron-phonon intersubband scattering rates [86]. The modes form an orthogonal and complete set that reduce in the bulk limits to the Frölich description. The slightly improved accuracy of the Huang-Zhu model appears only to be necessary when calculating energy loss rates for electron cooling calculations when taking account of hot-phonon distributions [116]. In these cases calculations have shown that use of the dielectric continuum model slightly overestimates the hot-phonon effect [117].

In the dielectric continuum model, each material layer  $i$  is described by a dielectric function  $\epsilon_i(\omega)$  as given by the Lyddane-Sachs-Teller relations for a binary semiconductor:

$$\epsilon_i(\omega) = \epsilon_i(\infty) \frac{(\omega^2 - \omega_{\text{LO}}^2)}{(\omega^2 - \omega_{\text{TO}}^2)}, \quad (3.4)$$

and ternary semiconductor [112]:

$$\epsilon_i(\omega) = \epsilon_i(\infty) \frac{(\omega^2 - \omega_{\text{LOA}}^2)(\omega^2 - \omega_{\text{LOB}}^2)}{(\omega^2 - \omega_{\text{TOA}}^2)(\omega^2 - \omega_{\text{TOB}}^2)}, \quad (3.5)$$

where  $\omega_{\text{LO}}$  and  $\omega_{\text{TO}}$  are the longitudinal and transverse optical phonon frequencies in the binary materials  $A$  and  $B$  that compose a ternary semiconductor. In regions absent of free charges, the potential must satisfy  $\epsilon \nabla^2 \phi(\mathbf{r}) = 0$ . In a multiple quantum well system, phonons are unconfined in the  $x-y$  plane, and hence the phonon potential can be written as

$$\Phi(\mathbf{r}) = \sum_{\mathbf{q}} e^{-i\mathbf{q}\cdot\boldsymbol{\rho}} \phi(q, z), \quad (3.6)$$

where  $\mathbf{q}$  and  $\boldsymbol{\rho}$  are the in-plane wave and position vectors. The function  $\phi(q, z)$  must satisfy the relation

$$\epsilon(\omega) \left( \frac{\partial^2}{\partial z^2} - q^2 \right) \phi(q, z) = 0, \quad (3.7)$$

where  $q = |\mathbf{q}|$  is the magnitude of the in-plane wavevector. Two types of solutions satisfy Eq. (3.7): interface (IF) modes, for which  $[(\partial^2/\partial z^2) - q^2] \phi(q, z) = 0$ , and confined modes, for which  $\epsilon(\omega) = 0$ .

For interface modes,  $\epsilon(\omega) \neq 0$ , and the modes have frequency  $\omega(q) \neq \omega_{\text{LO}}$ . The

solution for the potential is a linear combination of exponential terms, written as

$$\phi_i(q, z) = c_{i-} e^{-qz} + c_{i+} e^{+qz}, \quad (3.8)$$

where the coefficients  $c_{i-}$  and  $c_{i+}$  are chosen to match the electromagnetic boundary conditions ( $\phi$  and  $\epsilon(\omega)\partial\phi/\partial z$  are continuous). By requiring that the potentials decay in the outermost layers, and applying the boundary conditions at each interface for various  $q$ , we can calculate the dispersion  $\omega(q)$  for all IF modes. The coefficient amplitudes are also chosen to satisfy the generalized normalization condition for phonon mode strength [118]:

$$\frac{\hbar}{2\omega} = \sum_i \frac{\epsilon_0}{2\omega} \frac{\partial\epsilon_i(\omega)}{\partial\omega} \int dz \left( q^2 |\phi_i(q, z)|^2 + \left| \frac{\partial\phi_i(q, z)}{\partial z} \right|^2 \right). \quad (3.9)$$

This constraint is determined from the orthonormality and completeness conditions imposed on the phonon eigenfunctions as specified in the microscopic description [110, 111, 112]. The strength of the electron-phonon coupling is contained in this normalization, and it is equivalent to the Frölich interaction strength for bulk modes. Just as for bulk phonons, the total scattering rate due to all the IF and confined modes decreases rapidly with increasing  $q$ .

For confined modes,  $\epsilon(\omega) = 0$  and therefore  $\omega = \omega_{\text{LO}}$ , where  $\omega_{\text{LO}}$  is the bulk LO-phonon frequency in the layer of interest. Since  $\omega_{\text{LO}}$  changes at the heterointerfaces, the potential must go to zero there, and  $\phi(q, z)$  can be described in terms of sine wave modes. In each layer of thickness  $a_i$  that begins at position  $z_i$  we have

$$\begin{aligned} \phi_i(q, z) &= \left( \frac{\hbar}{\epsilon_0} \frac{1}{\partial\epsilon_i(\omega)/\partial\omega} \right)^{1/2} \left( \frac{1}{q^2 + \left( \frac{m\pi}{a_i} \right)^2} \right)^{1/2} \left( \frac{2}{a_i} \right)^{1/2} \sin \frac{m\pi}{a_i} (z - z_i), \\ m &= 1, 2, 3, \dots, \quad z_i < z < z_i + a_i. \end{aligned} \quad (3.10)$$

The amplitude for the confined modes is determined by the same normalization condition as for the IF modes (Eq. (3.9)). Halfspace modes are similar to confined modes, except they exist only in the outermost layers and are semi-infinite in extent. For

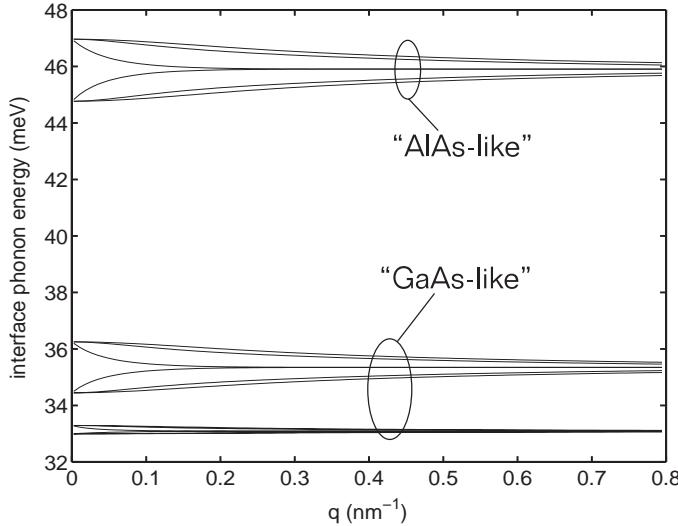


Figure 3-2: Dispersion relation of interface phonon modes for M100, where  $q$  is the in-plane wavevector.

the calculated structures in this chapter, the halfspace modes are  $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$  modes, and exist directly to the left and right of the triple well region. They do not contribute significantly to the total scattering rate, and will not be discussed further.

We used the transfer matrix approach described in Ref. [118] to account for the electromagnetic boundary conditions and obtain the mode potentials and dispersion relations for the interface phonon modes. The material constants used for the  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  system were also taken from the same source. The structures described in this chapter consist of six interfaces between binary and ternary materials ( $\text{GaAs}/\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ ), each of which contributes two GaAs-like modes and one AlAs-like mode, for a total of 18 IF modes. The dispersion relation for the IF modes for M100 are shown in Fig. 3-2. 12 of the modes are clustered around 33-36 meV and are considered “GaAs-like,” and 6 of the modes are clustered around 45-47 meV, and are considered “AlAs-like.” Some associated IF mode potentials  $\phi(z)$  (6 GaAs-like, 6 AlAs-like) are shown for  $q = 3 \times 10^{10} \text{ m}^{-1}$  in Fig. 3-3. The lowest six modes are not plotted as they have very small magnitudes, and hence contribute little to the total scattering rate. Examples of the two lowest order ( $m = 1, 2$ ) confined mode potentials for M100 are shown in Fig. 3-4. The modes for each layer are independent

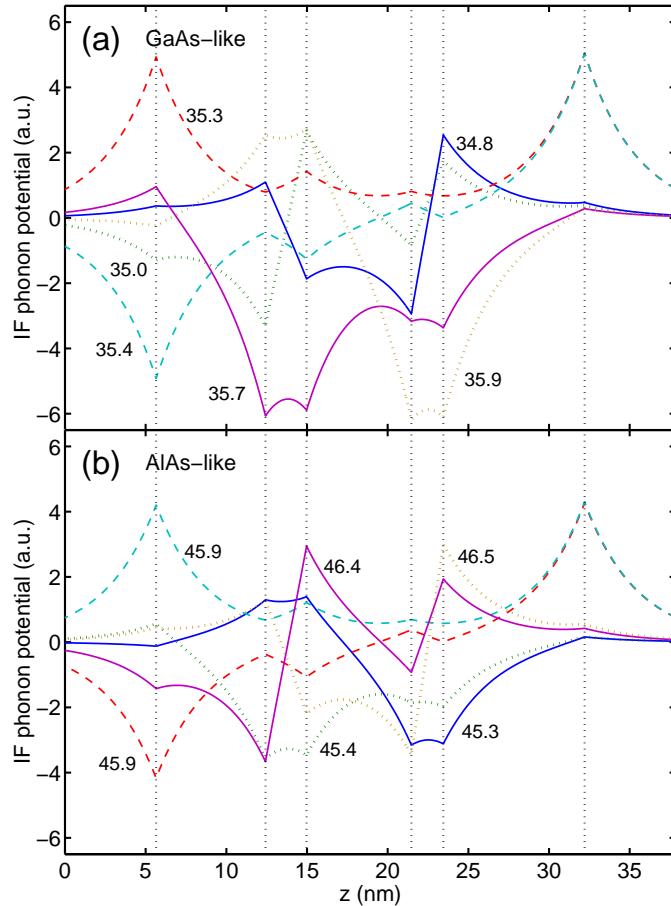


Figure 3-3: Interface phonon potential for (a) six “GaAs-like” modes and (b) six “AlAs-like” modes plotted for GaAs at  $q = 0.3 \text{ nm}^{-1}$ . Each mode is labeled with its energy in meV, and the vertical lines represent the heterointerfaces. The six modes of lowest energy contribute very little to scattering and are not plotted.

and are calculated separately, although they are plotted together for ease of viewing. The confined modes in the well have  $\hbar\omega_{\text{LO}} = 36.25 \text{ meV}$ , and the confined modes in the barrier have  $\hbar\omega_{\text{LO}} = 34.45 \text{ meV}$ , which is the bulk GaAs-like optical phonon energy in  $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ . AlAs-like confined modes also exist in the barrier regions at  $\omega_{\text{LO}} = 44.5 \text{ meV}$ , but contribute little and are not plotted. We included scattering due to the confined modes with indices  $m=1\text{--}30$ ; the contribution of higher order modes is negligible.

Once the IF and confined mode potentials have been found, the interaction Hamil-

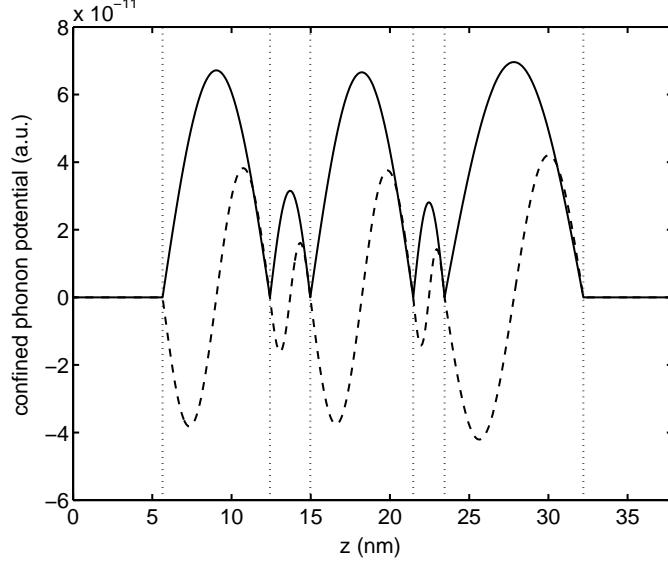


Figure 3-4: Confined “GaAs-like” modes for  $m = 1$  (solid) and  $m = 2$  (dashed) at  $q = 0.3 \text{ nm}^{-1}$ . The vertical lines represent heterointerfaces. Although the confined mode in each layer is calculated separately, they are plotted together for simplicity.

tonian can be written in terms of the potential [118]:

$$H_F = e \sum_{\mathbf{q}} e^{-i\mathbf{q}\cdot\boldsymbol{\rho}} \phi(q, z) (a_{-\mathbf{q}}^\dagger + a_{\mathbf{q}}), \quad (3.11)$$

where  $a_{-\mathbf{q}}^\dagger$  and  $a_{\mathbf{q}}$  are the phonon creation and annihilation operators. Phonon emission rates between two subbands  $n$  and  $n'$  with initial and final wavevectors  $k$  and  $k'$  are calculated using Fermi’s golden rule [113]

$$W_{n,n'}(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} (N_q + 1) |F(q)|^2 \delta(E_{n'} + E_{k'} + \hbar\omega - E_n - E_k), \quad (3.12)$$

where  $E_n$  and  $E_{n'}$  are the energies of the initial and final subbands, and  $E_k$  and  $E_{k'}$  are the kinetic energies of electrons in the initial and final subbands.  $N_q$  is the Bose-Einstein factor for phonons of frequency  $\omega(q)$ . In-plane momentum is conserved with the relation  $\mathbf{k} = \mathbf{k}' + \mathbf{q}$ . The form factor  $F(q)$  is given by

$$F(q) = -e \int dz \phi(q, z) \psi_n^*(z) \psi_{n'}(z), \quad (3.13)$$

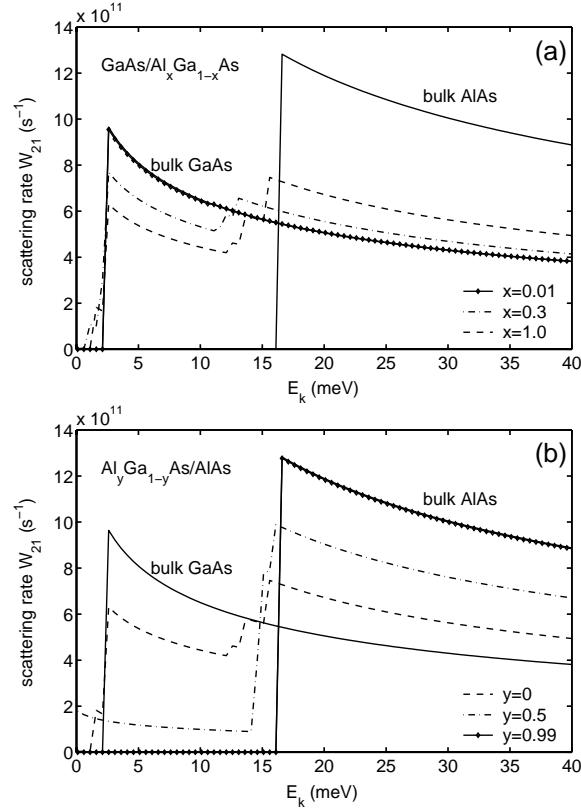


Figure 3-5: The scattering rate from  $n = 2$  to 1 for M100 is calculated as a function of initial in-plane electron energy  $E_k$  using both bulk and complex phonon spectra. Solid lines are scattering rates calculated using bulk GaAs and AlAs modes. Scored lines indicate calculations using the phonon spectra of (a) GaAs/ $\text{Al}_x\text{Ga}_{1-x}\text{As}$  and (b)  $\text{Al}_y\text{Ga}_{1-y}\text{As}/\text{AlAs}$  structures for various values of  $x$  and  $y$ . The line for  $x = 0.01$  is nearly indistinguishable from that of bulk GaAs, just as the line for  $y = 0.99$  is nearly indistinguishable from that of bulk AlAs.

where  $\psi_n(z)$  and  $\psi_{n'}(z)$  are  $z$ -component of the the initial and final electron envelope functions obtained by solving the one-dimensional Schrödinger equation. Integration over final states leads to

$$W_{n,n'}(\mathbf{k}) = \sum_{\mathbf{k}'} W_{n,n'}(\mathbf{k}, \mathbf{k}') = \frac{1}{2\pi} \frac{m}{\hbar^3} (N_q + 1) \int_0^{2\pi} d\theta |F(q)|^2, \quad (3.14)$$

where  $\theta$  is the angle between  $\mathbf{k}$  and  $\mathbf{k}'$ , and  $q^2 = k^2 + k'^2 - 2kk' \cos \theta$ . This expression must be evaluated for each allowed mode, confined as well as IF modes, and the total intersubband scattering rate is the sum of all these rates.

The effect of the complex mode structure is to redistribute some scattering power from lower energy GaAs modes into higher energy AlAs modes. Once electrons have sufficient energy such that all the IF and confined modes are accessible ( $> 50$  meV), the total scattering rate is bounded by the scattering rates calculated using bulk GaAs and bulk AlAs phonon modes. The relative contribution of the GaAs-like and AlAs-like modes depends sensitively on the barrier thickness and composition, as well as the electronic wavefunction overlap with the barrier regions. An example of this is shown in Fig. 3-5, where the scattering rate  $W_{21}(k)$  is plotted for M100 as a function of  $E_k$ , the initial in-plane kinetic energy. For this calculation,  $E_{21} = 34$  meV. The solid lines indicate the scattering rates for electrons calculated using either GaAs or AlAs bulk modes. The scored lines indicate rates calculated using the complex phonon spectra for heterostructures with various well and barrier compositions. Note that the same electronic wavefunctions were used for all calculations, and the change in heterostructure composition only affected the phonon calculations. Scattering rates were calculated using the phonon spectra of GaAs/ $\text{Al}_x\text{Ga}_{1-x}\text{As}$  structures (Fig. 3-5(a)) with  $x = 0.01, 0.3$ , and  $1.0$ , and for  $\text{Al}_y\text{Ga}_{1-y}\text{As}/\text{AlAs}$  structures (Fig. 3-5(b)) with  $y = 0.0, 0.5$ , and  $0.99$ . As the Al content of the barrier layers is increased, the scattering due to GaAs-like phonons decreases, and the scattering due to AlAs-like phonons increases. Note that in one limit the scattering rate for the GaAs/ $\text{Al}_{0.01}\text{Ga}_{0.99}\text{As}$  case (dotted line) is virtually identical to the bulk GaAs rate. In the other limit, the rate for the  $\text{Al}_{0.99}\text{Ga}_{0.01}\text{As}/\text{AlAs}$  structure is almost identical to the bulk AlAs rate. Since the rates due to bulk phonon interaction are computed independently, the convergence of the complex phonon results in these limiting cases provides excellent verification of our complex phonon calculations.

### 3.3 Optimized design of three-level structure

How does the inclusion of the complex phonon spectra change the design criteria for a THz three-level system? To investigate this, we computed the intersubband scattering rates using confined and interface modes for both the M100 and L30 structures. To

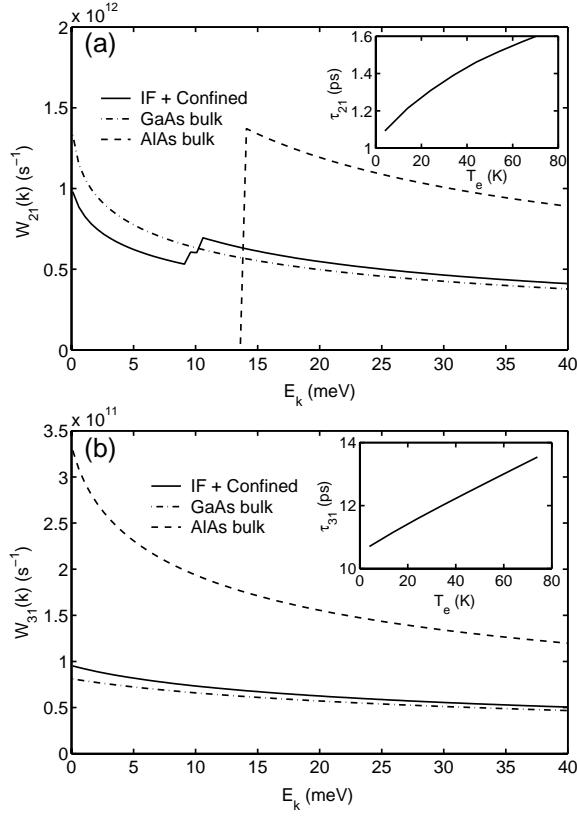


Figure 3-6: Scattering rate (a)  $W_{21}$  and (b)  $W_{31}$  for M100 as a function of  $E_k$  for  $E_{21} = 36.25$  meV ( $E_{31} = 53$  meV). Rates calculated using bulk modes are shown for comparison. The insets give scattering time (a)  $\tau_{21}$  and (b)  $\tau_{31}$  as a function of electron temperature for an upper subband population of  $10^{10} \text{ cm}^{-2}$ .

isolate the effects of subband energy separation on scattering, we adjusted  $E_{21}$  as an independent variable, and the other separations accordingly ( $E_{31} = E_{21} + E_{32}$ ). The electron wavefunctions and layer thicknesses were kept unchanged. While this is of course unphysical, it prevents contamination of the results with changes due to differences in wavefunction overlap that would occur if different structures were simulated. We believe that it does not detract from the validity of the results.

If  $E_{21}$  is set to 36.25 meV, the AlAs-like modes are inaccessible, and scattering from  $n = 2$  to 1 is potentially reduced. Furthermore, for  $E_{32} \approx 10 - 15$  meV, the transition from  $n = 3$  to 1 is placed in approximately resonance with the “AlAs-like” modes causing faster scattering and reducing  $\tau_{31}$ . The scattering rates  $W_{21}$  and  $W_{31}$  for M100 are plotted versus the initial electron kinetic energy  $E_k$  in Fig. 3-6 (a)

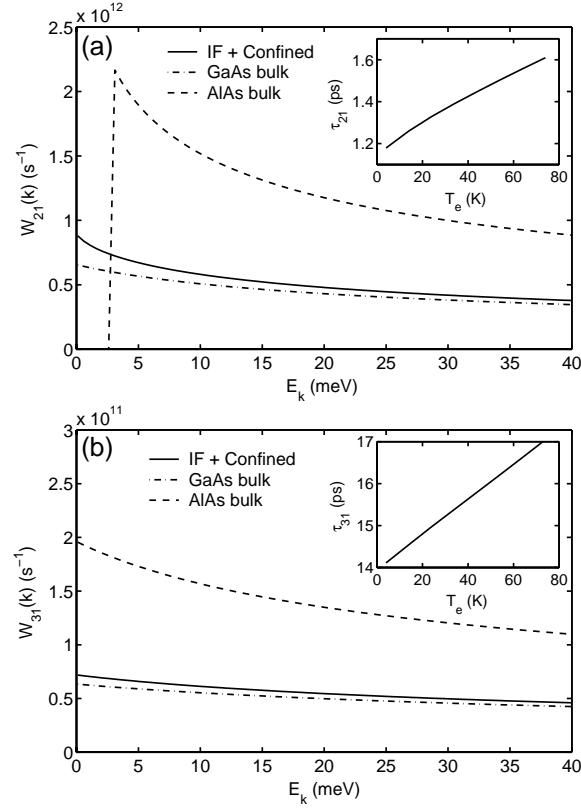


Figure 3-7: Scattering rate (a)  $W_{21}$  and (b)  $W_{31}$  for M100 as a function of  $E_k$  for  $E_{21} = 47.0$  meV ( $E_{31} = 64$  meV). Rates calculated using bulk modes are shown for comparison. The insets give scattering time (a)  $\tau_{21}$  and (b)  $\tau_{31}$  as a function of electron temperature for an upper subband population of  $10^{10} \text{ cm}^{-2}$ .

and (b) respectively. For  $W_{21}$ , note that the AlAs-like modes are not activated until  $E_k > 10$  meV. The insets display the average scattering time for an electron in the initial subband with temperature  $T_e$ . As the electron temperature rises, electrons populate higher  $E_k$  with a corresponding reduction in scattering. As a result, the average scattering time increases with temperature.

The alternative solution is to set  $E_{21} = 47$  meV so that all the phonon modes are accessible for electrons with  $k = 0$ . This has the additional advantage of moving  $E_{31} = 64$  meV away from any phonon resonances. Since the scattering rate decreases with the magnitude  $q$  of the emitted phonon, this greater energy separation can help to reduce  $3 \rightarrow 1$  scattering. To examine this scenario, the scattering rates for M100 were calculated with  $E_{21} = 47$  meV and  $E_{31} = 64$  meV, without changing

Table 3.2: Average intersubband scattering rates for M100 at T=10 K for different subband separations.

M100	$E_{21} = 36 \text{ meV}$ ( $E_{31} = 53 \text{ meV}$ )	$E_{21} = 47 \text{ meV}$ ( $E_{31} = 64 \text{ meV}$ )
$\tau_{21}$	1.2 ps	1.2 ps
$\tau_{31}$	11.5 ps	14.3 ps

the electronic wavefunctions. The scattering rates  $W_{21}$  and  $W_{31}$  for this scenario are shown in Fig. 3-7(a) and (b) respectively. The average scattering rates for M100 with the two different energy level separations are listed in Table 3.2.

Surprisingly, comparison of the calculated phonon scattering rates for the two scenarios reveals no significant change in  $\tau_{21}$ . Although setting  $E_{21} = 47 \text{ meV}$  allows scattering via AlAs-like modes, the GaAs-like modes still account for most of the scattering, since our structure employs thin barriers with a moderate aluminum concentration ( $x = 0.3$ ). The increase in scattering due to the addition of the AlAs-like modes is balanced by the reduction in the strength of the GaAs-like modes, which are now off resonance. The dependence of the scattering rate on  $E_{21}$  for M100 can be seen in Fig. 3-8(a), where the peak scattering rate (at any  $E_k$ ) is plotted as a function of the subband separation. The maximum scattering rate due to GaAs bulk modes is shown for comparison. Two peaks of similar magnitude are seen at  $E_{21} = 36.5 \text{ meV}$  and  $47.0 \text{ meV}$ . Note that even at the lower resonant energy, the total IF+confined maximum scattering was less than that of the bulk, and at the upper resonant energy the maximum scattering is only slightly larger than the GaAs bulk, but is still much less than the maximum bulk AlAs scattering rate (not shown). This is in contrast to Ref. [113] where single IF modes had calculated scattering rates more than an order of magnitude greater than the bulk rate. No such dramatic increase in scattering was seen in our calculations. In fact, such a strong scattering is unphysical, since it violates the effective bounds placed on scattering rates by the bulk AlAs rates.

Therefore, for M100 the only benefit of raising  $E_{21}$  from 36 meV to 47 meV is the decrease in  $3 \rightarrow 1$  scattering, as seen in Fig. 3-6(b) and Fig. 3-7(b). At  $T_e = 10 \text{ K}$   $\tau_{31}$  changes from 11.5 to 14.3 ps, as  $E_{31}$  changes from 53 to 64 meV. In order to examine overall effect, the quantity  $\tau_3(1 - \tau_{21}/\tau_{32})$  is plotted in Fig. 3-9(a) for M100 for both

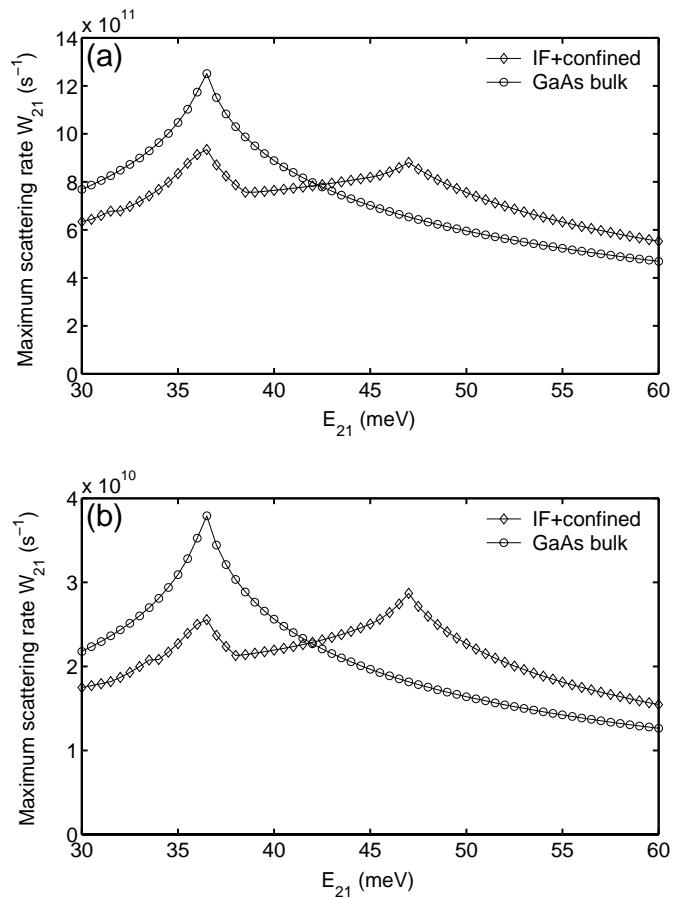


Figure 3-8: Maximum scattering rate versus subband separation  $E_{21}$  for (a) M100 and (b) L30. Maximum scattering rate calculated with GaAs bulk modes is present for comparison.

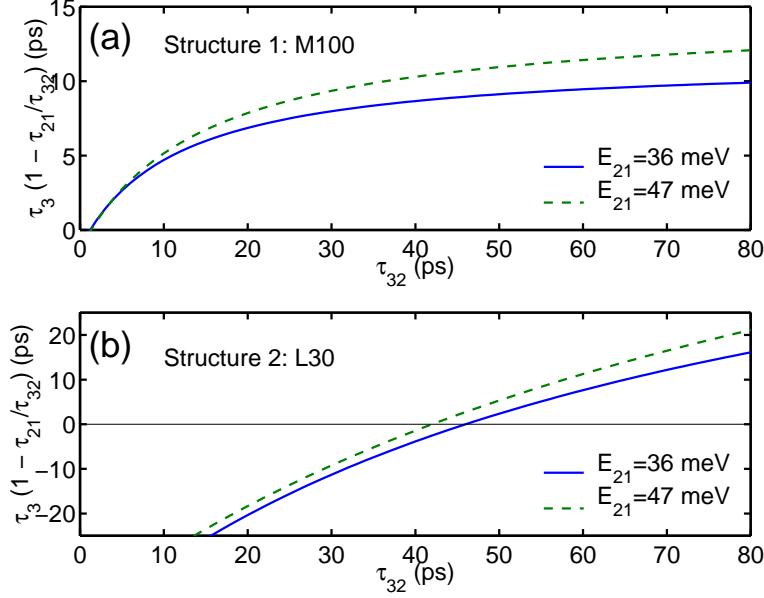


Figure 3-9: Plot of the quantity  $\tau_3(1 - \tau_{21}/\tau_{32})$ , which is proportional to the population inversion (Eq. (3.3)), versus the lifetime  $\tau_{32}$  for (a) M100 and (b) L30.

cases. This quantity is proportional to the population inversion, as seen in Eq. (3.3).  $\tau_{32}$  is used as an independent variable, since this time constant is highly dependent on the injection level (Eqs. (3.2), (3.1)) and is the least known. However, for practical devices, it is unlikely that  $\tau_{32}$  will ever be much greater than 30 ps at peak injection, and it will likely be much less. For  $\tau_{32} < 10$  ps, there is little difference between the two cases, but for longer  $\tau_{32}$ , there is a noticeable improvement in population inversion for  $E_{21} = 47$  meV.

However, even this marginal advantage of raising  $E_{21}$  is not necessarily so straightforward. In M100, even when  $E_{21} = 47$  meV, 74% of the  $2 \rightarrow 1$  scattering takes place through GaAs-like modes. As a result, most electrons from subband 2, and all electrons from subband 3 are injected into subband 1 with an extra  $\sim 11$  meV of energy. This extra energy will further heat the electron gas, and potentially decrease  $\tau_{32}$  due to optical phonon scattering of hot electrons. Examination of Fig. 3-9 reveals that reducing  $\tau_{32}$  could undo any advantage conferred by raising  $E_{21}$  from 36 to 47 meV. Even when the devices are tested with a bath temperature down to 4.2 K, estimates based on cooling rates suggest that the temperature of the electron gas may be any-

Table 3.3: Average intersubband scattering rates for L30 at T=10 K for different subband separations.

L30	$E_{21} = 36 \text{ meV}$ ( $E_{31} = 47 \text{ meV}$ )	$E_{21} = 47 \text{ meV}$ ( $E_{31} = 58 \text{ meV}$ )
$\tau_{21}$	46 ps	42 ps
$\tau_{31}$	70 ps	101 ps

where from 20–70 K higher, depending on the device structure and applied current density [83]. It is difficult to be more precise about the effects of increasing  $E_{21}$  without knowing the exact electron temperature of operational devices, however extra energy dissipation should generally be avoided unless necessary.

Examination of L30 reveals similar results as M100, with some small differences. In general, L30 has thicker barriers and hence the optical phonon scattering times are much longer. Also, greater overlap between states  $n = 3$  and 2 results in a smaller  $\tau_{31}/\tau_{21}$  ratio. The maximum scattering rate from  $2 \rightarrow 1$  versus  $E_{21}$  for L30 is plotted in Fig. 3-8(b). In this case slightly faster scattering is obtained by setting  $E_{21} = 47 \text{ meV}$ , since the Al<sub>0.3</sub>Ga<sub>0.7</sub>As barriers occupy a larger fraction of the structure. Hence the AlAs scattering contribution is greater. In L30,  $\tau_{31}$  also increases from 72 ps to 102 ps as  $E_{31}$  is moved from 47 meV to 58 meV. A summary of the average scattering times for L30 is presented in Table 3.3. Due to the slow depopulation time  $\tau_{21}$ , L30 is poorly suited to obtain population inversion, as can be seen from examination of Fig. 3-9(b). As in M100, the increase of  $E_{21}$  from 36 to 47 meV increases the population inversion for a given  $\tau_{32}$ . However the same caveat applies as in M100; the increased dissipated energy may heat the electron gas sufficiently to reduce  $\tau_{32}$  and offset any improvement. The primary advantage of L30 is the vertical optical transition, which offers a large oscillator strength. The vertical transition structure also displays narrower emission linewidths compared to the diagonal (interwell) structures due to the reduction in interface roughness scattering (see Sec. 6.3.1).

### 3.4 Conclusion

Awareness of the complex phonon spectra is important for the design of THz intersubband devices that use phonon-assisted depopulation, since the subband energy separations are close to the various optical phonon energies. When electrons have sufficient energy such that all the modes are available for scattering, the scattering rate is bounded by the scattering rates calculated using the bulk modes for the well and barrier materials, as has been frequently reported [106, 86, 119]. In our GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As structures, the relative contributions of GaAs-like and AlAs-like modes are determined by the barrier composition and thicknesses, as well as the electron wavefunction overlap with the barriers. It has been suggested that raising the subband separation to  $E_{21} \geq 47$  meV would benefit THz intersubband laser operation in two ways: reduced  $\tau_{21}$  and longer  $\tau_{31}$ . This proposal was examined by calculating these scattering rates for two structures that were previously experimentally tested, which were typical of our early designs. Due to the small energies involved in THz emission, electrically pumped THz intersubband emitters usually have wide wells and relatively thin and low barriers, which limits the influence of the stronger AlAs-like electron-phonon interaction. In the structures we examined, raising  $E_{21}$  to 47 meV provided little if any reduction in  $\tau_{21}$ , since the additional scattering from the AlAs-like modes was balanced by the reduced contribution of the off-resonance GaAs-like modes. As anticipated,  $\tau_{31}$  was favorably lengthened, as the energy separation  $E_{31}$  was pushed away from the AlAs-like resonance. However, raising  $E_{21}$  from 36 to 47 meV would result in the transfer of 11 meV of extra energy per carrier into heating the electron gas. Depending on the initial electron temperature, it is possible that the raised temperature will reduce  $\tau_{32}$  sufficiently to offset any benefit from an increased  $\tau_{31}$ . We did not observe the strong resonant scattering enhancement calculated for narrow quantum wells [113], nor did we observe the strong resonant enhancement in  $\tau_{21}$  as calculated for a similar three-level structure [109]. In fact, those calculated should be viewed skeptically, since they violate the effective sum rule in which the bulk AlAs scattering rate is an upper limit. When designing three-level intersubband

THz emitters, the choice of subband separations must be evaluated on a case-by-case basis, taking into account the results of the complex phonon scattering. For example, structures with large  $E_{32}$  ( $> 15$  meV) will have less to gain by setting  $E_{21} = 47$  meV, since the  $3 \rightarrow 1$  transition is already off resonance with the AlAs modes, and such structures are more susceptible to reduction of  $\tau_{32}$  by electron heating. Also, although not emphasized here, care must be taken when raising  $E_{21}$  so that subband  $n = 3$  is not brought into resonance with level  $n = 4$ , thus reducing  $\tau_3$ .

In conclusion, for the structures examined above, there does not appear to be a strong advantage in setting  $E_{21} \sim 47$  meV, rather than  $E_{21} \sim 36$  meV. Furthermore, for our terahertz devices that successfully lased (see Chap. 7), 15% Al barriers were used, which would result in an even weaker contribution of AlAs-like modes. As a result, most phonon-depopulation devices reported in this thesis are designed with an energy separation corresponding to the bulk GaAs phonon energy.

For optically pumped intersubband THz lasers however, because of thicker barriers with high Al concentrations ( $x \geq 0.4$ ) and relatively narrow wells, the influence of the AlAs-like modes will be stronger; therefore the benefit of setting  $E_{21} \sim 47$  meV may be more significant, as suggested in Ref. [108]. The development of an optically pumped intersubband Raman laser may have provided evidence suggesting important contributions of the AlAs-like interface modes to the intersubband scattering process [120]. Another example is the necessity of using complex phonon spectra to explain the thermal behavior of a GaAs/AlAs quantum cascade laser [121].

# Chapter 4

## Waveguide design

### 4.1 Introduction

A laser is composed of two principal components: an amplifier (gain medium), and a resonator (mirrors and waveguide). The purpose of the resonator is to confine the radiation and provide optical feedback while contributing a minimum of absorptive loss. The threshold condition for lasing is achieved when the wave reproduces itself after one round trip. That is, for a cavity mode to sustain oscillation, the round trip phase delay must be a multiple of  $2\pi$ , and the gain at that frequency must equal the total losses. This threshold condition is written as

$$\Gamma g_{th} = \alpha_w + \alpha_m, \quad (4.1)$$

where  $g_{th}$  is the threshold bulk gain of the active region and  $\Gamma$  is the confinement factor of the mode in the active region. The waveguide absorption loss is given by  $\alpha_w$ , and  $\alpha_m$  is the mirror loss due to the finite mirror reflectivities. This can be written as

$$\alpha_m = -\ln(R_1 R_2)/2L, \quad (4.2)$$

where  $R_1$  and  $R_2$  are the facet reflectivities, and  $L$  is the cavity length. For a GaAs/air interface,  $R = 0.32$ , and hence a typical 2-mm-long Fabry-Pérot ridge laser will have

$\alpha_m \approx 5.7 \text{ cm}^{-1}$ . This number may vary slightly as the effective mode index changes for various modes and waveguide types. By using highly reflective mirrors or increasing the cavity length the mirror loss can be minimized at the expense of the coupled out light. Unlike  $\alpha_m$ , the waveguide absorption loss  $\alpha_w$  doesn't depend on the cavity length, and it should be kept as small as possible.

## 4.2 One-dimensional mode solver

In order to design low-loss, high confinement waveguides, the one-dimensional multi-layer slab waveguide problem was solved using complex dielectric constants to represent the effects of free carriers and loss. The one-dimensional approximation is reasonable since for typical ridge waveguides the vertical dimension ( $\sim 10 \mu\text{m}$ ) is much smaller than the lateral dimension ( $\sim 60 - 200 \mu\text{m}$ ). Clearly this approximation is best for the widest waveguides. However, there are some additional cases where one must pay attention to the entire two-dimensional mode. One notable case is the semi-insulating surface plasmon waveguide described in Sec. 4.5.3. Only the transverse magnetic (TM) mode need be considered, since for intersubband lasers the dipole selection rule allows only radiation with the electric field polarized normal to the epitaxial layers.

The waveguide geometry is composed of slabs of infinite extent in the  $y$  and  $z$  directions, each characterized by a uniform dielectric constant  $\epsilon_i$  and thickness  $h_i$ , except for the upper and lower layers, which are semi-infinite. Propagation takes place along  $\hat{\mathbf{z}}$ , and the mode profile varies along  $\hat{\mathbf{x}}$  (the growth direction). Note that this coordinate system is standard for electromagnetic texts, and differs from that used in Chaps. 2 and 3 where  $\hat{\mathbf{z}}$  is the growth direction.

For TM modes of the waveguide, it is most convenient to work with the  $H$ -field, which is non-zero only in the  $y$  direction. The mode is given by the field solution:

$$\mathbf{H}(x, y, z) = \hat{\mathbf{y}} H_{y,\nu}(z) \exp(i\beta_\nu z), \quad (4.3)$$

where  $\nu$  is the mode index, and  $\beta_\nu$  is the propagation constant of the mode. For simplicity, we will generally omit the mode index  $\nu$  in future equations. Harmonic time dependence  $e^{-i\omega t}$  is assumed. In each layer  $i$  of homogeneous  $\epsilon_i$ , the solution must satisfy the scalar wave equation for  $H_y$

$$\frac{\partial^2 H_y}{\partial x^2} = (\beta^2 - \mu_0 \epsilon_i \omega^2) H_y. \quad (4.4)$$

Electromagnetic boundary conditions require the continuity of  $H_y$  (tangential component of  $H$ ) and  $(1/\epsilon) \frac{\partial H_y}{\partial x}$  (tangential component of  $E$ ) at each interface. The other two non-zero field components are given by

$$E_x = \frac{\beta}{\epsilon \omega} H_y, \quad E_z = \frac{i}{\epsilon \omega} \frac{\partial H_y}{\partial x}. \quad (4.5)$$

In each layer, the transverse mode profile takes on the form

$$H_y(x) = A_i \exp(i\kappa_i x) + B_i \exp(-i\kappa_i x), \quad (4.6)$$

where  $\kappa$  is the transverse wavenumber given by

$$\kappa_i^2 = \mu_0 \epsilon_i \omega^2 - \beta^2 \quad (4.7)$$

There are several types of solutions to this wave equation, which are analogous to the solutions of Schrödinger's equation in quantum mechanics. We are most interested in the guided (bound) modes, where the energy is localized inside the core waveguide material. There are also radiation (unbound) modes, where the energy is spread throughout space. These modes are generally not relevant for our waveguiding, as there is very little confinement in the gain medium. The exception is for slightly leaky modes, where only a small fraction of the radiation escapes, corresponding to an additional loss term. The multilayer equation is solved according to a transmission matrix formalism[122], and the resulting transcendental equation is numerically solved in MATLAB[79] to yield the propagation constants  $\beta$  for the guided modes of the

waveguide. The propagation constant (or equivalently the effective index :  $\beta = n_{eff}k_0$ ) is the eigenvalue for the mode. For lossless guides,  $\beta$  is real. For guides with complex valued  $\epsilon_i$ , the propagation constant is complex, and is given by

$$\beta = \beta' + i\beta'', \quad (4.8)$$

where  $\beta'' > 0$  corresponds to loss, and  $\beta'' < 0$  to gain. The power loss of the waveguide is given by

$$\alpha_w = 2\beta'', \quad (4.9)$$

indicating that the optical power in the waveguide decays with propagation along  $z$  as  $\exp(-\alpha_w z)$ .

Once the field solutions are known, we can obtain the confinement factor  $\Gamma$ , which indicates the optical overlap of the mode with the gain medium. We define a somewhat non-traditional confinement factor to reflect the fact that our MQW gain medium will only couple with  $\hat{\mathbf{x}}$  polarized light:

$$\Gamma = \frac{\int_{active} |E_x|^2 dx}{\int_{-\infty}^{\infty} |E|^2 dx}, \quad (4.10)$$

where the modal gain is given by  $\Gamma g_{bulk}$ .

### 4.3 Free carrier effects

The Drude theory for conductivity is used to obtain the contribution to the complex permittivities for the semiconductor and metal layers that results from free carrier effects. This theory assumes each independent electron moves with some momentum relaxation time  $\tau$ . This leads to a good model of the AC conductivity of a plasma where  $\mathbf{J}(\mathbf{r}, \omega) = \sigma(\omega)\mathbf{E}(\mathbf{r}, \omega)$ , provided that the wavelength of light  $\lambda$  is longer than the electron mean free path. This condition is easily satisfied inside GaAs, where  $\lambda > 20 \mu\text{m}$ , but is less good in high purity metals at low temperature. For example, using the time  $\tau = 0.12 \text{ ps}$  for Au at 77 K[123], the mean free path is  $\sim 1400 \text{ \AA}$ , compared

to a skin depth of  $\sim 200$  Å at 5 THz. While a nonlocal theory of conductivity should properly be used, in practice the observed optical response of Au and many other metals is still well fit using Drude theory. The other major deviation from the Drude model appears in the frequency dependence of  $\tau$ . Acoustic and optical phonon scattering mechanisms have separate frequency dependences, as does impurity scattering. Jensen considered the quantum contributions to the optical properties of polar semiconductors such as GaAs, and concluded that the Drude theory can be extended from DC to the near-infrared merely by using a frequency dependent electron scattering rate chosen to replicate the quantum result when used in classical Drude theory[124]. In the far-infrared limit this scattering rate reduces to a frequency independent constant that corresponds to the DC mobility.

The Drude model has in fact been found to accurately predict the optical properties of bulk GaAs in the terahertz frequency range. Time domain spectroscopy was used to measure the room temperature reflection from moderately doped wafers ( $n = 4 \times 10^{16}$  cm $^{-3}$ ) from 0.1 to 2.5 THz [125]. Similar measurements were made of the transmission for heavily doped thin film samples ( $n = 10^{16} - 2 \times 10^{18}$  cm $^{-3}$ ) from 0.2 to 3 THz [126]. In both cases the optical properties were well fit by using carrier densities and scattering times obtained from resistivity and mobility measurements. It should be noted that Huggard *et al.* [126] found that the optically determined mobilities tended to underestimate the Hall mobilities by approximately  $\sim 10\%$ . They attribute this effect to the fact that the optical technique determines the drift mobility, which is smaller than the Hall mobility by the Hall factor  $r_H = \mu_H/\mu_d$ . The differences between drift mobilities and Hall mobilities arise from transport differences that occur in a weak magnetic field, such that the transverse conductivity is proportional to  $\tau^2$ , so that the Hall factor is often defined as  $r_H \equiv \frac{\langle\langle\tau^2\rangle\rangle}{\langle\langle\tau\rangle\rangle^2}$ , where  $\langle\langle\rangle\rangle$  represents an energy average [127, 128]. As such, the value of the Hall factor represents the complicated interplay of various scattering mechanisms and their energy dependences. Fortunately, the Hall factor is typically close to unity for GaAs [129], especially at 77 K [128]. We can therefore confidently use the Drude model where the value of the relaxation time  $\tau$  is obtained from mobility data at the relevant

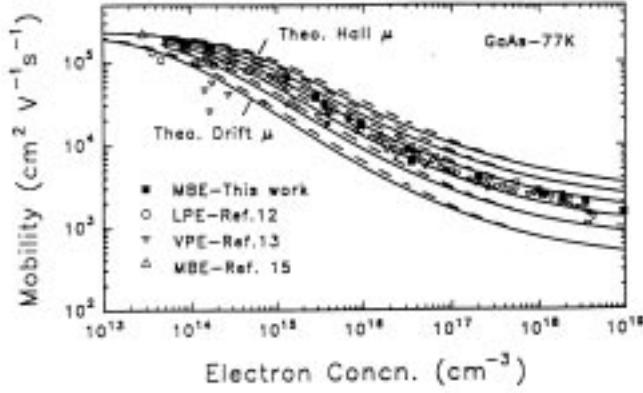


Figure 4-1: GaAs calculated electron Hall (dashed curves) and drift (solid curves) mobilities at 77 K as a function of electron concentration with compensation as a parameter ( $N_A/N_D = 0.0, 0.15, 0.30, 0.45, 0.60, \text{ and } 0.75$ ) (after Chin *et al.* [128]). Experimental Hall mobilities for a variety of grown structures are also shown. Note that the references cited in the figure do not correspond to those of this thesis.

temperature, using drift mobility data if possible.

Calculated and measured mobility data for bulk MBE grown GaAs at 77 K was compared by Chin *et al.* [128] and is shown in Fig. 4-1. The experimental data is best fit by a compensation ratio of 0.3, and at high concentrations there is no difference between the Hall and drift mobilities. Although the waveguides are designed to operate down to 4 K, we will in general use 77 K data. The primary reason for this is the lack of available data at lower temperatures. In general, losses will only decrease at lower temperatures as phonon scattering is frozen out, so use of the 77 K parameters represents a not best case scenario. Furthermore, for doping concentrations higher than  $10^{17} \text{ cm}^{-3}$  ionized impurity scattering dominates the scattering, and thus  $\tau$  will not change appreciably at lower temperatures.

The Drude model also accurately predicts the infrared dielectric constants of many metals, including gold [130]. Room temperature measurements of  $\epsilon$  in the terahertz were performed on Au and found a good fit using the value  $\tau = 0.024 \text{ ps}$ , which is close to the value  $\tau = 0.03 \text{ ps}$  obtained from the dc conductivity [123]. At low temperatures the situation is less clear cut. The 77 K dc conductivity gives a Drude scattering time of 0.12 ps [123], but optically fit data in the mid-infrared gives a time of at best 0.043 ps [131]. Below 77 K the scattering time depends heavily on the

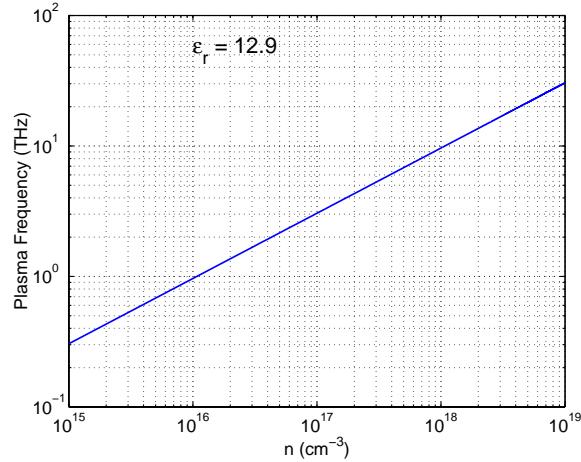


Figure 4-2: Bulk plasma frequency  $f_p$  in GaAs for  $\epsilon/\epsilon_0 = 12.9$  and  $m^*/m_0 = 0.067$ .

purity of the sample, but in very pure samples, the dc conductivity at 10 K indicates  $\tau > 10$  ps. To be conservative, a scattering time of  $\tau = 0.05$  ps for gold is used in the following loss calculations.

### 4.3.1 Drude model

The Drude model, described in numerous texts[123], predicts a frequency dependent conductivity

$$\sigma(\omega) = \frac{ne^2\tau}{m^*(1 - i\omega\tau)}, \quad (4.11)$$

when an electron is placed in a driving field  $\mathbf{E}(t) = \text{Re}(\mathbf{E}(\omega))e^{-i\omega t}$ . In this formula,  $n$  is the free carrier density,  $m^*$  the effective carrier mass, and  $e$  is the carrier charge. If the frequency dependent permittivity  $\epsilon(\omega) = \epsilon_{core} + i\frac{\sigma}{\omega}$  is used, we obtain

$$\epsilon(\omega) = \epsilon_{core} \left( 1 - \frac{\omega_p^2 \tau^2}{1 + (\omega\tau)^2} + i \frac{\omega_p^2 \tau}{\omega(1 + (\omega\tau)^2)} \right), \quad (4.12)$$

where  $\epsilon_{core}$  is the core permittivity excluding free carrier contributions, and

$$\omega_p^2 = ne^2/m^* \epsilon_{core} \quad (4.13)$$

defines the plasma frequency for the material. The dependence of the plasma fre-

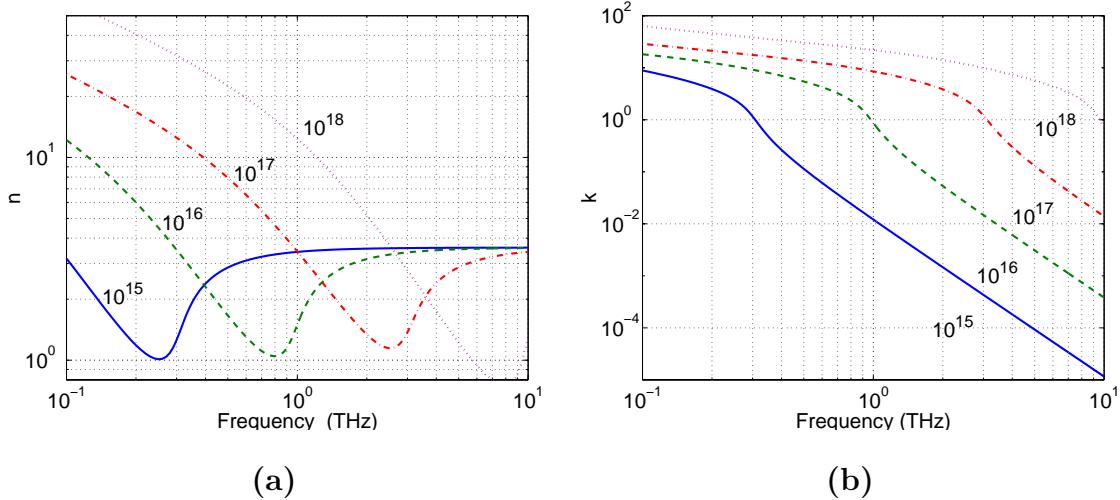


Figure 4-3: Real and imaginary part of the refractive index ( $\sqrt{\epsilon(\omega)} = n + ik$ ) for GaAs for free carrier densities  $10^{15}, 10^{16}, 10^{17}, 10^{18} \text{ cm}^{-3}$ , using  $\tau = 2.3, 0.7, 0.2, 0.1 \text{ ps}$  respectively, as given by mobility data in Fig. 4-1.

quency  $f_p = \omega_p/2\pi$  on the free carrier density (doping) is shown in Fig. 4-2. For the range of n-type doping available in MBE grown GaAs ( $< 5 \times 10^{18} \text{ cm}^{-3}$ ) it is seen that the plasma frequency spans the terahertz range. Hence, for a given frequency the semiconductor can be made to behave either as a dielectric ( $\text{Re}\{\epsilon\} > 0$ ) or a plasma ( $\text{Re}\{\epsilon\} < 0$ ) by the proper choice of carrier concentration.

### 4.3.2 Propagation in lightly doped regions

At low temperatures, for the level of doping over which typical terahertz frequencies can propagate, i.e.  $\omega > \omega_p$ , the Drude scattering time is generally long enough so that  $\omega\tau \gg 1$  and we have

$$\epsilon = \epsilon_{core} \left( 1 - \frac{\omega_p^2}{\omega^2} + i \frac{\omega_p^2}{\omega^3 \tau} \right). \quad (4.14)$$

For  $\omega \gg \omega_p$  the dielectric constant is slightly reduced by the factor  $1 - \omega_p^2/\omega^2$  and the free carrier loss is given by

$$\alpha_{fc} = \frac{\omega_p^2}{\omega^2} \frac{1}{c\tau} \sqrt{\frac{\epsilon_{core}}{\epsilon_0}} = \frac{ne^2 \lambda^2}{4\pi^2 m^* c^3 \tau \epsilon_0} \sqrt{\frac{\epsilon_0}{\epsilon_{core}}} \propto \frac{\lambda^2}{\tau} \quad (4.15)$$

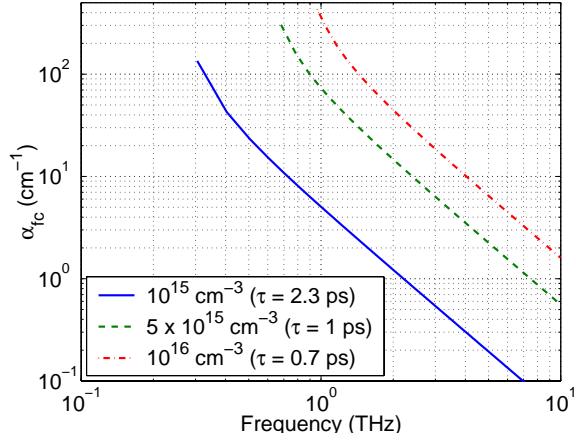


Figure 4-4: GaAs bulk free carrier loss above the plasma frequency for several carrier densities ( $\epsilon_r = 12.9$ ) using 77 K Drude scattering times[128].

The bulk free carrier loss is plotted in Fig. 4-4 for several doping densities for frequencies above the plasma frequency. The calculation was performed using Drude momentum relaxation times obtained from the 77 K mobility curves in Fig. 4-1.

### 4.3.3 Propagation in MQW active region

The loss due to free carrier absorption inside the active region remains one of the large uncertainties in waveguide loss calculations. Upon first examination, this loss might be expected to be negligible, since restriction of electron motion in the growth direction would limit its ability to interact with TM-polarized radiation. However, this is an overly optimistic view, especially since the degree of carrier confinement depends greatly on the nature of the active region. For example, a single quantum well injector differs greatly from a wide-miniband superlattice injector, which is quasi-three dimensional in nature. For this reason, loss inside the active region is generally considered using the bulk Drude model, where an average carrier density and the corresponding bulk mobility is used. This is the method used to calculate waveguide losses in Sec. 4.5. For short wavelength THz QCLs the cladding layer free carrier loss dominates over that from the active region, and this approximation is adequate. However, at longer wavelengths, the active region loss grows stronger and may be the

principal contribution to  $\alpha_w$ . Hence it is worthwhile to examine the validity of this approximation.

One of the few analyses of free carrier absorption at long wavelengths in a quantum well for TM-polarized radiation was undertaken by Vurgaftman and Meyer [132]. Along with the first order loss that results from the tails of the intersubband absorption, they considered the second order process of interface-roughness scattering assisted intrasubband absorption. In general, they calculated that the loss from second-order processes was nearly always much less than that from the broadened intersubband transition. This suggests that considering the tails of the intersubband absorption should be sufficient.

A phenomenological model for the relationship between intersubband and free carrier absorption can be found in Sec. III.C of Ando, Fowler and Stern [133]. In this treatment, a quantum well is described by an electron gas of sheet density population difference  $N_s \equiv N_0 - N_1$  over a width  $d_{eff}$  with a conductivity of

$$\sigma_{zz}(\omega) = \frac{N_s e^2 f_{10}}{m^* d_{eff}} \frac{-i\omega}{\omega_{10}^2 - \omega^2 - i\omega/\tau}, \quad (4.16)$$

where  $\hbar\omega_{10}$  is the intersubband transition energy,  $f_{10}$  is the oscillator strength, and  $\tau$  is a phenomenological relaxation time chosen to agree with the Drude relaxation time. In the limit of  $\omega_{10} = 0$ , a Drude-like expression for the conductivity is recovered:

$$\sigma_{zz}(\omega) = \frac{N_s e^2 f_{10}}{m^* d_{eff}} \frac{\tau}{1 - i\omega\tau}. \quad (4.17)$$

This expression is equivalent to Eq. (4.11) scaled by  $f_{10}$ .

The loss or gain associated with this transition is found by first using the expression for the complex permittivity  $\epsilon(\omega) = \epsilon_{core} + i\frac{\sigma}{\omega}$ . The permittivity can be rewritten as

$$\epsilon(\omega) = \epsilon_{core} + \epsilon_0(\chi' + i\chi''), \quad (4.18)$$

where the susceptibility  $\chi$  is due only to the intersubband oscillator. If the approxi-

mation  $\omega \approx \omega_{10}$  is made, the imaginary part of the susceptibility is given by

$$\chi'' = \frac{N_s e^2 f_{10}}{m^* d_{eff} \epsilon_0} \frac{1/4\tau\omega_{10}}{(\omega_{10} - \omega)^2 + (1/2\tau)^2}. \quad (4.19)$$

The loss is given by  $\alpha = 2\text{Im}\{n\}\omega/c$ , where  $n$  is the refractive index. If the condition  $|\chi|\epsilon_0 \ll \epsilon_{core}$  holds, then the loss is simply

$$\alpha = \frac{\chi''\omega}{nc}. \quad (4.20)$$

This leads to an expression for loss

$$\alpha = \frac{N_s e^2 f_{10}}{4m^* d_{eff} c n \epsilon_0} \frac{1/\tau}{(\omega_{10} - \omega)^2 + (1/2\tau)^2} = \frac{N_s e^2 f_{10}}{4m^* d_{eff} c n \epsilon_0} \frac{\Delta\nu/2\pi}{(\nu_{10} - \nu)^2 + (\Delta\nu/2)^2}, \quad (4.21)$$

where the equation has equivalently been written using frequency  $\nu$ , and the full-width-half-max linewidth  $\Delta\nu = 1/(2\pi\tau)$ . Note that this expression is equivalent to the gain formula given in Eq. 2.35, provided  $N_s$  is taken to be the population difference between the two subbands. Note also that the time  $\tau$  used above is chosen for agreement with Drude theory, and is different than the traditional dephasing time  $\tau_{dp}$  for homogeneously broadened transitions, which is related to the linewidth by  $\Delta\nu = 1/(\pi\tau_{dp})$ .

As shown above, the effect of carrier confinement along the growth direction is to shift the oscillator frequency up from zero to various intersubband transition frequencies, each weighted by their oscillator strength. For typical QCLs, most electrons reside in some multi-well injector of closely spaced energy levels. As photon energies are reduced to nearer the injector subband spacings, the effect of confinement may significantly affect  $\chi$ , i.e. the loss will increase and  $\epsilon(\omega)$  will decrease compared to what one would expect from the Drude model. This effect may be mitigated by the fact that injector oscillator strengths are typically less than unity. Also, any equalization of the subband populations will reduce these effects. Furthermore, caution should be taken in extrapolating losses due to intersubband transitions using Drude relaxation times  $\tau$  obtained from mobility data, as  $\tau$  is likely to be quite different in a MQW

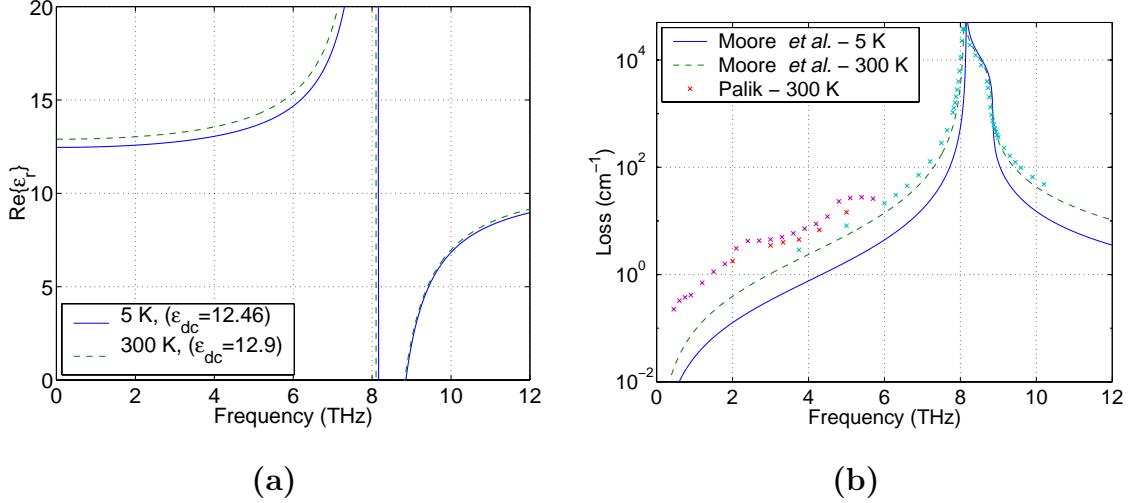


Figure 4-5: The relative permittivity  $\epsilon_r$  and loss for GaAs in the terahertz due to phonon interactions are shown in (a) and (b) respectively. The solid lines are a dual oscillator fit to experimental data by Moore *et al.* [134] and the individual points are experimentally measured from several sources collected in Palik [135].

system compared to that in bulk.

In light of the uncertainties involved in calculating the dielectric free carrier response inside the active region, approximating  $\epsilon(\omega)$  in the active region with the bulk Drude model is reasonable. However, for longer wavelength QCLs these effects will become more important, and the injector should be designed with this in mind.

## 4.4 Phonon coupling effects

Coupling with optical phonon modes causes dispersion and optical loss in the vicinity of the *Reststrahlen* band ( $\sim 33\text{--}36$  meV (8–9 THz) in GaAs) and places an upper limit on the operating frequency of THz lasers. The experimentally measured permittivity and loss of semi-insulating GaAs at 5 K and 300 K is shown in Fig. 4-5. For lasers operating below 4 THz, at cryogenic temperatures the phonon loss should not be significant. For slightly higher frequencies, it may contribute to the waveguide loss, but is still usually dominated by the free carrier contribution for all but the most lightly doped regions. In all waveguide calculations below, the effects of phonon

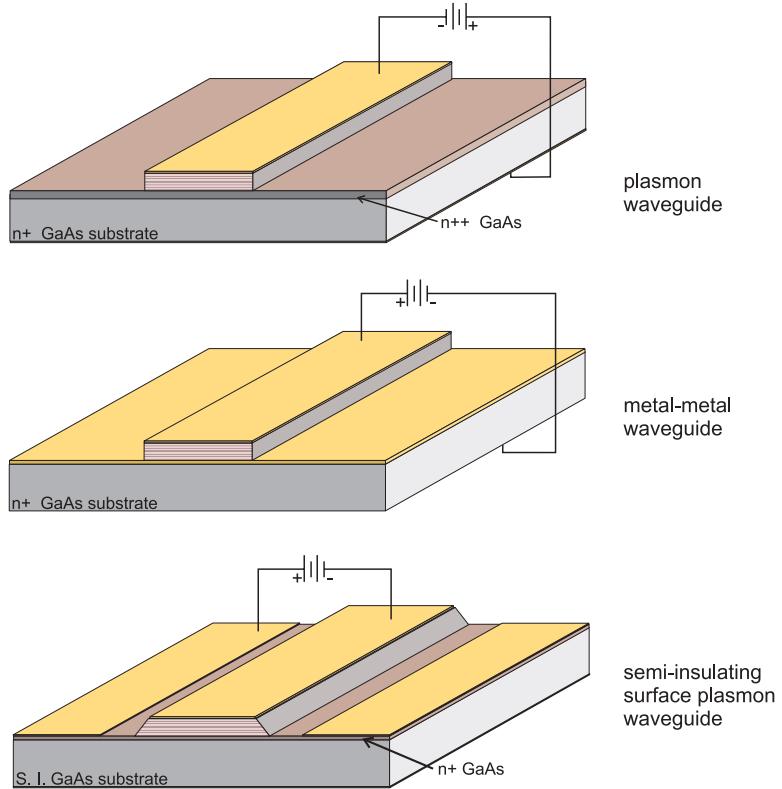


Figure 4-6: Drawing of three terahertz waveguides.

coupling on the permittivity and loss are neglected. Since for all the waveguides discussed, the mode resides almost entirely in GaAs, the phonon loss can simply be added to the calculated waveguide loss.

## 4.5 Terahertz waveguides

The difficulty of developing a low-loss waveguide was one of the primary stumbling blocks in the development of terahertz intersubband lasers. Mid-infrared quantum cascade lasers, like near-infrared and visible diode lasers, have traditionally used dielectric mode confinement to confine radiation in the vertical (growth) direction. For example, in InGaAs/AlInAs QCLs, the InP substrate has a lower refractive index than the active region, which provides natural cladding, and cladding above the active region is provided by MBE grown InP or AlInAs. These layers are also doped to allow electrical bias of the device.

Such a scheme is not feasible at terahertz frequencies for two reasons. First, to be effective, the thickness of the cladding layers must be on the order of the wavelength of the radiation inside the semiconductor, which would be larger than  $10\ \mu\text{m}$ . Such thick layers would take an unreasonably long time to grow by MBE. Second, loss due to free carrier absorption increases as  $\lambda^2$  (Fig. 4-4) for frequencies above the plasma frequency. To have a chance at lasing, the mode overlap with doped regions must be minimized. Furthermore, for GaAs/AlGaAs growth systems, the GaAs substrate has a higher refractive index than AlGaAs, and hence the substrate provides no natural mode confinement.

In the sections below, I will compare three types of waveguides for use in the terahertz: the plasmon waveguide, the semi-insulating-surface-plasmon waveguide, and the metal-metal waveguide. Sketches of these three waveguides are shown in Fig. 4-6. The modes in these waveguides are all designed around surface plasmons, which are localized modes that propagate along the interface between two materials with dielectric constants of opposite sign. The plasma layers are either metals, or GaAs layers heavily doped to act as “metals” ( $\Re\{\epsilon\} < 0$ ). In the simulations below, loss due to phonon coupling is neglected, and free carrier loss in the cladding layers and active region is accounted for using the Drude model as described in Sec. 4.3.

#### 4.5.1 Plasmon waveguide

In the plasmon waveguide, radiation is confined between the upper metallic contact and a thick heavily semiconductor doped layer. Due to the amphoteric nature of *n*-Si doping in GaAs, the maximum possible bulk doping level is approximately  $n = 5 \times 10^{18}\ \text{cm}^{-3}$ . Since the plasma frequency for this carrier density is 21.5 THz, it is possible to take advantage of reflective properties of the plasma at terahertz frequencies. For the plasmon waveguide the confinement below the active region is provided by a thick ( $1\text{--}2\ \mu\text{m}$ ) highly doped GaAs layer grown over the *n*<sup>+</sup> substrate ( $n = 1 - 2 \times 10^{18}\ \text{cm}^{-3}$ ). Upper confinement is provided by the metallic contact, which is placed directly above the active region, separated only by a thin heavily doped contact layer.

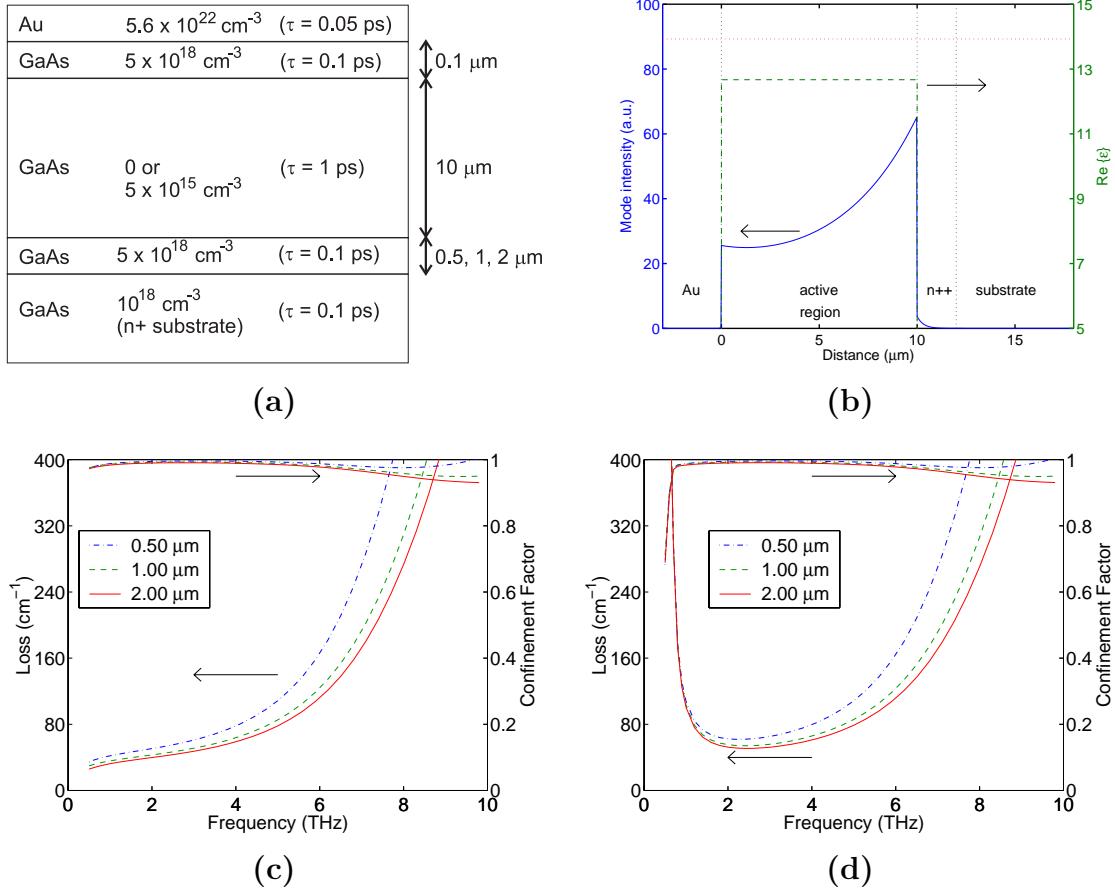


Figure 4-7: Calculated loss and confinement of plasmon waveguide versus frequency for structure schematically shown in part (a). (b) Typical mode profile for plasmon waveguide at 5 THz. (c) and (d) Loss and confinement for structures with various lower cladding layer thicknesses both without (c) and with (d) active region doping at  $5 \times 10^{15} \text{ cm}^{-3}$  ( $\tau = 1 \text{ ps}$ ).

The properties of this waveguide were obtained versus frequency for several thicknesses of lower cladding layer and are shown in Fig. 4-7. Part (a) displays the schematic structure along with Drude parameters for the layers, and a typical mode profile is shown in part (b). The loss and confinement factor are shown in parts (c) and (d) both without and with active region doping respectively. The majority of the loss is due to free carrier absorption in the lower cladding layer due to the significant skin depth. The 2- $\mu\text{m}$  cladding layer is optimal, as it prevents the field from leaking into lower doped substrate, where the skin depth is even larger. The cladding improves at lower frequencies as the dielectric constant is reduced according to  $\Re\{\epsilon\} \propto 1 - \omega_p^2/\omega^2$  (Eq. 4.14). However, the free carrier loss in the active region will increase as the frequency is reduced, and will offset this advantage.

In general, the limit on the GaAs carrier density to  $5 \times 10^{18} \text{ cm}^{-3}$  results in a large penetration depth of the mode into the cladding ( $\delta \approx 0.5 - 1 \mu\text{m}$ ). Losses therefore tend to be large, which was confirmed by an absorption measurement to be approximately  $50 \text{ cm}^{-1}$  at 4 THz performed by Rochat *et al.* [136].

### 4.5.2 Metal-metal waveguide

One improvement to the lossy plasma waveguide can be made by replacing the lower heavily doped GaAs layer with a layer of metal. This forms what is essentially a microstrip waveguide, as is used in microwave and millimeter-wave waveguiding. Since the skin depth in gold is only several hundred Angstroms in the terahertz, this metal-metal waveguide can be made much less lossy with a confinement factor close to unity. These advantages come at the expense of processing complexity, as a flip-chip wafer bonding technology is necessary, as described in Sec. 5.3. Such a waveguide was been used to successfully fabricate a mid-infrared QCL at 21–24  $\mu\text{m}$  [137]. However, in the mid-infrared, the operating frequency is closer to the plasma frequency and the metal is lossier, so there is no clear advantage of a metal waveguide. Its true benefits are realized in the terahertz.

The properties of a 10- $\mu\text{m}$ -thick gold-gold waveguide are shown in Fig. 4-8. The waveguide schematic is shown in (a), and a typical mode profile is shown in (b). The

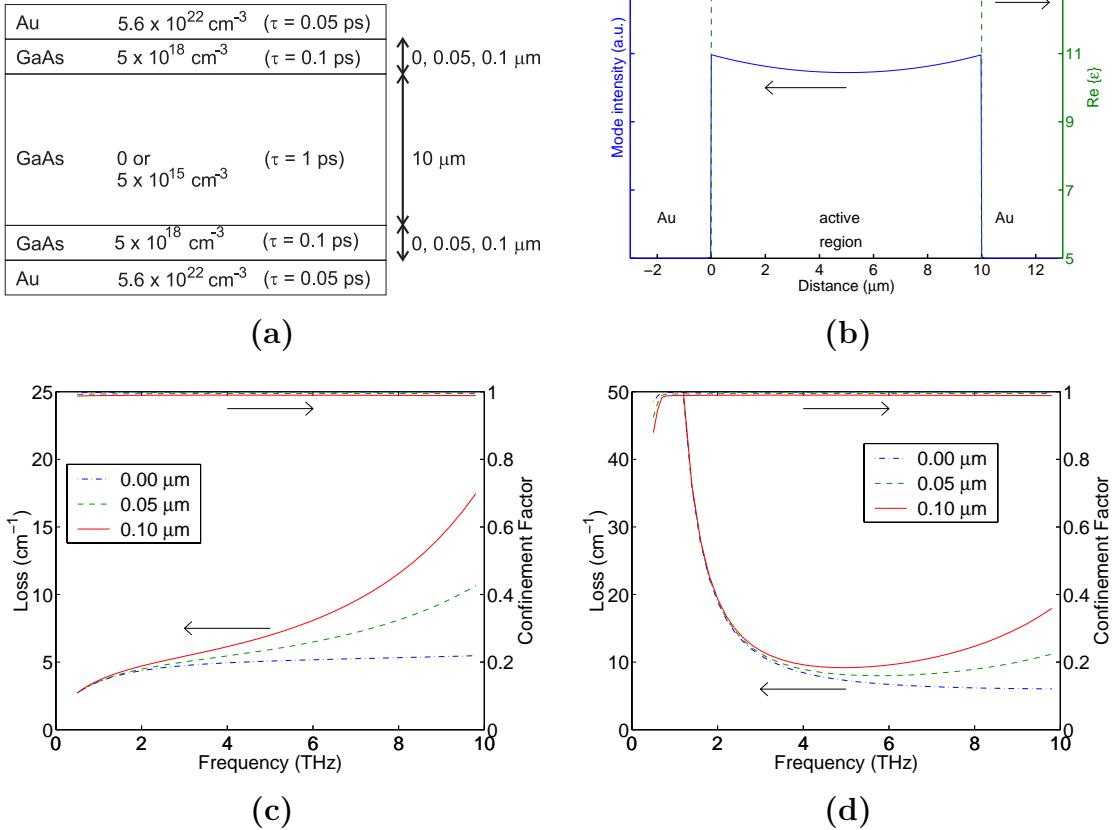


Figure 4-8: Calculated loss and confinement of metal-metal waveguide versus frequency for structure schematically described in part (a). (b) Typical mode profile for metallic waveguide at 5 THz. (c) and (d) Loss and confinement for structures with various contact layer thicknesses both without (c) and with (d) active region doping at  $5 \times 10^{15} \text{ cm}^{-3}$  ( $\tau = 1 \text{ ps}$ ).

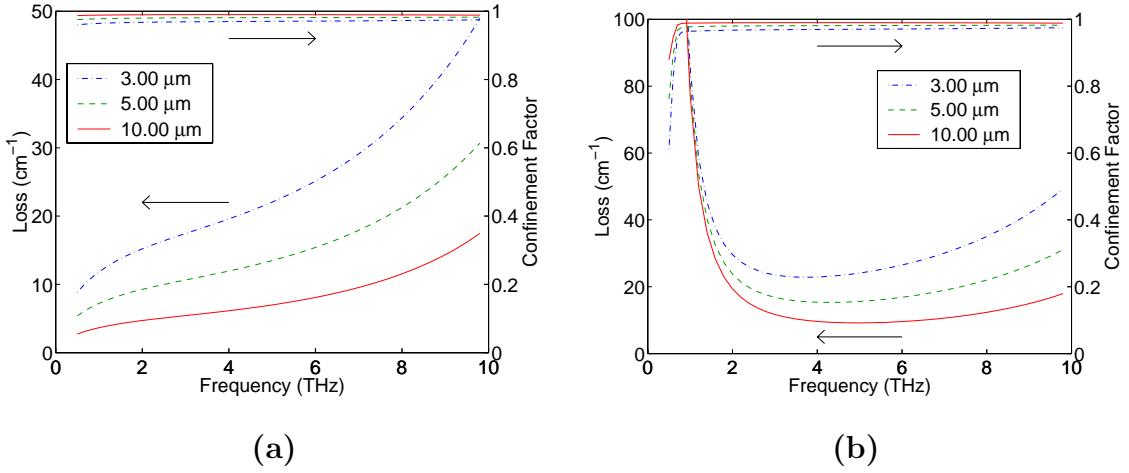


Figure 4-9: Loss and confinement of metal-metal waveguide versus frequency for active regions of various thickness for an (a) undoped active region and (b) with the active region doped at  $5 \times 10^{15} \text{ cm}^{-3}$  ( $\tau = 1 \text{ ps}$ ) (b). The model waveguide structure is the same as in Fig. 4-8 with 0.1- $\mu\text{m}$ -thick contact layers.

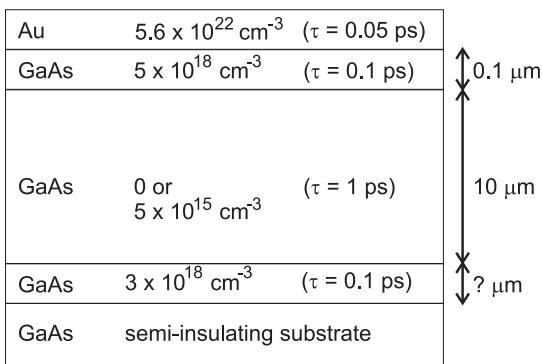
loss and confinement for both undoped and doped active regions are shown in (c) and (d). The need for electrical contacts requires the presence of heavily doped GaAs contact layers adjacent to the metallic films. Unfortunately, these contacts can be a significant source of loss, and their thickness should be minimized. This is illustrated in (c) and (d), where the loss is plotted for several thicknesses of the contact layers. The intrinsic waveguide loss is so low ( $\sim 5 \text{ cm}^{-1}$ ) that any free carrier loss in the active region may easily dominate.

The loss and confinement in metallic waveguides of different thicknesses is shown in Fig. 4-9. For thinner waveguides, the loss increases because the long wavelength radiation is forced further into the lossy contact and metallic regions. If this increase in loss can be tolerated, the active region can be made thinner, which shortens the excessively long MBE growth times, which can be longer than 12 hours for a 10  $\mu\text{m}$  growth. A thinner active region will use fewer modules and will have a lower threshold voltage and dissipate less power. While the metal-metal waveguide possesses potentially excellent optical properties, it is necessary to take great care during processing to ensure adequate thermal and mechanical properties. These issues are discussed in section 5.3.

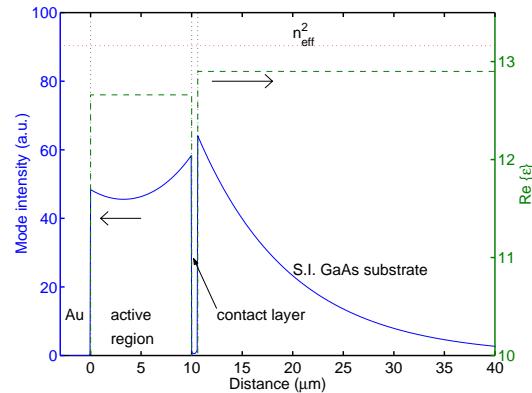
### 4.5.3 Semi-insulating-surface-plasmon waveguide

The semi-insulating-surface-plasmon waveguide, first described by Ulrich *et al.*, was a key enabling component of the first terahertz QCLs [64, 65, 66]. In this structure, the mode is bound to the upper metallic contact and a thin ( $0.1\text{--}1\ \mu\text{m}$ ) heavily doped contact layer grown directly beneath the active region and above the semi-insulating (SI) GaAs substrate. Because the doped layer is thinner than its own skin depth, the mode extends substantially into the substrate. However, because the mode overlap with the doped contact layer is small, the free carrier loss is minimized.

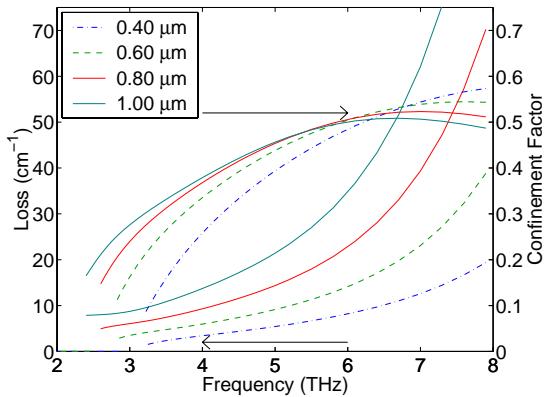
An overall description of the SI-surface-plasmon waveguide is shown in Fig. 4-10. The schematic for the structure is shown in part (a) and a typical mode pattern for a waveguide at 5 THz is shown in part (b). The waveguide design is optimized for a given laser by adjusting the thickness and doping of the lower contact layer. Waveguide performance is shown in parts (c), (d), and (e) for a typical waveguide with an average active region doping of  $5 \times 10^{15}\ \text{cm}^{-3}$ , and a contact layer of various thicknesses doped at  $3 \times 10^{18}\ \text{cm}^{-3}$ . Examination of part (c) reveals that both the loss  $\alpha_w$  and confinement  $\Gamma$  tend to decrease with reduced frequency. The loss is mostly due to lower contact layer, although the loss from inside the active region is significant at lower frequencies. In general, thicker contact layers tend to increase confinement, at the cost of higher loss. Waveguide performance is often analyzed by examining the term  $\alpha_w/\Gamma$  (see Fig. 4-10(d)), since this term appears in the expression for threshold gain  $g_{th}$  (see Eq. (4.1)). However, this particular figure of merit neglects the mirror losses  $\alpha_m$  from the cavity, and thus underestimates the penalty of having a low confinement factor  $\Gamma$ . The term  $\alpha_w/\Gamma$  changes little as confinement drops, since much of the loss  $\alpha_w$  is due to overlap with the contact layer. However, the term  $\alpha_m/\Gamma$  can become quite significant for cases where  $\alpha_w \leq \alpha_m$ . The sum of these two terms is shown in Fig. 4-10(e) for  $\alpha_m = 5.5\ \text{cm}^{-1}$ , which is the value for an uncoated 2-mm-long Fabry-Pérot cavity. While this is merely a typical case, and  $\alpha_m$  may be made smaller by using a longer cavity or high-reflectivity (HR) coatings, the reduction of  $\Gamma$  at low frequencies is problematic for using SI-surface-plasmon waveguides at very



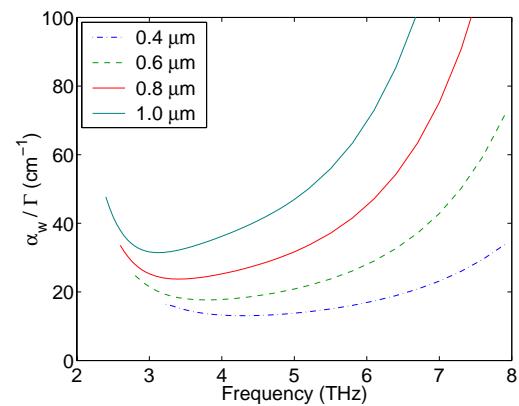
(a)



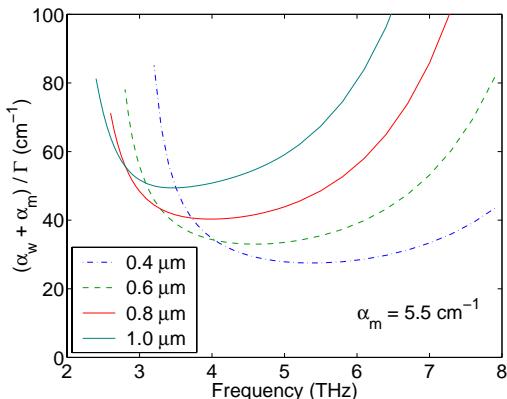
(b)



(c)



(d)



(e)

Figure 4-10: (a) Schematic and (b) typical mode pattern at 5 THz for SI-surface-plasmon waveguide. (c) Loss and confinement for structures with various contact layer thicknesses with active region doping at  $5 \times 10^{15} \text{ cm}^{-3}$  ( $\tau = 1 \text{ ps}$ ). (d) and (e) Loss divided by confinement both without (d) and including (e) facet loss from an uncoated 2-mm-long Fabry-Pérot cavity.

long wavelengths.

The reduction in  $\Gamma$  with frequency has two causes. First, the geometric effect of a longer wavelength with respect to the active region thickness reduces the overlap. Second, the active region doping reduces the dielectric constant in the active region compared to the SI substrate. According to the Drude model, the presence of free carriers approximately reduces  $\Re\{\epsilon(\omega)\}$  by the factor  $1 - \omega_p^2/\omega^2$  (see (4.14)). This reduction in  $\epsilon(\omega)$  can already be seen at 5 THz in Fig. 4-10(b), and grows more severe at longer wavelengths. As a result, the effective index  $n_{eff} = \beta/k_0$  is reduced, and the mode extends further into the substrate (see Eq. (4.7)). In fact, at some cutoff frequency,  $n_{eff}$  falls below the refractive index of the substrate, and the mode becomes unbound (leaky). In Fig. 4-10(c-e), the waveguide performance is only plotted down to this cutoff wavenumber. The cutoff frequency can be minimized by reducing the active region doping or by increasing the contact layer thickness and reducing its doping. However these modifications come at the cost of reduced design flexibility and increased loss. Alternatively, the SI-surface plasmon waveguide can be operated even beyond cutoff by lapping the substrate extremely thin and using the backside metallization for confinement. This method was used by the Neuchâtel group recently to demonstrate a 2.3 THz QCL using an 18- $\mu\text{m}$ -thick active region where the substrate was reduced to a thickness of 80  $\mu\text{m}$  [45].

In general, the SI-surface-plasmon waveguide can be adequately modeled by a one-dimensional solver. taken. However, since it is grown on a SI-GaAs wafer, lateral electrical contacts are made adjacent to the ridge structure (Fig. 4-6), and extra care must be taken to prevent coupling with lossy surface plasmons associated with these contacts. To prevent this, most masks were fabricated with a 50- $\mu\text{m}$  separation between the contacts and the ridge. These issues and others related to two-dimensional modeling of terahertz waveguides are discussed in Kohen *et al.* [138].

#### 4.5.4 Waveguide comparison

While the plasmon waveguide was the first type proposed for terahertz use, in GaAs the limit on the  $n$ -type doping level allows the mode to penetrate too far into the

cladding, which leads to unacceptably high losses. One is left only with a choice between the metal-metal waveguide and the SI-surface-plasmon waveguide. As shown above, the metal-metal waveguide demonstrates superior performance at all frequencies, i.e. its  $\alpha_w/\Gamma$  is usually lower or equivalent to that of the SI-surface-plasmon waveguide. Since its confinement factor is almost unity, there is a reduced need for facet coatings, which are a primary failure mode in lasers, especially during cryogenic cycling. Additionally, using a metal-metal waveguide, thinner and narrower waveguides are more feasible, which leads to reduced power dissipation and easier heat removal. Furthermore, the use of lateral heat removal techniques on the laser ridge such as epitaxial regrowth, ion implantation isolation [48], or thick gold electroplating [139], or epi-layer down mounting are more easily facilitated using the metal-metal waveguide. The lateral contact geometry for the SI-surface-plasmon guide makes these techniques more difficult. The principal disadvantage for the metal-metal waveguide is the added processing complexity and potential for reduced yield. For example, imperfect bonding will lead to a high thermal contact resistance and reduced performance (Sec. 5.3.3). However, this should be correctable by improved fabrication. Indeed, the thermal conductivity of metal improves relative to semiconductor at higher temperatures. For example,  $\kappa_{Au} > \kappa_{GaAs}$  at  $T > 100$  K. While at intermediate terahertz frequencies (3–6 THz) either waveguide could be used, below roughly 2.5 THz, the performance of the SI-surface-plasmon waveguide degrades dramatically, and the metal-metal waveguide becomes increasingly advantageous. The high confinement factor of the metal-metal waveguide may also become more advantageous at higher terahertz frequencies ( $> 6$  THz) near the *Reststrahlen* band. In this case the modal overlap with lossy GaAs that doesn't contribute gain should be minimized. We have realized the first terahertz QCL using a metallic waveguide (see FL152Fmm in Sec. 7.3) [70]. The advantage of the metal-metal waveguide was demonstrated by processing pieces of the same wafer into both types of waveguide; lasing was only observed in the metal waveguide device. Furthermore, a metal-metal waveguide was used to obtain pulsed lasing up to a record high temperature of 137 K, while the same device in a SI-surface-plasmon waveguide lased only to 92 K (see Sec. 7.6).

# Chapter 5

## Experimental Setup and Fabrication

### 5.1 Experimental setup

A variety of measurements were performed to characterize various laser and would-be laser structures. The primary measurements involve measuring the transport characteristics of a device, such as current versus voltage ( $I$ - $V$ ) and conductance versus voltage ( $G$ - $V$ ) characteristics. Equally as important is measuring the power level and spectrum of radiation emitted due to spontaneous and/or stimulated emission. These measurements are done almost exclusively at cryogenic temperatures. Also, capacitance measurements (CV) are used to measure the average doping in a structure, and transmission line model (TLM) measurements measure contact resistances.

#### Device dewar and electric biasing

Due to the intersubband selection rule (Sec. 2.3), emitted radiation is polarized in the transverse magnetic direction and for structures with unpatterned surfaces it is emitted only from the edges. However, in devices where the waveguide loss is much greater than the mirror loss  $\alpha_w \gg \alpha_m$ , only a small fraction of emitted light escapes the waveguide. In many devices tested,  $\alpha_w \approx 100 - 200 \text{ cm}^{-1}$  so that most of the

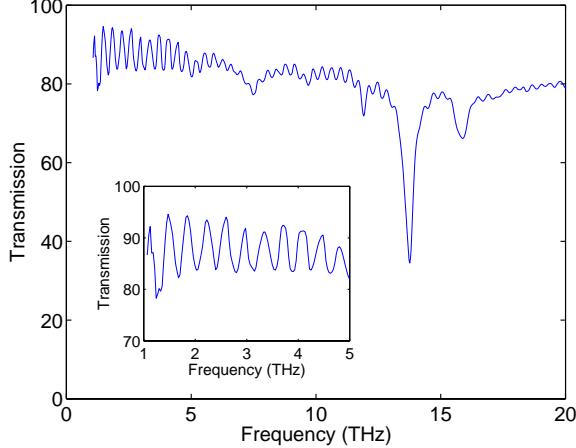


Figure 5-1: Transmission of 0.25-mm-thick polypropylene sheet at room temperature used as window on device dewar (inset zooms in on 1–5 THz region).

photons emitted further away from the edge than  $\alpha_w^{-1} \approx 50 - 100 \mu\text{m}$  are absorbed. For this reason, some devices were fabricated with a metallic grating on top of the device, typically with a period of 10–25  $\mu\text{m}$  and a 50% fill factor. The purpose of the grating is to efficiently couple radiation from the surface of a device, even in the presence of high losses. This was one of the earliest methods for measuring intersubband emission [36], and is described in detail in Refs. [59, 140]. However, the edge emission configuration is preferred whenever possible, because it eliminates the possibility of any grating induced artifacts in the emission spectrum.

The optical measurements were performed with the device mounted in vacuum on a copper cold plate, that could be cooled to 4.2 K with liquid helium (LHe) or to 77 K with liquid nitrogen (LN). The dewar output window was a 0.25-mm-thick polypropylene sheet with approximately a 85% transmissivity in the terahertz (Fig. 5-1). Custom device mounts made from oxygen free high conductivity (OFHC) copper were designed to hold the device to the cold plate in both edge emitting and surface emitting configurations. The edge emitting mount with cone is pictured in Fig. 5-2(a). The surface emitting mount (not shown) is an L-shaped bracket designed so that the same type of chip carriers could be mounted perpendicular to the cold plate surface. The devices were indium soldered epi-layer side up to an OFHC copper chip carrier in order to provide a good thermal and electrical contact to the heat sink.

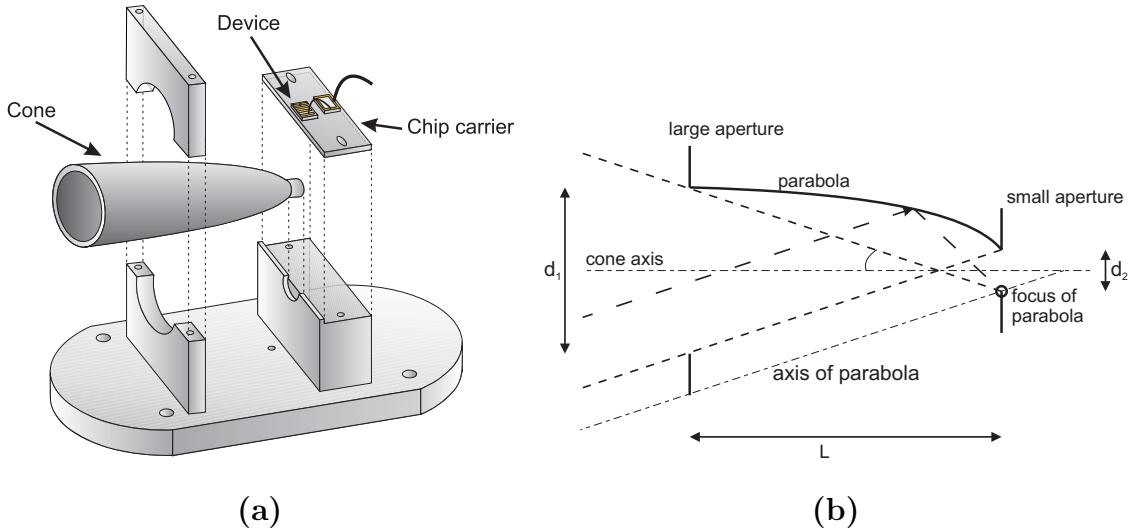


Figure 5-2: (a) OFHC copper mount used for edge emission measurements to hold cone and chip carrier with device. (b) Diagram of Winston cone used for output coupling. The cone is formed by rotating a parabolic section about an axis (different from the parabolic axis). Radiation originating within the smaller aperture will emerge with an angle of divergence  $\theta < \theta_{max}$ .

The use of indium solder rather than conductive silver epoxy noticeably improved heat sinking, and dramatically reduced the amount of blackbody radiation observed. This is not surprising, as the thermal conductivity of conductive epoxies tends to be dominated by the epoxy matrix, rather than the suspended conductive particles. Electrical contact was made from a bonding pad mounted on the chip carrier to the emission structures with aluminum or gold wire bonds. A silicon diode temperature sensor (not shown) was mounted on the side of the copper mount just below the mounting point for the chip carrier. A layer of soft indium foil was placed at all metal-metal interfaces to improve the thermal contact. A one foot length of coaxial cable inside the dewar<sup>1</sup> was used to deliver the electrical bias and monitor the voltage across the device, which allowed the delivery of pulses with a fast ( $< 10$  ns) rise time.

Since the free space wavelength of the emitted radiation ( $\sim 100$   $\mu\text{m}$ ) is much larger than the thickness of the waveguide ( $\sim 10$   $\mu\text{m}$ ), the radiation strongly diverges upon exit from the edge emitting device. For surface emitting structures, due to the larger

---

<sup>1</sup>Ultra-miniature coaxial cable type C (32 AWG copper center conductor, 50  $\Omega$  characteristic impedance), with part number 9002-001 available from Lake Shore.

size of the emission area, the divergence is not as strong. In order to collimate the emitted light for coupling out of the dewar, a Winston cone manufactured by Infrared Laboratories was used for output coupling for edge emitting devices (a straight cone was used for surface emitting devices). A Winston cone is an off-axis parabolic condenser usually used for collection of far-infrared and sub-millimeter radiation [141, 142]. A schematic diagram is shown in Fig. 5-2(b). For a device placed at the smaller aperture of the cone, radiation emitted at any angle will emerge from the larger aperture at an angle  $\theta$  smaller than some maximum angle

$$\tan \theta_{max} = \frac{d_1 + d_2}{2L}, \quad (5.1)$$

where  $d_1$  and  $d_2$  are the diameters of the larger and smaller apertures, and  $L$  is the length of the cone. Such a cone does not have a well defined focal point however, making it somewhat more difficult to focus the emitted radiation. The dimensions chosen were  $d_1 = 0.366''$ ,  $d_2 = 0.063''$ , and  $L = 1.219''$ , yielding  $\theta_{max} = 10^\circ$ . The cone was electroformed from copper and plated with gold, ensuring high reflectivity.

Electrical bias was provided in a number of ways. A homemade bias box constructed by Jurgen Smet [37] and modified by Hans Callebaut was used to provide high duty cycle pulses with rise times of approximately 1  $\mu$ s. For duty cycles up to 10% with  $\sim$ 10 ns rise times, a pulser from AVTECH systems (model AV-1011-B) was used. A second AVTECH pulser (mode AVO-6C-B) was occasionally used to provide especially high power pulses (up to  $\sim$ 5 A) but required the use of a  $50\text{-}\Omega$  impedance-matching resistor. Alternatively, when optical measurements were not necessary, small devices were mounted on a dipstick and immersed in LHe or LN. This allowed the rapid testing of devices without the need for pumping down the vacuum of a dewar. Current was measured using various current-sensing resistors for long pulsed measurements, and using a inductive current sensing loop<sup>2</sup> mounted just outside the dewar for the fast pulsed measurements. Pulsed measurements were taken using calibrated boxcar averagers, and the data was collected electronically.

---

<sup>2</sup>Model 711S wideband current probe, American Laser Systems, Santa Barbara, CA.

## Terahertz detectors

Direct detection of the picowatt level terahertz spontaneous emission was performed using either a silicon composite bolometer, or a gallium doped germanium (Ge:Ga) photodetector, both of which required cryogenic cooling with LHe. Measurements were usually performed in pulsed mode (i.e. electrical chopping) and lock-in detection was used to reduce the noise bandwidth.

The Si bolometer is a broadband thermal detector where the spectral response is controlled by a series of cut-on filters. Radiation is collected via a Winston condensing cone onto the detector element,<sup>3</sup> which is a blackened diamond thermal absorber 2.5 mm in diameter thermally bonded to a small silicon element. It operates at 1.6 K. Since the detection is limited by the thermal response of the element, the response time is on the order of milliseconds. For this reason, we usually operate the bolometer with a chopping frequency of not higher than 400 Hz, at which point the manufacturer reports that the response has fallen to 84%. The reported responsivity is  $8.47 \times 10^5$  V/W, not counting the preamplifier gain, which is usually set to 200. However, this calibration was performed electrically not optically, and is off by at least a factor of 2, and perhaps as much as 25 in the mm-wave region. The noise equivalent power is reported to be  $\text{NEP} = 4.36 \times 10^{-14}$  W Hz<sup>-1/2</sup>. The bolometer window is composed of 1-mm-thick high density polyethylene with a diamond coating, which is reported to have a transmission of greater than 80% below 6 THz (200 cm<sup>-1</sup>). The four filters available have cut-on frequencies of 600 cm<sup>-1</sup> (18 THz), 200 cm<sup>-1</sup> (6 THz), 100 cm<sup>-1</sup> (3 THz), and 35 cm<sup>-1</sup> (1 THz) and are labeled 1–4 respectively. One can obtain an idea of their spectral responses by examining 300 K blackbody spectra taken with filters 1, 2, and 3, as shown in Fig. 5-3. Keep in mind that these are blackbody radiation spectra and not transmission curves for the filters. Filter 2 should be used with care since the filter is quartz and has an absorption at 130 cm<sup>-1</sup> [143].

The Ge:Ga photodetector measures the conductivity change due to photoexcitation of electrons from shallow impurity states into the valence band. Operation takes

---

<sup>3</sup>Detector element is unit number 2765 from Infrared Laboratories, Tucson, AZ.

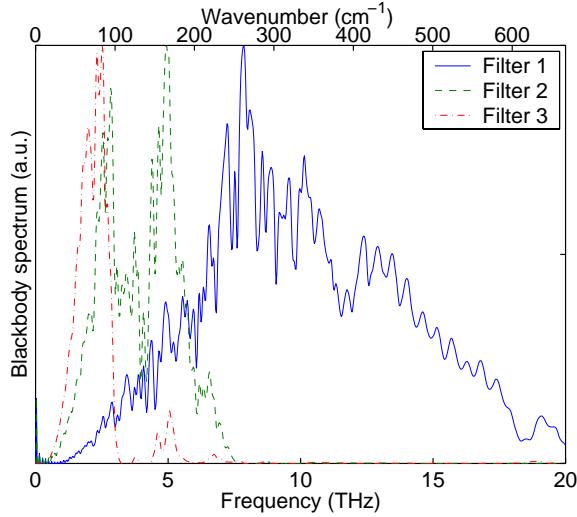


Figure 5-3: Purged blackbody spectra taken using Si bolometer at  $4 \text{ cm}^{-1}$  resolution with filters 1,2, and 3 (ignore relative intensity).

place at 4.2 K. Light is collected through a high density polyethylene window and a  $310 \text{ cm}^{-1}$  (9.4 THz) cut-on filter, into a Winston condensing cone, and onto the  $1 \text{ mm} \times 1 \text{ mm}$  detector element.<sup>4</sup> The NEP is reported to be  $1.70 \times 10^{-13} \text{ W Hz}^{-1/2}$ . The principal advantage of this detector over the bolometer is its faster response. Although the manufacturer reports a maximum chopping frequency of 5 kHz for the detector and amplifier circuit, sub-microsecond response times have been observed. One of the more difficult measurement problems is measuring picowatt level intersubband emission in the presence of the blackbody radiation that results from the ohmic heating of the device. Since many of the thermal time constants associated with the substrate and chip carrier are on the order of a millisecond or longer, if an electrical chopping frequency of 10 kHz or greater is used, one can operate beyond the thermal  $f_{3dB}$  point. Since the response of intersubband emission is virtually instantaneous with bias, the blackbody power detected by the lock-in amplifier is reduced while the intersubband emission signal is unchanged. In many cases, this method has significantly reduced blackbody spectral contamination. One disadvantage to the Ge:Ga photodetector is its sharp spectral cutoff at  $80 \text{ cm}^{-1}$  (2.4 THz) which corresponds to the acceptor energy. A 300 K blackbody spectrum taken using the photodetector can

---

<sup>4</sup>Detector element is unit number 2189 from Infrared Laboratories.

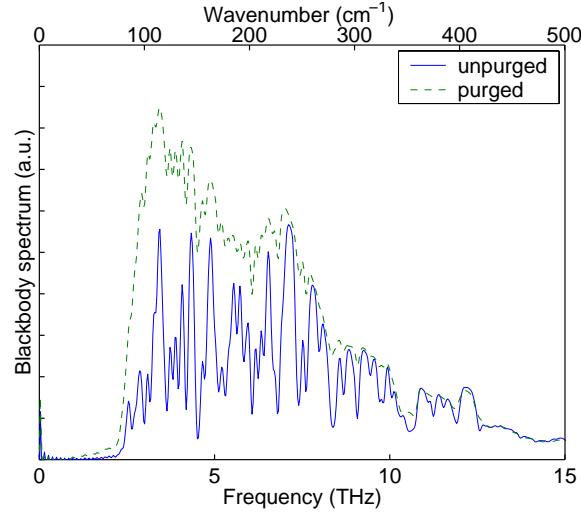


Figure 5-4: 300 K blackbody spectra taken with the Ga:Ge photodetector at  $4 \text{ cm}^{-1}$  resolution before and after purging of the optical path. The sharp edge at  $80 \text{ cm}^{-1}$  is due to the detector cut-off.

be seen in Fig. 5-4.

Once lasing was achieved, room temperature detectors such as pyroelectrics and thermopiles could be used to measure radiation. Both of these are broadband thermal detectors, whose operation is extended into the terahertz with the use of a semimetallic or blackened coating on the element. Pyroelectric detectors operate by detecting the change in a spontaneous electric polarization caused by temperature change. Chopped radiation was collected onto the 2-mm diameter element of the pyroelectric detector<sup>5</sup> by a condensing cone, and the detector signal was extracted using a lock-in amplifier. A thermopile power meter is essentially an array of thermocouple devices that generate a voltage for a heat flow corresponding to a given temperature difference. A thermopile operates only in dc mode, with a thermal response time on the order of seconds. As a result it is useful only for measuring average power. However, because it is sensitive to the environment, it is the noisiest of the detectors, and can only resolve power with a resolution of  $10 \mu\text{W}$ . Because of its larger element area of 25 mm diameter, this detector<sup>6</sup> was used to calibrate the absolute power level of the lasers.

---

<sup>5</sup>Model P4-42, Molelectron, Campbell, CA.

<sup>6</sup>Model 360203, Scientech Inc., Boulder, CO.

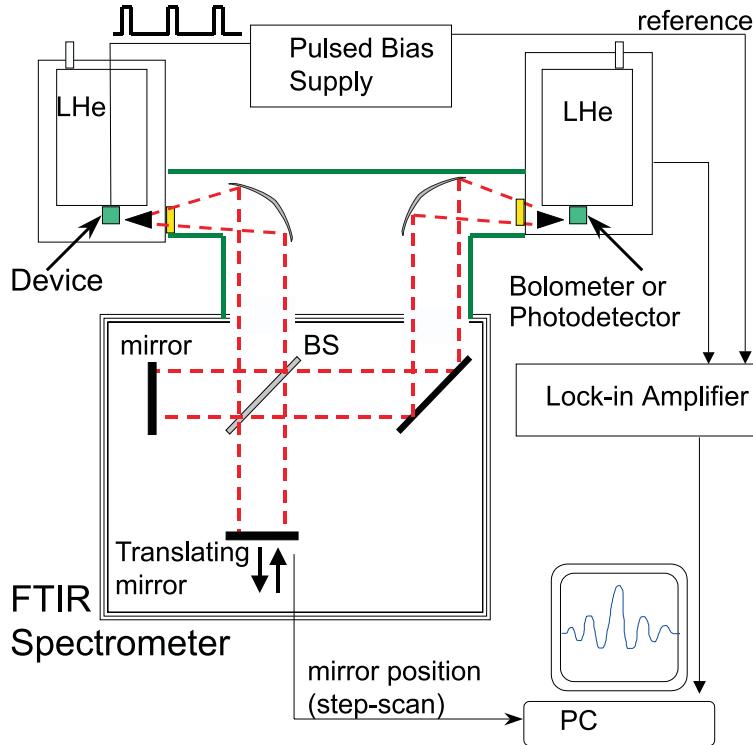


Figure 5-5: Experimental setup for emission measurements.

### Spectral measurements

The experimental setup for measuring spectra is shown in Fig. 5-5. The emission spectra were resolved using a Nicolet Magna-IR 850 Fourier transform infrared spectrometer (FTIR), as shown in Fig. 5-5. The FTIR operates by using a Michelson interferometer to trace out an interferogram, which is Fourier transformed to yield the power spectrum of the emitted radiation. Spontaneous emission measurements are performed in step-scan mode, where the translating mirror pauses at each step to allow processing of the small detector signal with the lock-in amplifier or the boxcar averager. This technique is essential for observing the small spontaneous emission power above background radiation. In general, the step time of the interferometer is set to be roughly one third of the lock-in integration time; this effectively low-pass filters the interferogram (i.e. suppresses the spectrum at high frequencies). However, this has a negligible effect on the frequencies of interest, and results in a shorter data collection time and reduced cryogen consumption. Laser emission experiments and

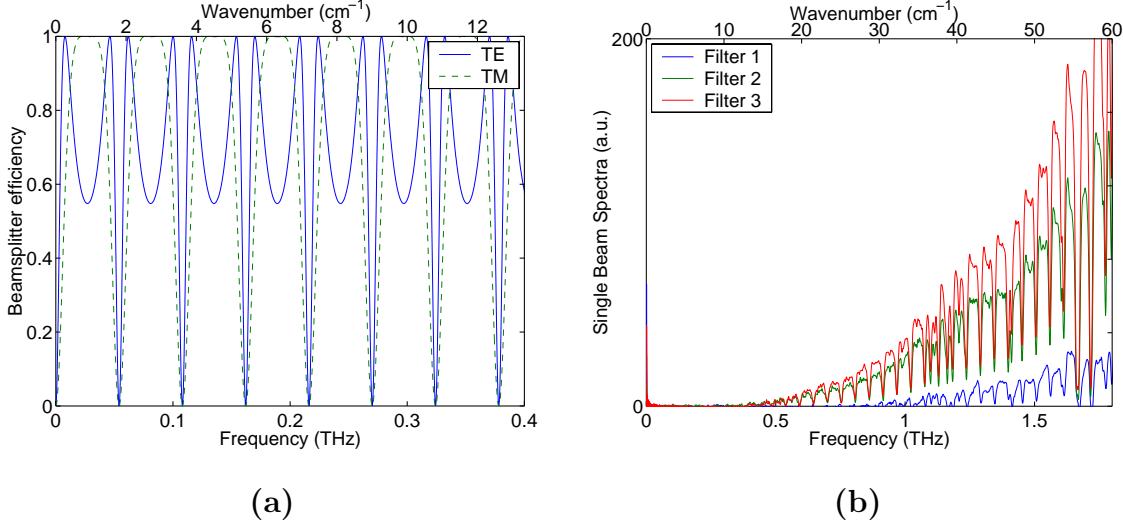


Figure 5-6: (a) Calculated beamsplitter efficiency ( $4R \times T$ ) for a 0.83-mm-thick silicon wafer at  $45^\circ$ . (b) Blackbody spectra taken with the bolometer at  $0.125 \text{ cm}^{-1}$  which display dips regularly spaced at  $1.8 \text{ cm}^{-1}$  due to the beamsplitter.

blackbody are performed in linear-scan mode, where the mirror translates smoothly while tracing the interferogram. When sufficient power is available, this method allows for the much faster collection of data. In this thesis, most displayed spectra have been processed with one level of zero padding using a Happ-Genzel windowing function and a Mertz phase transform[144].

Spectral measurements in the terahertz spectral range are complicated by the presence of strong atmospheric absorption, most of which is due to water. Not only does this reduce the total detected power, but the presence of pressure broadened (but still quite sharp!) absorption lines can make interpretation of spectra more difficult. Since it was beyond our capability to evacuate the spectrometer, nearly all spectral measurements were performed with the optical path purged with dry nitrogen gas. This can minimize, but not completely remove, the absorption lines. An example of a 300 K blackbody spectrum taken with the photodetector at a resolution of  $4 \text{ cm}^{-1}$  (0.12 THz) before and after purging the optical path is shown in Fig. 5-4. If spectra are taken with higher resolution, the lines become deeper and more distinct.

In addition, the beamsplitter efficiency is not flat, which can add complications to high resolution measurements, especially at very low frequencies (< 2 THz). The

Nicolet “Solid Substrate” far-infrared beam splitter is a polished high-resistivity silicon wafer. As a result, Fabry-Pérot resonances are present, and nulls can appear in the beamsplitter efficiency, which is shown calculated in Fig. 5-6(a) for a 0.83-mm-thick silicon wafer. This corresponds to a  $1.8 \text{ cm}^{-1}$  spacing free spectral range, where the thickness was chosen to agree with the observed spacing. The effect of these nulls can be seen in a single beam blackbody spectra taken with the spectrometer using the bolometer with three different filters (Fig. 5-6(b)). Regularly spaced dips can be clearly seen every  $1.8 \text{ cm}^{-1}$ , which cannot be simply attributed to absorption. At higher frequencies, the regular pattern of dips “washes out,” probably due to beam-splitter misalignment or imperfectly parallel surfaces. For low resolution spectra, a smooth average of beamsplitter response is obtained, but for very high resolution spectra (i.e. laser sources) spectral distortion may occur.

## 5.2 Fabrication

Our designs were grown by molecular beam epitaxy (MBE) first by Prof. Michael Melloch of Purdue University, and later by Dr. John Reno of Sandia National Laboratories. Either semi-insulating or silicon doped  $n^+$  ( $1 - 2 \times 10^{18} \text{ cm}^{-3}$ ) (100) oriented GaAs wafers were used. For MBE growth of terahertz quantum cascade structures, precision in the doping levels and barrier height is essential. Also, experience indicates that the sharpness of barriers is critical, especially in structures with very thin barriers (2 ML). Since the growths tend to be very long (up to  $10 \mu\text{m}$ ), long-time stability of the growth is necessary to prevent non-uniformities among modules.

All fabrication took place at MIT in the Technology Research Laboratory (TRL) at the Microsystems Technology Laboratory (MTL), with the exception of the lapping, which was performed at the Center for Materials Science and Engineering (CMSE). Fabrication proceeded via standard contact photolithographic techniques to create ridge and mesa structures. The exception is the more complicated metal-metal waveguide fabrication, which is described separately in Sec. 5.3. The most commonly fabricated structures were edge emitting ridges and surface emitting grat-

ing structures. Grating structures were most commonly used with devices grown on an  $n^+$  substrate, i.e. plasmon waveguide structures, because of high waveguide losses make edge coupling inefficient.

The most generic processing sequence proceeds as follows. Following MBE growth, the first step is to deposit the metallic non-alloyed ohmic contacts on the epitaxial layers. Because of the delicacy of the LTG-grown epilayer, this step is best performed first. Following an image reversal photolithography step, Ti/Au (200/4000 Å) layers are deposited via electron-beam evaporation, and patterned via lift-off in acetone. Then mesa structures are defined by either wet etching or dry etching through the epitaxial layers down to the substrate. In the case of wet etching, an additional photostep is performed first to provide a photoresist etch mask, whereas for dry etching, often the metal itself is used as a self aligned etch mask. The details of various etching processes are described below in Sec. 5.2.1. NiGeAu alloyed ohmic contacts are then made to the back of the  $n^+$  substrate.

The processing for semi-insulating surface plasmon waveguide devices (Fig. 4-6) is similar to the generic process, but involves additional steps to enable the use of lateral contacts. The etching must be carefully performed to reach the contact layer without etching through it. Both under- or over-etching can lead to parasitic resistance, or in a worst case, an open circuit. Following the etch, another photostep and lift-off metallization is performed to provide lateral NiGeAu alloyed ohmic contacts.

### 5.2.1 Etching

Wet etching is performed most often for mesa (ridge) etching, and is generally preferred over dry etching due to its reproducibility and ability to leave smooth surfaces. Primarily though, unlike dry etching, there is no worry about contamination or damage of the surfaces, which can lead to unpredictable electric properties. The main disadvantage is that wet etching will produce sloped sidewalls, which results in current spreading and bias non-uniformity from module to module.

The most common wet etchant used for mesa definition was NH<sub>4</sub>OH:H<sub>2</sub>O<sub>2</sub>:H<sub>2</sub>O (peroxide-ammonia (PA)), usually in the ratio of 20:12:480. The etchant is relatively

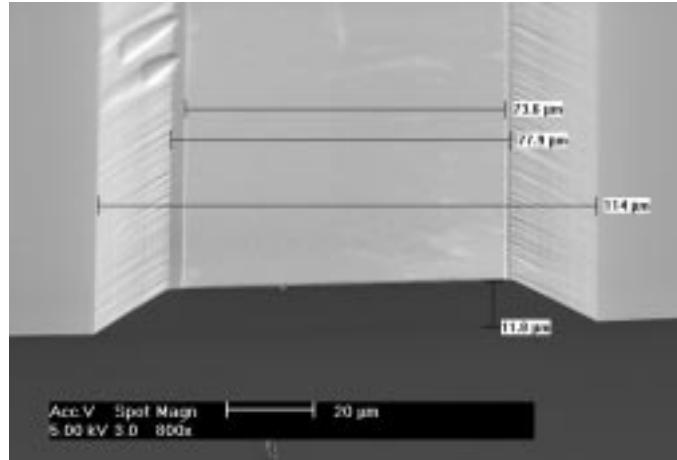


Figure 5-7: Typical SI-surface-plasmon ridge structure wet etched in  $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (20:12:480). Device is 12- $\mu\text{m}$ -high FSL120.

non-selective between GaAs and  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  for  $x < 0.4$ . It etches at approximately 400 nm/min and is generally reaction-limited, although it displays some diffusion limited traits such as trenching (see below). Outwardly sloped sidewalls are produced along the  $(01\bar{1})$  section (parallel to the major flat on US (100) wafers), and undercut sidewalls are produced along a  $(011)$  section. This sidewall slope depends somewhat on the wafer composition, i.e. devices with 15% Al barriers (i.e. FL175C) had slopes of  $\sim 40^\circ$ , and a device with 30% Al barriers (FL175B) had a slope of  $\sim 26^\circ$ , where the slope is the acute angle between the sidewall and substrate. The etch rate is fairly reproducible from run to run, although the etch depths might vary by as much as 10% over the area of a wafer piece. For some deep etches ( $\sim 10 \mu\text{m}$ ), sidewall waviness, or even large chips would sometimes appear (Fig. 5-7). This was likely related to photoresist adhesion, and possibly to the formation of microscopic bubbles under the photoresist during the etching. This phenomenon was not always a problem, and generally seemed to diminish when using more dilute etching solutions (such as 10:6:480). The best photoresist adhesion and overall results were obtained with Shipley 1813 photoresist.

The phenomenon of “trenching” is associated with many isotropic etch solutions where etching proceeds fastest adjacent to masked material. In our case, this leads

to deeper “trenches” immediately adjacent to ridge structures, the depth of which depends on the ridge width. Typically, for a 10- $\mu\text{m}$  etch, the region next to the ridge might be  $\sim 0.5 \mu\text{m}$  deeper than 50  $\mu\text{m}$  away from the ridge. This is only problematic for the SI-surface-plasmon waveguide fabrication, since one must deposit lateral contacts to the  $n^+$  contact layer 50  $\mu\text{m}$  away from the ridge, without etching through the contact layer.

This peroxide-ammonia etch also exhibits dramatically increased etching at metal-GaAs interfaces due to electrochemical enhancement [145, 146]. This effect prevents the use of metal for etch masks in wet processes, and requires metal surfaces to be covered by photoresist prior to etching. Sufficient overlap must be included to account for undercut.

The etchant  $\text{H}_3\text{PO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (1:1:25) was used much less frequently, but is potentially advantageous for wet etching QCL devices. An example SEM of this etch is shown in Sec. 5.3.3 for a metal-metal waveguide. This etch has the characteristic that it produces outwardly sloping sidewalls along both  $(011)$  and  $(01\bar{1})$  sections. Also, no evidence of trenching was observed, and there was much less variability in etch depth over the wafer piece area ( $\pm 0.25 \mu\text{m}$  over a 10  $\mu\text{m}$  etch). Steeper sidewalls were observed ( $\sim 48^\circ$ ), which reduces current spreading. Furthermore, when wet etching metal waveguide devices where the film is effectively upside down, this etch was far superior to the peroxide-ammonia etches, where dramatic photoresist undercut and “chipping” was observed.

However, this etch was observed to display relatively different etch rates from run to run. This requires close control over etch conditions, and careful calibration of the etch rate during each etch. Also, dramatically increased lateral etching at metal-GaAs interfaces was observed, perhaps even more so than with the peroxide-ammonia etch.

Dry etching was performed by electron-cyclotron-resonance reactive-ion-etching (ECR-RIE) in a Plasmaquest. This machine is fitted with a load lock, but is used for both etching and deposition, with both chlorines and fluorines. As such, chamber conditions were somewhat inconsistent, and long cleans and conditioning runs were used prior to each run.  $\text{BCl}_3:\text{N}_2$  (15:5 sccm) was used to etch deep mesas up to 12- $\mu\text{m}$

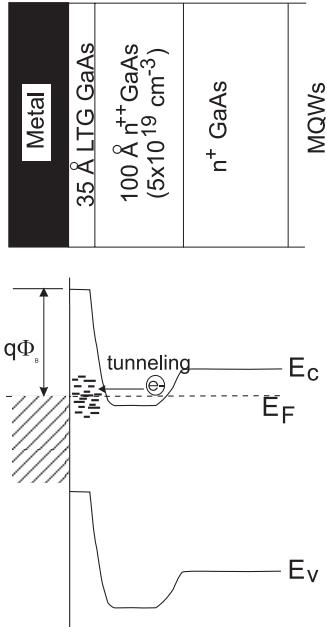


Figure 5-8: Schematic of band diagram for nonalloyed ohmic contact to  $n$ -GaAs.

deep, using Ti/Au (200/4000 Å) contacts as self aligned etch masks. The full recipe is given in Appendix B. Relatively vertical sidewalls were obtained, as is shown in the SEMs in Sec. 5.3.3. The sidewall roughness was not expected to be significant, because it was much smaller ( $\sim 1 \mu\text{m}$ ) than the radiation wavelength inside the semiconductor ( $\geq 15 \mu\text{m}$ ). The primary disadvantage of dry etching was the sometimes unreliable electrical characteristics of the resulting devices. This is assumed to take place due to deposition of contaminants from the chamber, or redeposition of sputtered mask material. A brief cleanup wet etch was sometimes used, but this is generally undesirable due to the rapid undercutting at the semiconductor-metal interface.

### 5.2.2 Non-alloyed ohmic contacts

Making ohmic contacts to  $n$ -GaAs has traditionally been difficult because the bulk doping density is limited to  $N_D - N_A \approx 5 \times 10^{18} \text{ cm}^{-3}$ . Hence alloyed ohmic contacts are used extensively, where a metal/dopant film is deposited and then annealed so that the dopants diffuse into the semiconductor. However, this type of contact is

highly undesirable for terahertz optical devices because of the potentially high free carrier loss caused by the uncontrolled diffused doping profile.

Therefore, whenever possible, low-temperature-grown (LTG) GaAs nonalloyed ohmic contacts of the type described by Patkar *et al.* in Ref. [147] were used. This method takes advantage of the very high doping levels ( $\sim 10^{20} \text{ cm}^{-3}$ ) that can occur in the thin GaAs layer at the surface due to the pinning of the Fermi level by surface states. However, usually this layer oxidizes quickly with exposure to air, and is removed with the oxide etch step that precedes metal deposition. To prevent this, 35 Å of LTG GaAs (grown at 250° C), which resists oxidation, is grown as a passivation layer. Due to the high concentration of defect states in LTG GaAs, the Fermi level remains pinned, and the surface doping remains high. Following a dip in BOE, Ti/Au layers (200/2000–4000 Å) are deposited. Using this method, contact resistances as low as  $\rho_c = 2.4 \times 10^{-7} \Omega \text{ cm}^2$  have been reported [147], although our results have not been as good.

### 5.3 Metal-metal waveguide fabrication

The fabrication of practical metal-metal waveguide structures is a non-trivial task, but has been successfully performed, and the first working terahertz quantum cascade lasers using such waveguide were demonstrated by the author and coworkers in March 2003 [70] (see Sec. 7.3). The advantages of the metal-metal waveguide in the terahertz spectral range were described in Sec. 4.5; any disadvantages lie in the additional fabrication complexity. In our group, the original fabrication of metal-metal waveguides using a gold-gold thermocompression bonding method was pioneered by Bin Xu [59] based on a method originally developed by Parameswaran [148]. While this method provided good results, there were problems with fabrication yield relating to bonding layer adhesion and mechanical strength of the waveguides. For example, wire bonding to devices would usually destroy them. Therefore, a fabrication method was developed that takes advantage of indium-gold reactive bonding. In this section, I will discuss the original Au-Au thermocompression bonding method, as well as the

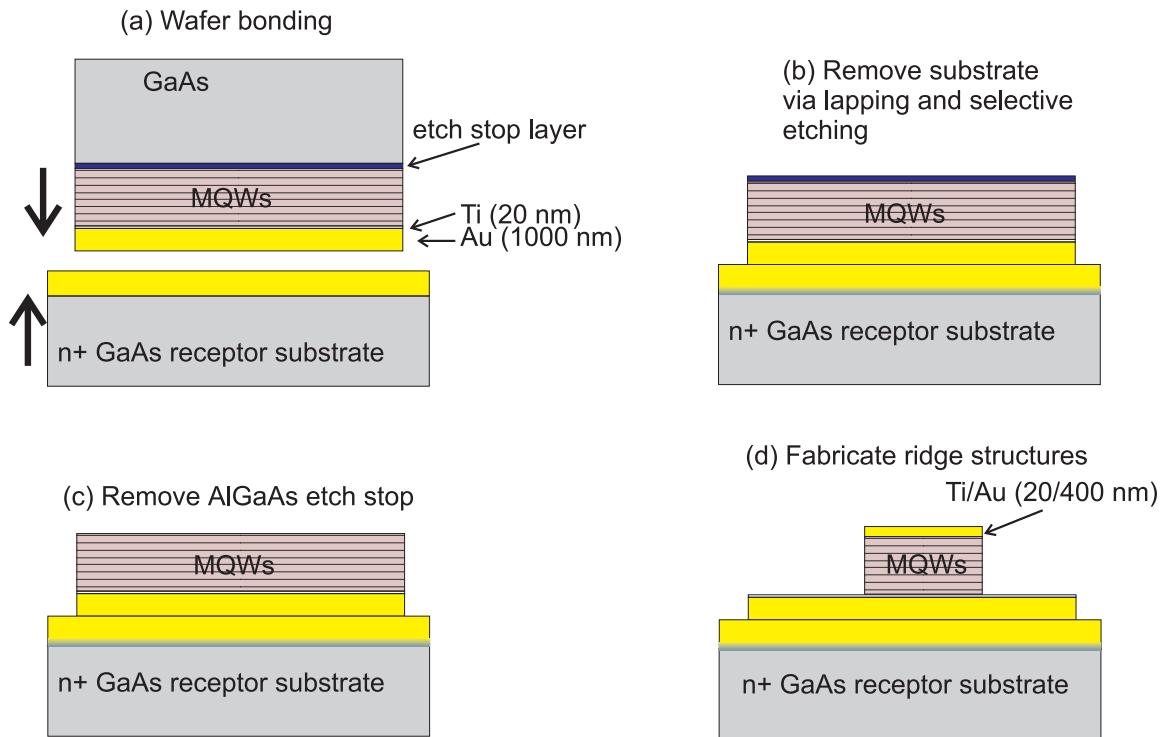


Figure 5-9: Schematic diagram metal-metal bonding process.

newly developed In-Au reactive bonding method.

The overall process for fabrication proceeds as follows:

1. Evaporate Ti/Au (200/10000 Å) on the MBE grown device wafer and evaporate the appropriate metallization on the  $n^+$  GaAs receptor wafer.
2. Bond wafers according to desired technique (see below).
3. Deposit  $\sim 2000$  Å SiO<sub>2</sub> on all sides and back of bonded wafer pair to protect edges from lateral etching.
4. Lap the device wafer until  $\sim 50\text{--}75$   $\mu\text{m}$  substrate thickness remains.
5. Remove the remaining substrate with NH<sub>4</sub>OH:H<sub>2</sub>O<sub>2</sub> (1:19) wet etch.
6. Remove the Al<sub>0.5</sub>Ga<sub>0.5</sub>As etch stop layer in HF. This will also remove the protective oxide.

7. The remaining thin film is consists of only the MBE grown active region. It can be processed into ridge structures using standard photolithographic techniques

This process flow is illustrated schematically in Fig. 5-9 and the details of the process are discussed below.

### 5.3.1 Wafer bonding

The initial preparation of the MBE grown device wafer is the same for either Au-Au thermocompression bonding, or In-Au reactive bonding. Wafers that are to be used for bonding are MBE grown with a 500-nm-thick  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  layer underneath the active region as an etch stop. Following a short  $\sim 10$  s buffered oxide etch (BOE) dip (BOE is an aqueous  $\text{NH}_4\text{-HF}$  solution, buffered for a constant pH), a Ti/Au (200/10000 Å) layer is deposited by e-beam evaporation on the device wafer. This Ti/Au layer forms a non-alloyed ohmic contact to the active region provided the MBE growth was capped with a low-temperature-grown (LTG) GaAs contact layer (see Sec. 5.2.2). The use of a non-alloyed ohmic contact is important, since the uncontrolled doping profile that is associated with standard annealed ohmic contacts to GaAs potentially adds significant waveguide loss. During all steps, great care should be taken to keep wafer surfaces clean, especially while scribing the wafers and afterward. Plastic tweezers should be used to minimize scratching of the metallic surfaces. When cleaving, nicks on the wafer edge should be made using single, clean motions, to avoid marking of the surface.

#### Au-Au thermocompression bonding

The process used for the Au-Au thermocompression bonding is described below. For further recipe details, see Ref. [59]. An  $n^+$  GaAs receptor wafer is coated with Ni/Ge/Au/Ni/Au (50/660/1330/400/10000 Å). The NiGeAu multilayer will anneal during bonding to form an alloyed ohmic contact with the receptor wafer. The  $\sim 1 \mu\text{m}$  thick Au layers on both wafers are necessary for successful wafer bonding, since the soft gold will deform to allow for any wafer surface or metal layer thickness non-

uniformities. Wafer pieces approximately  $1\text{ cm} \times 1\text{ cm}$  are cleaved from each wafer, blown with  $\text{N}_2$  to remove particulates, cleaned for 15 minutes in an UV ozone cleaner, and blown clean again. Clean surfaces are essential for successful bonding. The two wafer pieces are placed on a hot plate face to face and aligned along their crystal axes. A glass slide is placed on top of the stack so that pressure can be applied evenly. Bonding takes place at approximately  $370\text{--}380^\circ\text{ C}$  as pressure to the stack is applied by hand via a flat edged for approximately 10–15 minutes. The wafers are then allowed to cool. It is important not to overheat the wafers, as the LTG-GaAs non-alloyed ohmic contact begins to degrade above  $400^\circ\text{ C}$  [149].

### In-Au reactive bonding

Following evaporation of Ti/Au on the MBE grown device wafer, the metal sequence Pd/Ge/Pd/In/Au ( $250/100/250/12000/1200\text{ \AA}$ ) is evaporated on a  $n^+$  GaAs receptor substrate (see Fig. 5-10). The purpose of the Pd/Ge/Pd multilayer is to improve the electrical contact to the receptor substrate [150, 151]. Since this bonding process does not use temperatures above  $250^\circ\text{ C}$ , a traditional NiGeAu alloyed contact cannot be used. Palladium in particular is an attractive candidate for low temperature contacts since it is highly reactive at room temperature, and displaces the native oxide at the GaAs surface to form  $\text{Pd}_4\text{GaAs}$  [152]. The resulting contact resistance is unknown, and it is even likely that an ohmic contact is not achieved, since for most reported PdIn or PdGeIn contacts to  $n^+$  GaAs, annealing above  $400^\circ\text{ C}$  is required. However, due to the large area of this contact compared with the device area, an true ohmic contact is not necessary.

Wafer pieces approximately  $1\text{ cm} \times 1\text{ cm}$  are cleaved from each wafer, blown with  $\text{N}_2$  to remove particulates, cleaned for 15 minutes in an UV ozone cleaner, and blown clean again. The two wafer pieces are placed on a hot plate face to face and aligned along their crystal axes. A glass slide is placed on top of the stack and pressure was applied by hand using a flat tool. The temperature was raised to  $250^\circ\text{ C}$  over the course of approximately 10 minutes, while pressure was applied to various parts of the wafer stack. The wafers are then cooled to room temperature over the course of

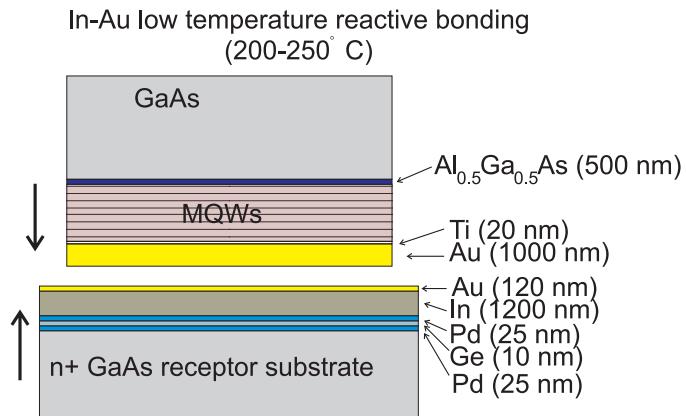


Figure 5-10: Schematic diagram deposited metal layers for In-Au reactive bonding.

20 minutes.

Bonding occurs as the indium melts at 156° C, and flows laterally to fill voids and encase any submicron particles present on the bonding surfaces. The liquid indium diffuses into the gold layer, where it quickly reacts and forms AuIn<sub>2</sub> and a variety of other Au-In alloys [154, 155]. Because the melting points of the resulting eutectic alloys are greater than 450° C (see Fig. 5-11), the resulting bond will be stable at very high temperatures. This is in contrast to other schemes where solder is used to bond wafers, in which the bonding temperature must be above the solder melting point, while subsequent processing temperatures are limited to below the melting point of the solder. Also, because of the role of the molten indium, this method has much less stringent requirements on the cleanliness and flatness of the bonding surfaces than the Au-Au thermocompression method does. Ideally, the layer thicknesses are chosen so that all the indium is consumed, leaving no low melting point material. Also, it is preferable that the gold next to the active region remain unconsumed, to prevent any increase in waveguide losses associated with the lower conductivity Au-In alloys.

One difficulty with use of indium is its propensity to oxidize in atmosphere. For this reason a 1200-Å-thick Au layer was evaporated above the 1.2-μm-thick In layer before vacuum was broken, as suggested in Ref. [154]. Indium is sufficiently reactive that the gold layer is consumed almost immediately upon deposition, and the surface appears a light-gray color, as well as being somewhat rough. This is consistent

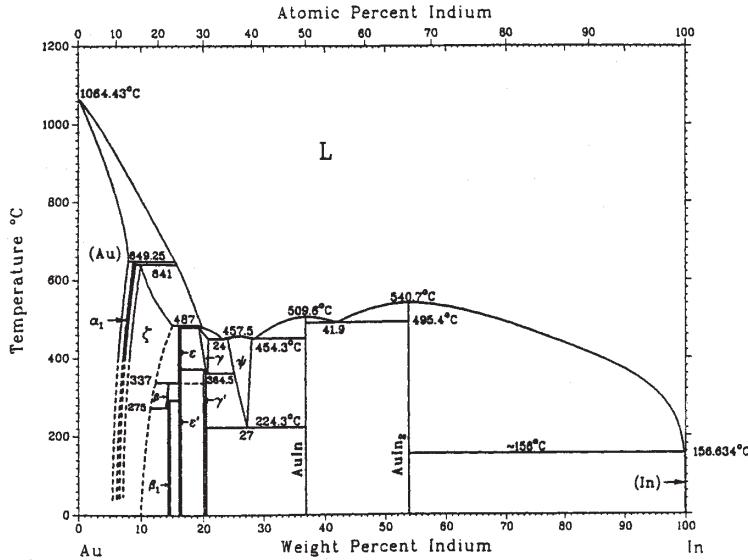


Figure 5-11: Gold-indium equilibrium phase diagram [153]

with the results reported in Ref. [154], where the  $\text{AuIn}_2$  formed immediately under similar conditions. However, they report that this compound is stable and able to protect the surface from oxidation. Ideally, the bonding should be done under an inert atmosphere to reduce oxidation as well. As the indium melts, this solid surface layer is broken up which forms a mixture of liquid and solid, which is incorporated into the bond layer.

### 5.3.2 Substrate removal

After wafer bonding, the next step is to remove the substrate of the device wafer via lapping and selective etching. First, the front and back of the bonded wafers are coated with  $\sim 2000 \text{ \AA}$  of  $\text{SiO}_2$  using low temperature plasma-enhanced chemical vapor deposition (PECVD). This helps to protect the back of the receptor wafer and the sides of the device wafer during the substrate removal etch. If this step is neglected, lateral undercut of the active region beneath the etch stop layer will occur. The advantage of applying oxide early in the process is that the lapping of the substrate automatically removes the oxide from the wafer top, leaving it only on the sides as desired. The wafers are then mounted with clear wax on a chuck and the substrate

is lapped using decreasing grit sizes until approximately  $\sim 50\ \mu\text{m}$  is remaining, as measured by a micrometer. After solvent cleaning in ultrasound, the backside of the receptor wafer is coated with photoresist (Shipley 1813) and hard-baked to further protect it during etching. One can optionally coat the sides of the device wafer as well, but it is extremely difficult to do this without also getting photoresist on top of the device wafer. This impedes etching and often leaves a “wall” around the outside of the film. For this reason, photoresist was not typically used to protect the sides.

The selective etchant  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2$  (1:19) is used to remove the GaAs substrate at an etch rate of approximately  $1\text{--}2\ \mu\text{m}/\text{minute}$ . When the  $0.5\text{-}\mu\text{m}$   $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  layer has been reached, the surface becomes mirror-like. Upon exposure to air, the etch stop layer quickly oxidizes, and colorful reflection fringes are evident. The exact selectivity of this etch for GaAs over  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  is unknown, however selectivities of up to 493 [156] and 1000 [157] have been reported for etching GaAs over AlAs.<sup>7</sup> The selectivity over  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  is almost certainly less. Furthermore, the selectivity is reported to be somewhat sensitive to the pH of the solution. However, when a  $0.5\text{-}\mu\text{m}$  pure AlAs layer was used, the result was unsatisfactory. The etchant oxidized the AlAs layer, which became thicker, built up stress, and cracked, which allowed the etchant through [159]. However, the stress buildup might have been less of a problem had a thinner AlAs layer been used. Although it was widely used, the 500-nm thickness of the  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  etch stop layer is somewhat of an overkill, and layers as thin as 100 nm have been used for successful substrate removal in our most recent fabrication.

Following substrate removal, the protective photoresist is removed in acetone. Then the  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  etch stop layer is removed by immersion in 48% HF acid for 1–2 minutes. Damage to the GaAs is negligible due to the extremely high selectivity ( $\sim 10^7$ ) for  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  with  $x > 0.45$  over GaAs [160]. This etch also removes the protective oxide layer from the back and sides of the bonded wafers.

---

<sup>7</sup>Citric acid: $\text{H}_2\text{O}_2$  solutions have been reported to have an even higher selectivity (up to 1450 GaAs over AlAs), and are widely used due to their flexibility and ability to selectively etch GaAs over  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  layers down to  $x = 0.15$  [158]. However, this solution was only rarely used in this research, as it was difficult to obtain in TRL, and etched GaAs at a considerably slower rate.

Now, all that remains is the several micron thick epitaxial layer bonded to the receptor wafer, which can be processed with standard photolithographic methods. Because the epitaxial film is now oriented upside down, it is impossible to use a LTG-GaAs non-alloyed ohmic contact on top as would usually be done. Since using an alloyed ohmic contact is highly undesirable due to additional optical loss, usually a Schottky diode contact is made simply by depositing Ti/Au on the heavily doped GaAs contact layer. Since this Schottky contact will be forward biased, the only effect on device operation is to add an extra 0.5–0.7 V voltage drop at the contact.

For metal-metal waveguides, the most common processing is a self-aligned dry etch process, where laser ridge structure are lithographically defined using evaporated Ti/Au (200/4000 Å) Schottky contacts in a lift-off step. The metal contacts are then used as an etch mask for a Plasmaquest (ECR-RIE) etch using  $\text{BCl}_3/\text{N}_2$  at 15:5 sccm. However, wet etching can be used as well; the best results have been obtained with the  $\text{H}_3\text{PO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (1:1:25) solution and Shipley 1813 photoresist. When etching, it is important to remember that since the film is upside down, the crystal axes are oriented opposite to what they would normally be. This will affect the direction of outwardly sloping sidewalls, i.e. ridges should be parallel to the major flat for US oriented wafers. Also, this etch should be stopped just before the metal layer is reached, since rapid lateral etching at the GaAs-Ti/Au interface can undercut the ridge. After a backside Ti/Au metallization, no further processing steps are necessary, although occasionally a silicon nitride layer is applied to the sidewalls of the ridge structures to prevent electrical shorting.

### 5.3.3 Metal-metal fabrication results

#### Au-Au thermocompression bonding

Au-Au waveguides were successfully fabricated, and many such devices were successfully tested, although no lasing was achieved due to lack of a adequate gain medium. Scanning electron micrographs of typical Au-Au thermocompression bonding waveguides are shown in Fig. 5-12. Dry etching was used, which is the cause of the roughness

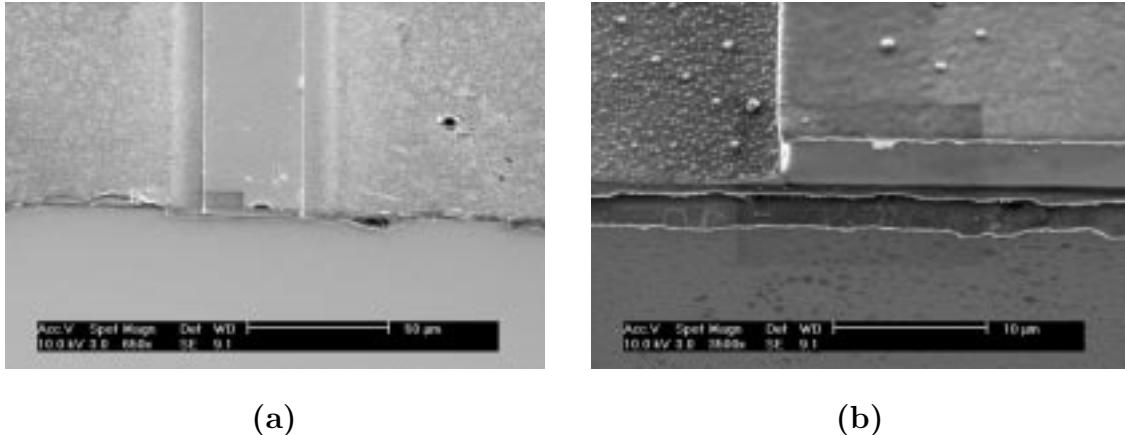


Figure 5-12: SEM micrographs of typical metal-metal waveguide structures fabricated by Au-Au thermocompression bonding and dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm) (ECR-RIE). Active region is  $\sim 7 \mu\text{m}$  thick. Device is T150Emm.

on the sidewalls and floor. Examples of metal-metal waveguides fabricated using Au-Au thermocompression bonding are T65mm (Sec. 6.4.1), T150E-2mm (Sec. 6.4.3), T150Fmm (Sec. 6.4.4), and FL125mm (Sec. 6.5.1). In many of these devices, such as T150E and T150F, the observed spontaneous emission power from metal-metal waveguides was a factor of two to four times greater than from plasmon waveguides, which sometimes allowed edge emitting operation where it was not previously possible. This is an indication that the metal-metal waveguides are much less lossy than the plasmon waveguides, as expected.

Some of the earlier fabrication runs had inconsistent *I-V* characteristics, often displaying several volt parasitics. This is attributed to the processing sequence, where the top Schottky contact was evaporated after the mesa was etched instead of before. It is likely the surface was contaminated during the additional photostep, which resulted in a poor contact. In later fabrication runs, such as T150E-2mm, T150Fmm, CSL65mm, and FL125mm, the Schottky contact was evaporated immediately, and there is no evidence of excessive contact resistance beyond a  $\sim 0.5$  V diode drop.

Facet cleaving was not observed to be a problem, provided the crystalline axes of the bonded wafers were reasonably well aligned, and the Au-Au layers were well adhered. Because gold is a soft metal, when a wafer was cleaved the gold tended to

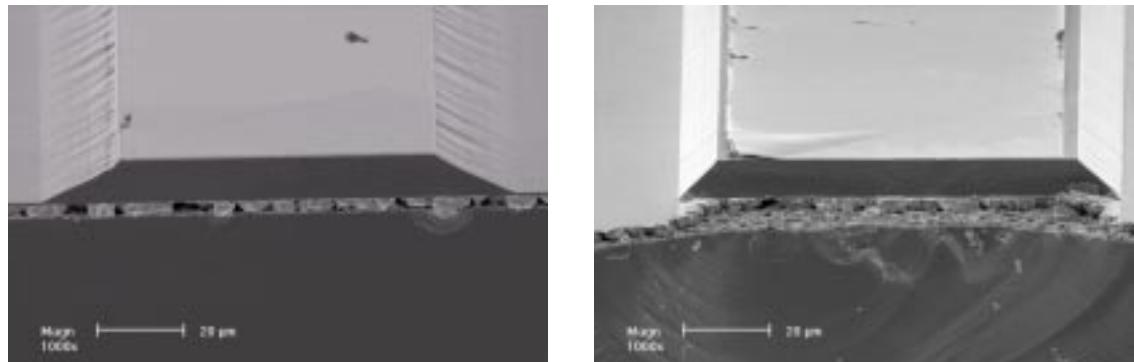
tear at the interface. Sometimes slight cracking in the epitaxial layers was observed over a few microns at the leading edge of the facet, before the cleave could stabilize. Other times, the epitaxial cleave was set back a few microns from the substrate cleave, as can be seen in Fig. 5-12(b). If the layer adhesion was poor, the cleaves might be as far as 50–100  $\mu\text{m}$  apart, resulting in a ridge with substantial overhang or under-hang from the substrate edge.

The largest drawback to the Au-Au thermocompression bonding method was the inconsistent adhesion. The quality of the bond is highly dependent on the quality of the metal, as well as the flatness and cleanliness of the surfaces. This affected the yield of the devices, as well as the thermal resistance at the bonding interface. Also, wire bonding directly to devices was usually impossible. Either the bonding tip would crack the rigid epitaxial layer as it “floated” on the 2- $\mu\text{m}$ -layer of soft gold, or the bond would pull apart the device at the bonding interface if the adhesion was poor. Instead, either bonding paths were used, or gold wires were epoxied to the top of the mesas. Despite these drawbacks, it is possible that the bonding quality could be improved by using the higher pressures and controlled atmospheres available with a bonding machine rather than a hot plate.

### In-Au reactive bonding

Metal-metal waveguides were successfully fabricated using In-Au reactive bonding. Ridge structures fabricated using wet and dry etching are shown in Figs. 5-13,5-14 respectively. The result of this process was a very mechanically robust waveguide that did not suffer from any adhesion problems. In fact, the first terahertz QCL fabricated in a metal-metal waveguide used this fabrication technique (Sec. 7.3). Unlike the Au-Au waveguides, these devices were able to survive many wire bonds without damage. Electrical characteristics were very good, with no evidence of excessive parasitic resistances, other than the Schottky voltage drop at the top contact. In general, this process was considered a success.

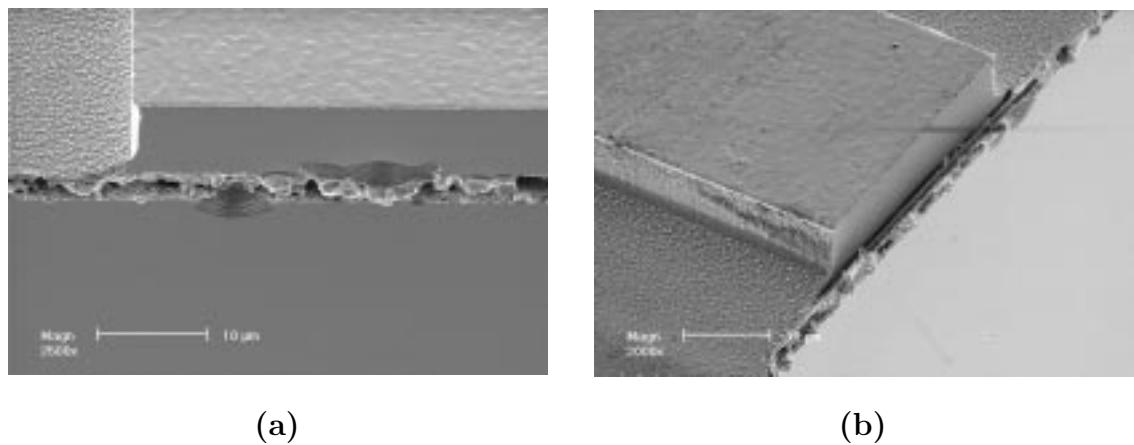
When the epitaxial layer was etched away, it was revealed that over most of the area, especially the center of the die, the metal was a grayish color rather than gold.



(a)

(b)

Figure 5-13: SEM micrographs of typical metal-metal waveguide structures fabricated by In-Au reactive bonding bonding and wet etching by (a)  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (10:6:480) (device is FL152Fmm) and (b)  $\text{H}_3\text{PO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (1:1:25) (device is FL162Gmm). Active regions are 10- $\mu\text{m}$  thick.



(a)

(b)

Figure 5-14: SEM micrographs of typical metal-metal waveguide structures fabricated by In-Au reactive bonding bonding and dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm) (ECR-RIE). Active region is 10  $\mu\text{m}$  thick. Device is FL137HMM.

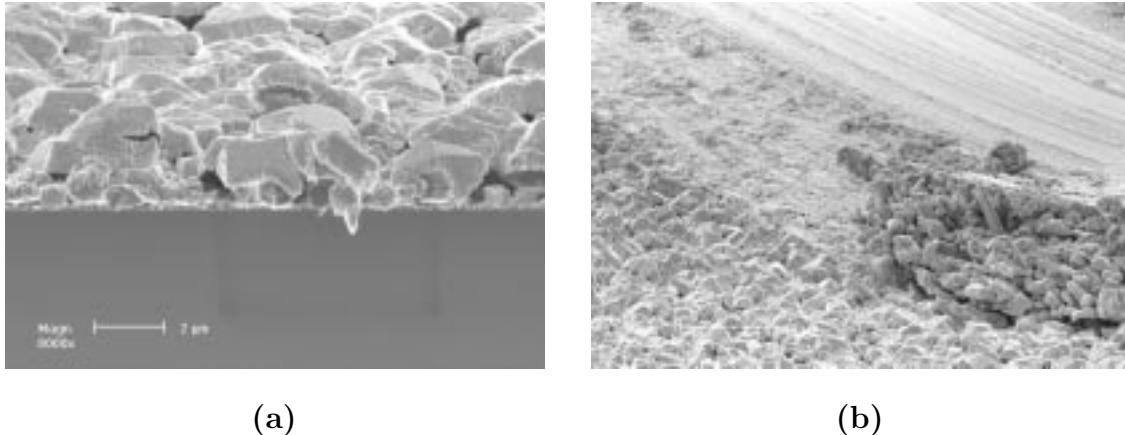


Figure 5-15: (a), (b) SEM micrographs of In multilayer coated receptor wafer prior to bonding. The wafer in (b) has been scratched with tweezers, which apparently reveals smooth In underneath the “rocky” In-Au alloys.

This indicated that the indium had consumed all of the pure gold, and the lower waveguide was composed of AuIn alloys. In fact, the places where some gold was still evident, i.e. the edges, corresponded with regions where the bonding and heat sinking was poor. Although it was very strong, the alloy itself appeared to be brittle rather than ductile in nature, and one could not wire bond to it. Examination of the SEM micrographs reveals that the bonding layer seemed to form clumps, possibly leaving some open space at interface. While facets were generally able to be cleaved well, occasionally chunks were taken out of the facet usually attached to one of the clumps in the bonding layer. In this case, this might be a case of the bonding layer being too strong, as the cleave would sometimes follow the structure of the bonding layer clumps rather than the crystalline axis of the epitaxial layers. Still, these problems were not significant enough to prevent lasing in FL152Fmm.

Also, there is evidence that the demonstrated metal-metal waveguides may introduce a significant thermal resistance at the bonding layer compared to SI-surface plasmon waveguides that are grown directly on SI GaAs substrates. There are two indications that this is a problem. First, the maximum lasing pulse width of FL152Fmm devices was roughly two microseconds for the best devices, much shorter than for FL175C devices, where lasing over hundreds of microseconds was observed. The las-

ing is quenched due to the temperature rise in the active region, which possesses a small thermal volume. Second, the maximum lasing pulse widths decrease rapidly and the threshold current densities increase rapidly for devices cleaved from outside the wafer piece center. This is well correlated with the observation that the bonding quality is the best in a  $\sim$ 5-mm diameter region in the center of the piece, and degrades further outside. To summarize, while the short pulse width is partly due to the intrinsic weak gain in the FL152F device, the correlation with bonding quality indicates that heat removal is impeded by the bond. Such a phenomenon has been reported Ref. [137] for a mid-infrared QCL in a metal waveguide. Examination of the SEM pictures supports this conclusion, as the bonding layer seems to have intermittent air gaps. SEM micrographs suggest that some of the clumped structure is present on the receptor wafer before bonding (see Fig. 5-15), which suggests the Au cap layer violently reacted with the indium on contact. Improvement of the receptor wafer metal morphology should be improved, perhaps by omitting the Au cap layer altogether.

In conclusion, viable metal-metal waveguides have been fabricated via both techniques, but the In-Au reactive bonding method produces more mechanically robust waveguides with a higher yield. Such a waveguide was used to obtain lasing in a terahertz QCL at 3.0 THz, which did not otherwise lase in a SI-surface plasmon waveguide structure (see Sec. 7.3), which confirms its expected advantage at low frequencies (Sec. 4.5.4). Also, a record high operating temperature of 137 K was observed in another device with a metal-metal waveguide, whereas operation only up to 92 K was obtained using a device from the same wafer with a SI-surface-plasmon waveguide (see Sec. 7.6). However, further improvement in fabrication is still needed to improve yield and improve the quality of the bonding interface. The metallurgy of the AuIn alloyed bonding layer should be improved, and a larger number of processing parameters, such as bonding temperature and time should be optimized. It is expected that using a bonding machine instead of a hot plate will improve quality, since higher and more uniform pressures are attainable, and an inert atmosphere will prevent oxidation of indium during bonding. It is likely that the Au-Au thermocompression

bonding process would benefit from the use of a bonding machine as well. Indeed, since the thermal conductivity of Au is better than GaAs for  $T > 100$  K, metal-metal waveguides will become increasingly advantageous for higher temperature operation.

# Chapter 6

## Survey of early designs

### 6.1 Introduction

This chapter details many of the early device designs that were designed and tested along the way to the development of a terahertz quantum cascade laser. None of the devices described in this chapter lased; discussion on those devices is deferred until Chap. 7. Despite their failure to lase, a great deal of information was gleaned about intersubband transport, emission, and waveguiding in terahertz quantum cascade devices.

A schematic overview of the device families described in this thesis is shown in Fig. 6-1. The M-series and L-series of devices are both three well designs where the depopulation of the lower radiative state takes place via interwell (diagonal) LO-phonon scattering. For the M-series, the radiative transition was more diagonal (smaller oscillator strength, longer upper state lifetime), and for the L-series, the radiative transition was vertical (larger oscillator strength, shorter upper state lifetime). These devices evolved into the FL and FFL series of devices, where resonant tunneling in combination with LO-phonon scattering was used to selectively depopulate the lower radiative state while maintaining a long upper state lifetime. The T-series of devices simply consisted of a two-well module with an intrawell radiative transition. In this device, there is no deliberate LO-phonon scattering event. Rather, a population inversion requires a cold electron distribution in the lower doublet. Several

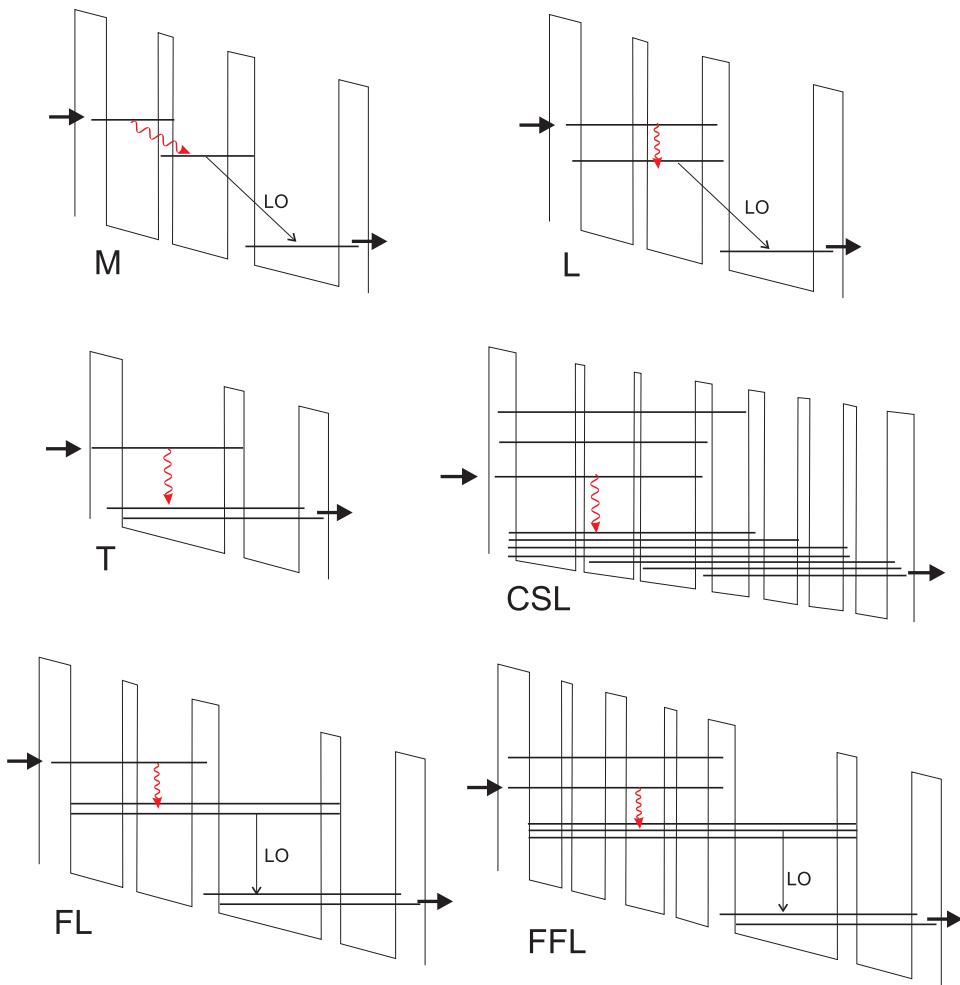


Figure 6-1: Schematic diagrams of the various multiple quantum well designs tested in this thesis.

copies of the first terahertz laser [64], a chirped superlattice design (CSL), were also tested.

Note that in this chapter that layer thicknesses are sometimes referred to in units of monolayers (ML). In the GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As material system, 1 ML=2.825 Å.

Over the course of this research, different values of the barrier height were used for various designs, and this is reflected in some of the calculations reported in this chapter. In the GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As material system, the barrier height of interest to us is the conduction band offset  $\Delta E_c$  (at the  $\Gamma$ -point), expressed as a fraction of the total bandgap difference between the well and barrier materials  $\Delta E_g$ , where  $\Delta E_g = 1.247x$  eV, provided  $x < 0.45$  [161]. The accepted value of the conduction band offset has changed substantially over the years, and even now there is some uncertainty in the correct value, as is described in the review article by Yu *et al.* [162]. The early accepted value of  $\Delta E_c \approx 0.85\Delta E_g$  gave way during the 1980's to a value of  $\Delta E_c \approx (0.6 - 0.7)\Delta E_g$ . Accordingly, throughout much of the early work described in this thesis, a 65% offset was used. However, we began to use an 80% offset for our calculations after certain experimental results were more easily explained by the larger barrier height (see Sec. 6.5.1). The use of the larger height was encouraged by the use of larger heights by other QCL groups, such as 80% as used by Faist's group [163]. Sirtori's group uses the formula  $\Delta E_c = 0.894x$  eV, which corresponds to a 72% band offset [48, 164]. In retrospect, the decision to use an 80% barrier height may have been premature, especially since the samples which indicated the use of a larger barrier height were not controlled for possible thickness deviations from design. A lower value such as 72% may be more appropriate, and may be a safe compromise between the currently accepted values and the values currently in use by the QCL community. Choice of the proper barrier height can affect such critical design parameters such as subband alignments, and anticrossing gaps between levels. In the next two chapters, the particular conduction band offset percentage that was used for a given calculation is always explicitly stated.

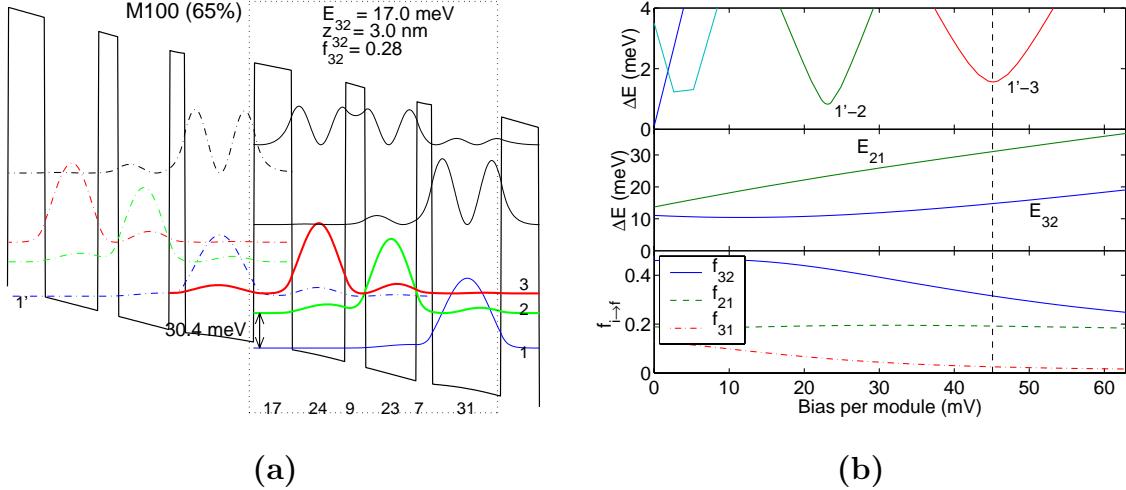


Figure 6-2: (a) Self-consistent conduction band profile of M100 calculated using 65% conduction band offset. The device is grown in the GaAs/Al<sub>0.3</sub>Ga<sub>0.7</sub>As material system. The layer thicknesses are given in monolayers. The 7 ML collector barrier was designed to be delta doped at a sheet density of  $8.8 \times 10^{10} \text{ cm}^{-2}$  per module. (b) Anticrossings, energy separations, and oscillator strengths calculated for M100 (without self-consistency).

## 6.2 M series

### 6.2.1 M100

The M-series of devices, including M100, was designed by Bin Xu, and is discussed extensively in his thesis [59] and in Ref. [60]. However, a brief discussion of this original device serves as a natural starting point for the evolution of terahertz QCLs that use LO-phonon scattering for depopulation. Furthermore, some M100 devices were retested using improved heat sinking on a cold plate with indium solder. This change reduced the amount of blackbody radiation in the spectra.

The M100 device is based around 100 three-well modules which are cascaded together to form a 3.1-μm-thick active region. The conduction band diagram for M100 is shown in Fig. 6-2(a) (calculated with 65% band offset). Electrons collect in the collector/injector state  $n = 1'$  and resonantly tunnel into the upper radiative state  $n = 3$ . The radiative transition is interwell (spatially diagonal), and takes place between states  $n = 3$  and 2, with a calculated oscillator strength of  $f_{32} = 0.28$ .

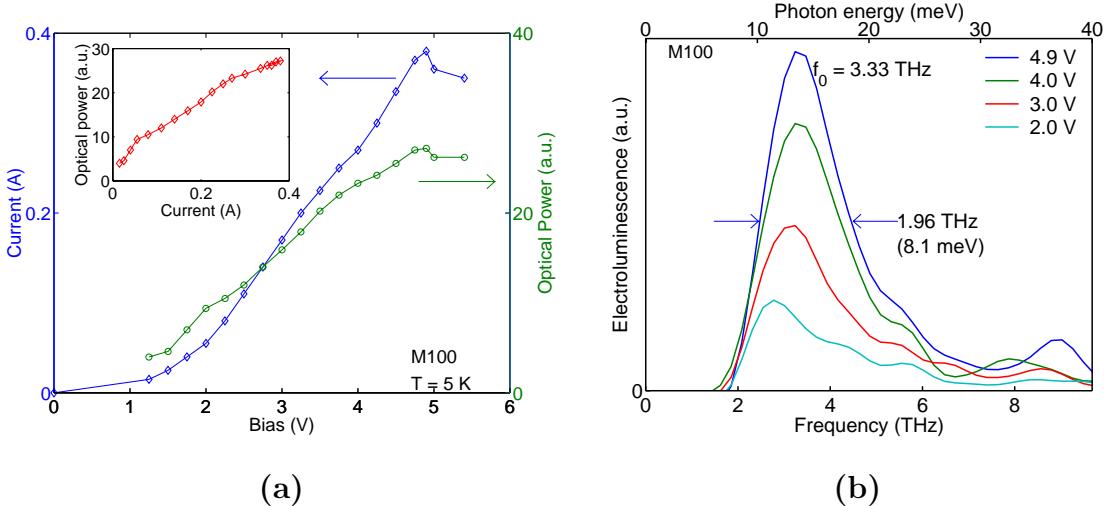


Figure 6-3: (a) Current and optical power versus voltage at 5 K for M100, taken at 50% duty cycle with a 40 kHz pulse repetition frequency. Inset gives  $L$ - $I$ . (b) Spectra at 5 K measured using the photodetector with  $32 \text{ cm}^{-1}$  (0.96 THz) resolution. The device is  $200 \times 400 \mu\text{m}^2$  with a 15- $\mu\text{m}$ -period grating with the conductive cap layer etched in between the grating.

Since the transition is diagonal, the  $E_{32}$  energy separation is very sensitive to applied electric field due to the Stark effect, but at design bias the separation is  $\sim 17$  meV. The energy separation  $E_{21} \approx 30$  meV is designed to be close to the LO-phonon energy  $E_{LO} = 36$  meV, so that fast LO-phonon scattering can depopulate level 2. Electrons with sufficient kinetic energy to emit an LO-phonon scatter with a time of  $\tau_{21} \approx 1$  ps. An examination of scattering rates for this structure that accounts for the complex phonon spectra is given in Chap. 3. The calculated anticrossings, energy separations, and oscillator strengths are shown in Fig. 6-2(b). The injection anticrossing gaps are  $\Delta_{1'2} = 0.83$  meV and  $\Delta_{1'3} = 1.5$  meV. The design bias occurs at approximately 45 mV/module when calculated without self-consistency, but can vary by several mV if self-consistency is included. The device was delta doped in the 7 ML barrier at  $8.8 \times 10^{10} \text{ cm}^{-2}$  per module.

This device was grown on an  $n^+$  substrate which serves as a plasmon waveguide. Surface emitting grating structures were fabricated to couple radiation out of the lossy cavity ( $\alpha_w \approx 1100 \text{ cm}^{-1}$ ,  $\Gamma = 0.76$ ).  $I$ - $V$  and  $L$ - $V$  characteristics taken in pulsed mode are shown in Fig. 6-3(a). A region of negative differential resistance

(NDR) occurs at 4.9 V as the injector level  $1'$  becomes misaligned with the upper state  $n = 3$ , preventing current flow. The NDR occurs at a bias that corresponds to 49 mV/module, which is close to the calculated design value of 45 mV/module.

The emitted output power level was extremely low, likely due to the lossy cavity and the small oscillator strength of the transition. Several spectra are displayed in Fig. 6-3(b), but because of extremely low signal level, only  $32 \text{ cm}^{-1}$  (0.96 THz; 4 meV) resolution could be used. The spectra consisted of a broad peak (FWHM = 8 meV) centered at approximately 3.3 THz ( $\sim 14$  meV), that displays a slight Stark shift with increasing bias. However, because of the breadth of the peak and the poor spectral resolution, it is impossible to rule out the presence of blackbody emission. Also, the low energy side of the peak is close to the  $80 \text{ cm}^{-1}$  cutoff for the photodetector. The broad linewidth of this transition is characteristic of interwell (diagonal) transitions, which are more subject to broadening due to interface roughness scattering than intrawell transitions. Also, the emission frequency is a sensitive function of bias, so module-to-module growth non-uniformities and bias inhomogeneities can contribute significant inhomogeneous broadening. The extremely broad linewidth and weak emitted power suggest that a more vertical transition with a larger oscillator strength is a better candidate for successful emission and lasing.

## 6.3 L series

### 6.3.1 L30

The L30 structure is based around a core three-level system constructed from a GaAs/Al<sub>0.3</sub>Ga<sub>0.7</sub>As triple-quantum-well structure, as shown in Fig. 6-4(a). This structure was originally designed by Bin Xu [59], but was retested under improved experimental conditions, i.e. in a cold-plate cryostat with improved heat sinking. The calculated anticrossing gaps, oscillator strengths, and energy separations are shown in Fig. 6-4(b). This device is similar to the structure M100, which featured a diagonal interwell radiative transition with an oscillator strength of  $f_{32} = 0.27$  [60, 59]

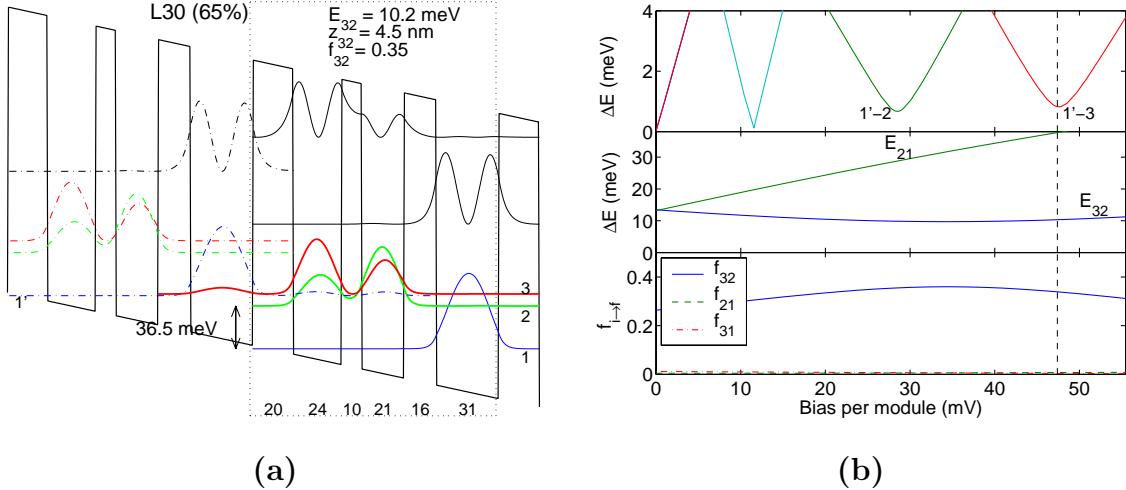


Figure 6-4: (a) Conduction band profile of L30 calculated using 65% conduction band offset. The device is grown in the GaAs/Al<sub>0.3</sub>Ga<sub>0.7</sub>As material system. The layer thicknesses are given in monolayers. The 16 ML collector barrier was designed to be delta doped at a sheet density of  $6.0 \times 10^{10} \text{ cm}^{-2}$  per module, although C-V measurements indicate a doping of  $8.4 \times 10^{10} \text{ cm}^{-2}$  per module. (b) Anticrossing plot, energy separations, and oscillator strengths for L30. All data was calculated using 65% offset, no self-consistency.

(see Section 6.2.1). For L30, The  $3 \rightarrow 2$  radiative transition was designed to take place between a pair of anticrossed states ( $\Delta_{32} = 9.7 \text{ meV}$ ), so that the transition is largely vertical, and has a larger oscillator strength  $f_{32} = 0.37$ . At a design bias of  $\sim 47 \text{ mV/module}$ , the ground state of a previous module  $n = 1'$  is aligned with  $n = 3$ , and electrons populate the upper level of the three-level system via resonant tunneling. The injection barrier is 5.6-nm thick (20 ML), which gives injection anticrossings  $\Delta_{1'3} = 0.8 \text{ meV}$  and  $\Delta_{1'2} = 0.67 \text{ meV}$ . The collector barrier is extremely thick (4.5 nm or 16 ML), which results in a slow phonon depopulation, calculated to be  $\tau_{21} = 30 \text{ ps}$ . The parasitic scattering time is  $\tau_{31} = 100 \text{ ps}$  at design bias. These rates are calculated using bulk GaAs phonon modes. See Chap. 3 for an analysis of the LO-phonon scattering rates in this structure that takes account of the heterostructure, including interface and confined modes.

To assure full carrier ionization and to eliminate radiative 2p-1s impurity transitions, the Si delta doping ( $N_D \approx 6 \times 10^{10}/\text{cm}^2$  design,  $\approx 8.4 \times 10^{10}/\text{cm}^2$  measured) was centered in the 16 ML collector barrier. Thirty modules were cascaded together

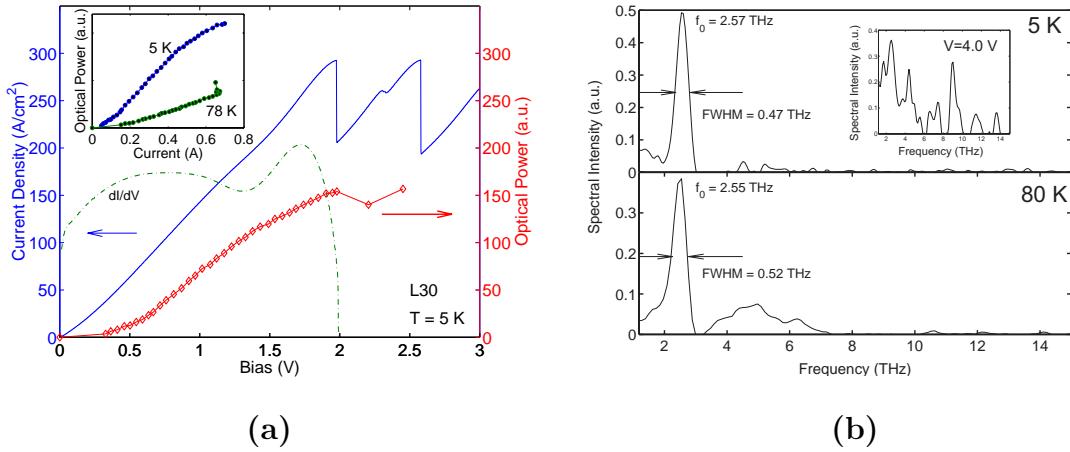


Figure 6-5: (a) Current density, conductance, and optical power versus voltage at 5 K in pulsed mode for L30. Inset gives  $L$ - $I$ . (b) Spectra at 5 K and 80 K for a bias of 1.6 V measured using the bolometer with  $8 \text{ cm}^{-1}$  resolution. The inset shows a spectrum taken at a 4.0 V bias. The device is  $400 \times 600 \mu\text{m}^2$  with a 15- $\mu\text{m}$ -period grating.

and grown by MBE on a  $n^+$  GaAs substrate between  $n^{++}$  contact layers. In order to couple out a large portion of the intersubband radiation power from a lossy cavity ( $\alpha_w \approx 1000 \text{ cm}^{-1}$ ,  $\Gamma \approx 0.83$ ), a diffractive metallic grating was used with a 15  $\mu\text{m}$  period and a 50% filling factor. This grating, which also served as an ohmic contact, was patterned on top of a  $400 \times 600 \mu\text{m}^2$  device mesa using an annealed Ni/Ge/Au alloy. Smaller unpatterned mesas with a size of  $115 \times 115 \mu\text{m}^2$  were used to measure the dc current-voltage ( $I$ - $V$ ) relation. Fig. 6-5(a) shows the dc  $J$ - $V$  curve of the MQW structure measured at 4.2 K, along with its differential conductance. The conductance peaks at 1.7 V, indicating the point at which  $n = 1'$  is aligned with  $n = 3$ . This is close to the design bias of 1.4 V ( $30 \times 47 \text{ mV}$ ), with the extra voltage likely due to the contact resistance. The peak current density is  $J_0 = 290 \text{ A/cm}^2$ . At biases greater than 2 V,  $n = 1'$  becomes misaligned with  $n = 3$ , resulting in a negative differential resistance (NDR). The sawtooth feature in the  $I$ - $V$  is due to the development of high-field domains in which groups of modules jump to higher bias points. The broader conductance peak centered at 0.7–0.8 V is likely due to the  $1' \rightarrow 2$  tunneling process.

Emission spectra, shown in Fig. 6-5(b), reveal a clear peak due to  $3 \rightarrow 2$  intersubband emission. The measured peak frequency of 2.57 THz (10.6 meV) is close to the calculated value of 10.3 meV. Since the  $3 \rightarrow 2$  transition is designed to be close to the anticrossing gap,  $E_3$  and  $E_2$  track each other as the bias voltage is changed (Fig. 6-4(b)), and we should expect to see little Stark shift of the emission frequency. This is indeed the case. In the bias range of 1.0–1.9 V, in which level 3 has an appreciable population, the emission spectra showed a narrow peak whose frequency barely changed with the bias voltage. The full-width half-maximum (FWHM) of the emission peak was measured to be as narrow as 0.47 THz (1.9 meV). Spectra were also taken with the cold stage cooled with liquid nitrogen to 80 K (Fig. 6-5(b)), and the linewidth remained narrow. The secondary broad feature is blackbody radiation due to device heating. At this elevated temperature, the  $I$ - $V$  characteristics remained principally unchanged, but the emitted optical power due to intersubband transitions was approximately a factor of 2.5 smaller. This difference is due to a reduced scattering time  $\tau_{32}$  at the elevated temperature.

The extremely long lifetime  $\tau_{21}$  (calculated to be 30 ps), is much slower than  $\tau_{32}$ , which is likely  $< 10$  ps. This makes obtaining a population inversion in this structure impossible. In fact, since  $2 \rightarrow 1$  scattering is the transport bottleneck, levels  $n = 2$  and  $n = 1'$  likely share most of the population. Since for most of the bias range the injection through the 20 ML barrier is not the bottleneck, the  $I$ - $V$  is nearly linear up to the NDR point, and the only signs of resonant tunneling injection are the slight indications of the conductance curve. For such a device to obtain gain, the 16 ML collector barrier must be drastically thinned to increase  $2 \rightarrow 1$  scattering. However, due to the vertical nature of the radiative transition, this comes at the cost of decreasing the parasitic lifetime  $\tau_{31}$ .

This measurement, performed in early 1999 and reported in Ref. [63], was the first observation of narrow-linewidth intersubband emission from a structure not based on an intrawell transition. For an intrawell transition, the transition energy is primarily determined by the well width, whereas for an anticrossed interwell transition, the energy primarily depends on the barrier thickness. Monolayer fluctuations in layer

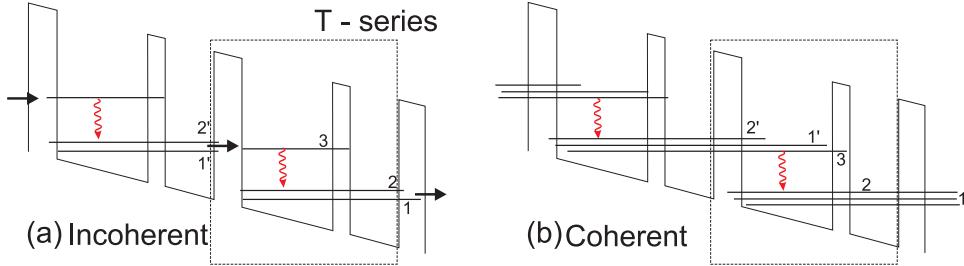


Figure 6-6: Schematic diagram for the T-series of devices for the case of (a) incoherent sequential or (b) coherent resonant tunneling injection.

thickness therefore have a proportionally greater impact on transition energy. The fact that such a narrow linewidth was observed, considering the extra interface roughness scattering and module non-uniformity, indicated the high growth quality of the sample. Also, the temperature insensitivity of the linewidth confirmed the viability of terahertz lasers at elevated temperatures. This structure also demonstrated the advantages of a vertical transition over the diagonal transition featured in the device M100. In that device, the level of intersubband emission was much smaller and the linewidth was broader, presumably due to interface roughness scattering and module non-uniformity. This fact illuminated the necessity of having a sufficiently large oscillator strength to obtain high quality intersubband emission (and lasing).

## 6.4 T series

The T-series of devices were based around a very simple two well module, as is shown in the schematic in Fig. 6-6 or the conduction band diagram for T65 (Fig. 6-7). The lifetime of the upper state  $n = 3$  is limited by e-e scattering and thermally activated LO-phonon scattering, and is expected to be rather long at low temperatures. In these structures, there is no explicit LO-phonon scattering channel for depopulation. Rather, obtaining a population inversion is dependent on the thermalization of the electron distribution in the lower states. As shown in Fig. 6-6, there are two ways to model this system, depending on whether injection from  $n = 2'$  and  $n = 1'$  into  $n = 3$  is modeled by incoherent sequential or coherent resonant tunneling.

In the incoherent case, each module is considered isolated, and carriers are injected into the excited state  $n = 3$ . Population inversion between  $n = 3$  and 2 requires a cold thermal electron distribution in the lower 2-1 doublet, and efficient injection into level  $n = 3$ . One might expect levels 2 and 1 to be in thermal equilibrium due to the relatively fast electron-electron (e-e) scattering that occurs between two closely spaced levels. Since the total population among the three levels remains constant, increased injection will raise the population  $n_3$  while simultaneously depleting  $n_2 + n_1$ .

If injection is coherent, then states  $n = 2', 1'$ , and 3 will form a coherent triplet that is delocalized across the injection barrier. Since the intersubband e-e scattering rate is inversely proportional to subband energy separation, intra-triplet scattering should be faster than inter-triplet scattering, which leads to a thermal distribution inside the triplet. In such a scenario,  $n_3$  will always be greater than  $n'_2$  thereby  $n_2$  by translational symmetry. Thus a population inversion is obtained, the degree of which depends strongly on the electron temperature.

As well as serving as a potential QCL gain medium, this two-well design served as a model system for investigating terahertz intersubband transport, subband separations, and optical properties. Because only three subbands per module participate in transport, analysis of the system becomes significantly more tractable than for devices with more subbands. Although lasing in a T-series device was not obtained, information gleaned from these devices, specifically regarding coherent and incoherent tunneling transport, was of direct assistance in designing operational laser devices.

#### 6.4.1 T65

The T65 structure, shown in Fig. 6-7, is based around a 25.4-nm (90 ML) wide well where the intrawell radiative transition takes place primarily between levels  $n = 3$  and 2. A narrower 14.4 nm (51 ML) collector/injector well and its associated subband  $n = 1$  are used to collect electrons and inject them into the next module. The narrower well is doped with Si to provide an electron density of  $n = 2.0 \times 10^{10} \text{ cm}^{-2}$  per module. As shown in the anticrossing plot in Fig. 6-7(b), the lowest two states anticross with  $\Delta_{21} = \min |E_2 - E_1| \approx 2.5 \text{ meV}$  at  $\sim 19 \text{ mV}$  bias per module. The

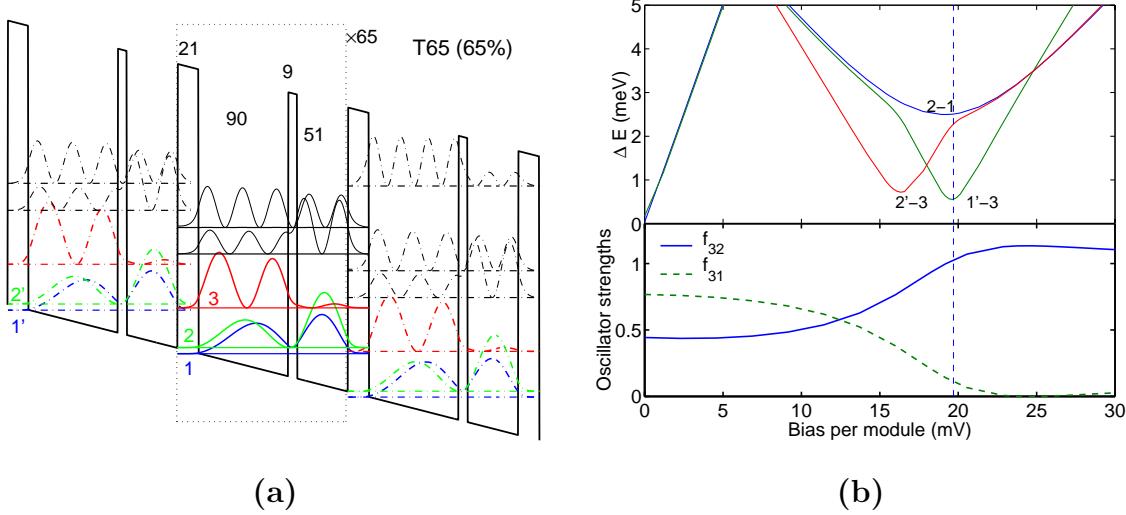


Figure 6-7: (a) Conduction band diagram and anticrossing diagram of T65 grown in GaAs/Al<sub>0.15</sub>Ga<sub>0.85</sub>As calculated with 65% conduction band offset. Layer thickness are given in monolayers. The 51 ML well was doped at  $1.4 \times 10^{16} \text{ cm}^{-3}$  to yield a sheet density of  $2.0 \times 10^{10} \text{ cm}^{-2}$ , which is approximately corroborated via *C-V* measurements. (b) Anticrossing plot and oscillator strengths for T65.

injection barrier is rather thick (5.9 nm or 21 ML), and as a result the interaction is weaker and the anticrossing gaps are smaller than  $\Delta_{21}$ , with  $\Delta_{2'3} = 0.75$  meV and  $\Delta_{1'3} = 0.5$  meV. Analysis of the oscillator strengths plotted in Fig. 6-7(b) reveals that at the designed bias,  $f_{32} \gg f_{31}$  even though  $n=2$  and  $n=1$  are at anticrossing and their two wavefunctions have a strong spatial overlap. This result is due to the sensitive dependence of the dipole moments on the wavefunction symmetry.

### Plasmon waveguide

Sixty-five cascaded two-well modules were grown in the GaAs/Al<sub>0.15</sub>Ga<sub>0.85</sub>As material system on an  $n^+$  GaAs substrate. A 500 Å  $n^{++}$  cap layer ( $n = 5 \times 10^{18} \text{ cm}^{-3}$ ) was grown above the active region, along with a LTG-GaAs non-alloyed ohmic contact. Below the active region, an identical  $n^{++}$  contact layer was grown, with a 2000 Å AlAs etch stop layer underneath, and a 9000 Å-thick  $n^{++}$  GaAs layer was grown beneath that to provide plasmon confinement. Various ridge and grating structures with a height of approximately 3 μm were formed by wet etching. The resulting plasmon waveguide was calculated to have  $\alpha_w = 185 \text{ cm}^{-1}$  with  $\Gamma = 0.86$ .

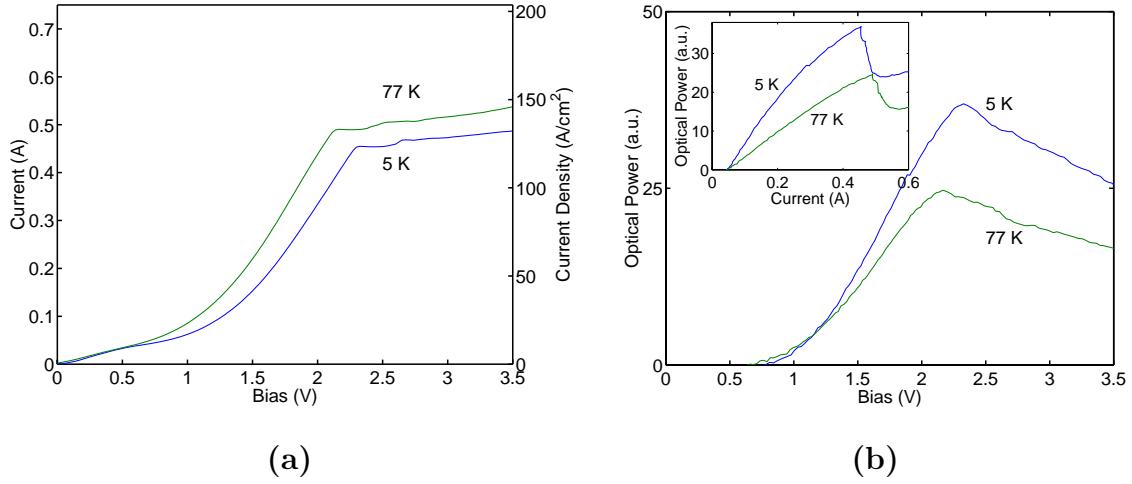


Figure 6-8: (a)  $I$ - $V$  and (b)  $L$ - $V$  and  $L$ - $I$  (inset) for T65 at 5 K and 77 K.  $715 \times 515 \mu\text{m}^2$  grating device with  $15 \mu\text{m}$  period was tested using photodetector at 50% duty cycle.

Typical  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  curves taken at 5 K and 77 K are displayed in Fig. 6-8. The sublinear behavior of the  $L$ - $I$  at 5 K is attributed to the reduction of the lifetime  $\tau_3$  of level  $n=3$  with increased injection due to either electron-electron scattering or heating of the electron gas. At  $T = 80$  K,  $\tau_3$  is further reduced due to thermal activation of LO phonon assisted scattering.

Due to the large waveguide loss, edge emission spectral measurements were impossible, and emission measurements were performed using surface emission grating structures with periods of 15, 20, and 25  $\mu\text{m}$ . Representative spectra taken at 5 K using 15- and 20- $\mu\text{m}$  period gratings are shown in Fig. 6-9, and reveal a narrow emission peak at 21.1 meV (5.04 THz) due to the intrawell transition. There is some discrepancy between the observed emission and the calculated  $E_{32} = 17.5$  meV (at design bias), which may be due to MBE undergrowth of the design. No x-ray data was available for this growth, so this conjecture cannot be verified. The linewidth is narrowest at low biases before the 2-1 anticrossing where the intrawell transition  $3 \rightarrow 1$  is relatively isolated. At these low biases, the  $3 \rightarrow 2$  transition is diagonal, there is emission power at lower frequencies. This peak shifts to higher frequencies with bias and eventually joins with the main peak. This Stark shift is obscured by the effect of the grating on the emission spectrum. The narrowest measured linewidth

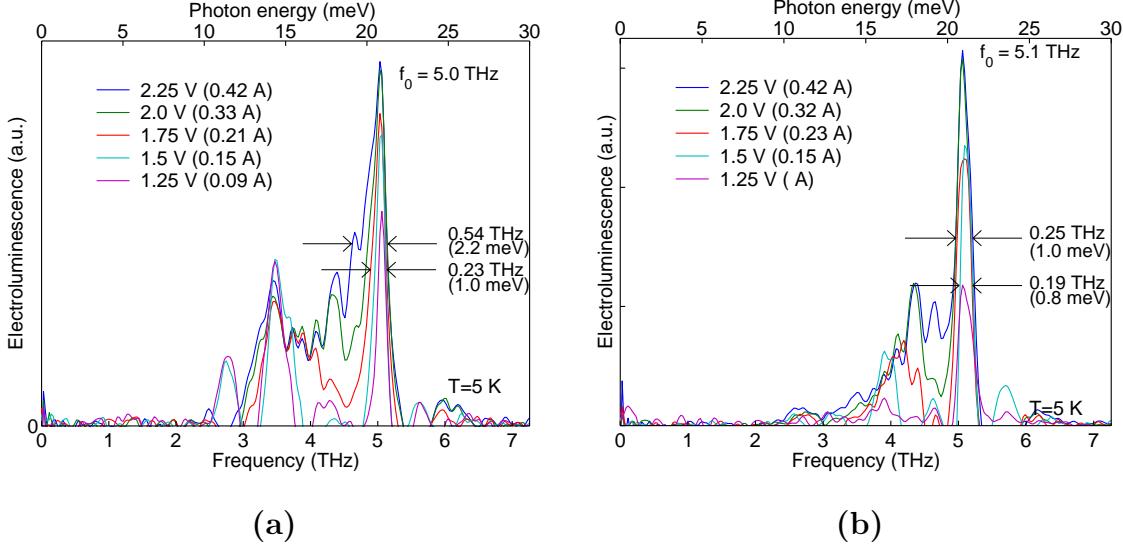


Figure 6-9: T65 spectra at 5 K from (a)  $715 \times 505 \mu\text{m}^2$  grating device with  $20 \mu\text{m}$  period and (b)  $715 \times 515 \mu\text{m}^2$  grating device with  $15 \mu\text{m}$  period collected using photodetector with  $4 \text{ cm}^{-1}$  ( $0.12 \text{ THz}$ ) resolution.

was 0.7 meV measured from a 25- $\mu\text{m}$ -period grating device, and linewidths of 0.8 and 1.0 meV were measured from the 15- $\mu\text{m}$  and 20- $\mu\text{m}$  gratings. Although the measured peaks might be slightly narrowed due to the effect of the grating, such narrow linewidths indicate the high quality of the growth. These are characteristic of intrawell designs [61], which are less sensitive to interface roughness scattering and well width fluctuations.

## Magnetotunneling

Transport measurements performed under high magnetic fields have proven to be useful for determining subband energy separations in quantum well structures, including THz emitters [165, 166, 167, 168]. Application of a magnetic field  $B$  along the growth direction will split each subband into discrete Landau levels separated by  $\hbar\omega_c$ , where  $\omega_c = eB/m^*$  is the cyclotron frequency. Intersubband scattering, and thereby transport, will be suppressed as  $B$  increases, except at resonant field values where the Landau levels are aligned. Hence, the current will exhibit peaks at values

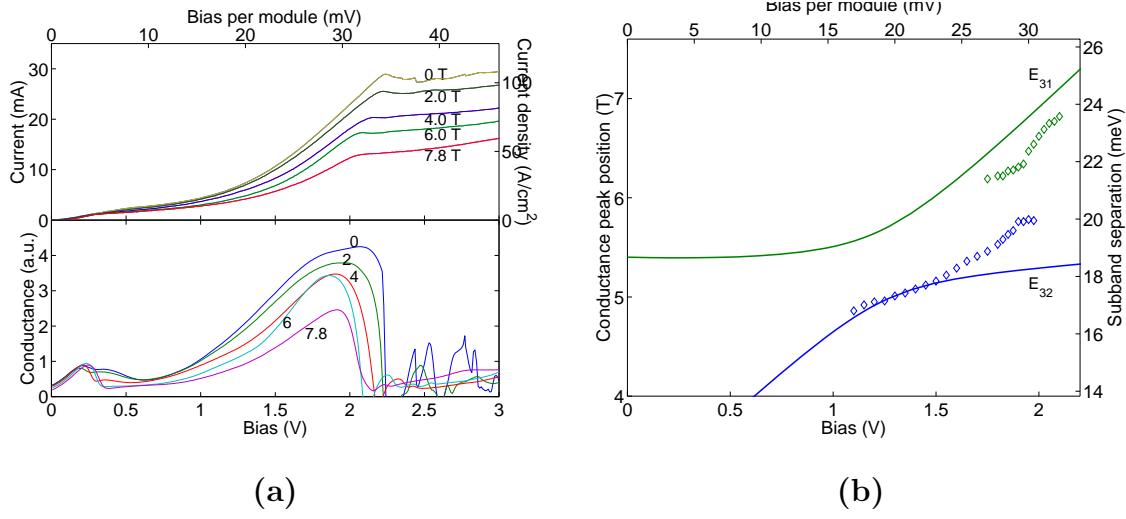


Figure 6-10: (a) Current and conductance versus voltage for several values of applied magnetic field at  $T = 4.2$  K. (b) Position of  $\Delta\ell = 2$  peaks in  $G$ - $B$  plots and associated energy difference near the anticrossing of subbands  $n=2$  and 1. The solid lines are the corresponding calculated energy differences  $E_{32}$  and  $E_{31}$ .

of  $B$  that obey

$$\Delta E = \Delta\ell\hbar\omega_c, \quad (6.1)$$

where  $\Delta\ell$  is the change in Landau index. Identification of these peaks in the current-magnetic field ( $I$ - $B$ ) relation will reveal the intersubband energy separation  $\Delta E$ . This method has been successfully used to measure the energy separation between the two radiative states [165, 166]. However, the energy separation between a radiative level and the injection/collection level is usually only a few meV, which requires measurements of the conductance  $G=dI/dV$  or higher-order derivatives to enhance the energy resolution and sensitivity. While  $I$ - $B$  magneto-oscillations reveal “bottlenecks” in the sum over all current paths,  $G$ - $B$  magneto-oscillations act as a local probe in energy space.

For magneto-transport measurements, devices were processed into  $150 \times 150 \mu\text{m}^2$  mesas, and mounted in liquid helium inside a magnet dewar. Magnetic fields up to 8 T were applied along the growth direction (parallel to current flow). Current-voltage and conductance-voltage characteristics are plotted in Fig. 6-10 for several values of applied magnetic field. At  $B = 0$ , the conductance curve flattens at 1.8–2.1

$V$ , as  $n=3$  becomes aligned first with  $n=2'$  and then with the injector state  $n=1'$ . At 2.25 V,  $n=1'$  becomes severely misaligned with  $n=3$ , and a negative differential resistance (NDR) occurs. While this is somewhat greater than the calculated design bias of 1.3 V, the difference is partially due to the fact that the intrawell transition energy is 21 meV (obtained from the emission measurements), rather than calculated 17.5–18.5 meV (depending on the bias). Parasitic resistances may also contribute to the voltage offset. As  $B$  is increased from 0 to 7.8 T, the current decreases due to the reduction in phase space for intersubband scattering. Additionally, the NDR shifts to lower voltages with increasing  $B$ . The formation of the Landau levels sharpens the density of states, and thus less misalignment of the subbands  $n=1'$  and 3 is tolerated before the device enters an NDR region.

$I$ - $B$  and  $G$ - $B$  characteristics were measured at many voltages, several of which are shown in Fig. 6-11. For example, at  $V = 1.90$  V  $I$ - $B$  maxima corresponding to  $\Delta\ell = 2, 3, 4$  are seen at 5.47 T, 3.64 T, and 2.73 T, which correspond to an energy separation of 18.9 meV. The alignment of the Landau levels is shown schematically in the fan diagram in Fig. 6-11(a). While the peaks in the  $G$ - $B$  curves generally mirror those in the  $I$ - $B$  curves, at  $V > 1.8$  V the  $G$ - $B$  data reveals the more complicated structure of an anticrossing between two  $\Delta\ell = 2$  peaks (indicated by arrows in Fig. 6-11(c)). As the bias is increased, the magnitude of the lower energy peak decreases while that of the higher energy peak increases. The limited resolution at  $\sim 4$  T largely obscures much of the anticrossing effect for the  $\Delta\ell = 3$   $G$ - $B$  peak. The dependence of these  $\Delta\ell = 2$  conductance peaks on bias is shown in Fig. 6-10(b), and it displays a typical anticrossing behavior between a Stark shifted peak and another fixed at  $\sim 6.0$  T ( $\sim 21$  meV). This 21 meV peak corresponds to the intrawell transition, which agrees with the emission data. At 1.9 V, the maxima are minimally separated by 0.5 T, which corresponds to an anticrossing gap of 1.7 meV. This likely underestimates the actual subband separation slightly as the apparent maxima are shifted together by the summation of two peaks of substantial linewidth. Since the  $1'-3$  anticrossing is calculated to be  $\Delta_{1'3} = 0.5$  meV, we conclude that the observed anticrossing of at least 1.7 meV is between subbands  $n=2$  and 1 (calculated  $\Delta_{21} = 2.5$  meV). We can

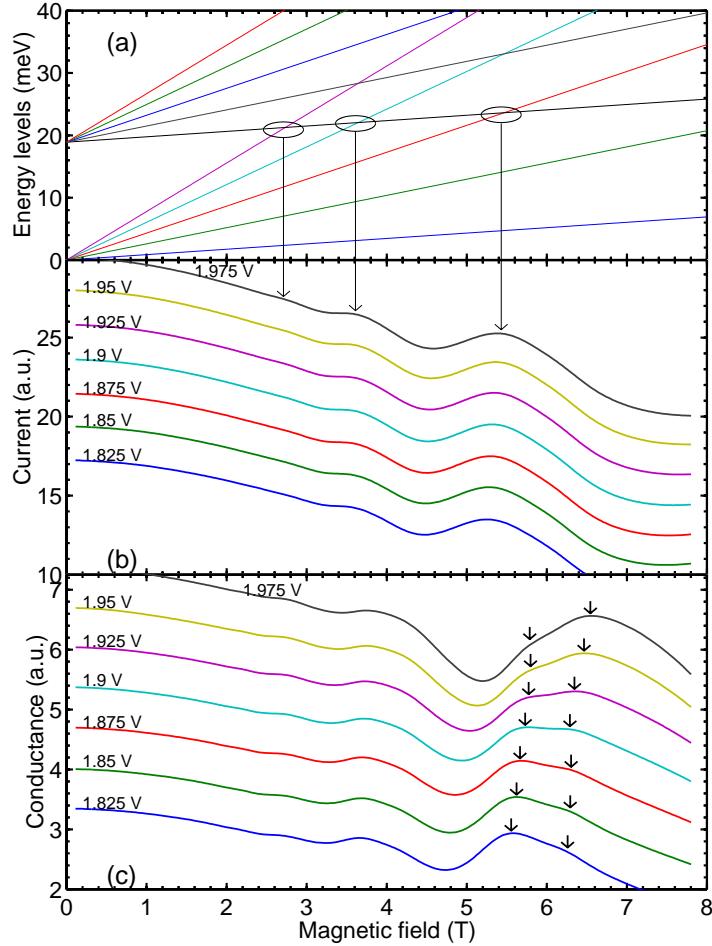


Figure 6-11: Current (b) and conductance (c) versus magnetic field ( $\mathbf{B} \parallel \mathbf{J}$ ) for several applied biases. A fan diagram that shows the crossing of Landau levels for an 18.9 meV energy separation is shown in (a). The arrows in (c) indicate the resonance peaks that correspond to the 2-1 anticrossing.

therefore conclude that the Stark-shifted peak represents the  $3 \rightarrow 2$  transition (below the anticrossing bias). The energy separations  $E_3 - E_2$ , and  $E_3 - E_1$  are plotted in Fig. 6-10(b) for comparison. No evidence of this anticrossing is seen in the optical emission spectra, although this is not surprising, since at the 2-1 anticrossing we calculate  $f_{32} \gg f_{31}$  (Fig. 6-7(b)). Also, some of the detailed structure of the peaks is obscured because of the gratings. The anticrossing behavior is seen more clearly in edge emission measurements of similar devices in Sec. 6.4.3. Furthermore, no evidence is seen of a  $1' - 3$  anticrossing, since the injection barrier is thick, and  $1' \rightarrow 3$  transport is likely due to incoherent sequential tunneling between spatially localized states. This is corroborated by the fact that when temperature is raised from 5 K to 80 K the emitted power drops noticeably (Fig. 6-8), while the current changes little. This indicates that the  $1' \rightarrow 3$  tunneling time, rather than the lifetime  $\tau_3$ , is the transport bottleneck, which is a signature of incoherent sequential tunneling (see Sec. 2.5.2 and Ref. [104]).

The presence of the anticrossing gap between levels  $n=2$  and 1 indicates that the tunneling in this structure should be modeled as a coherent resonant process rather than an incoherent sequential one. Rather than speaking of a “tunneling time” between wells, it is more appropriate to calculate the coherent anticrossed wavefunctions for  $n=2$  and 1, and describe transport in terms of intersubband scattering processes. As discussed in Sec. 2.5.2, the transition between resonant coherent tunneling and incoherent sequential tunneling between subband depends on both the dephasing time and the various state lifetimes. Eq. (2.70) describes the reduction of the damped oscillation frequency (effective anticrossing gap) due to dephasing. Clearly dephasing is not fast enough to over-damp this tunneling interaction, and the broadening of each state is less than  $\Delta_{21}$ , although it may be part of the reason that the measured anticrossing gap is smaller than the calculated one. Thus it is legitimate to describe the states  $n = 2$  and 1 as delocalized and separated in energy by an anticrossing gap. This knowledge is important for modeling transport and population distribution in these and similar structures. However, care must be taken when considering the coherence of various tunneling channels with similar calculated anticrossing gaps. Since

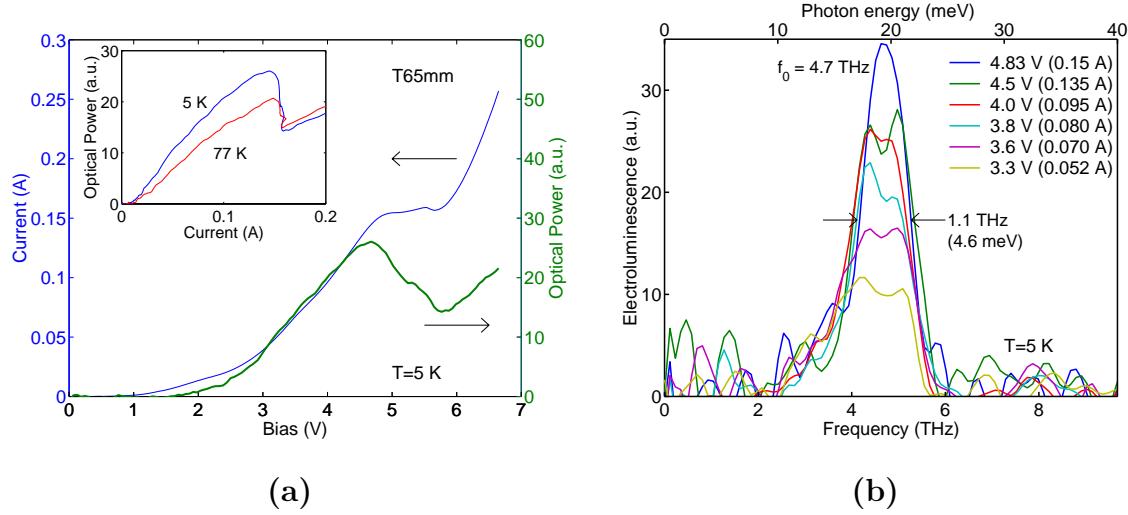


Figure 6-12: (a)  $I$ - $V$  and  $L$ - $V$  plots at 5 K and  $L$ - $I$  plots at 5 K and 77 K for T65mm edge emitting device 1.5 mm wide and approximately 70  $\mu\text{m}$  deep. (b) Edge emitting spectra from same device taken at 5 K at 50% duty cycle using the photodetector with a spectral resolution of  $16 \text{ cm}^{-1}$ .

these two states are low in the wells, they likely have relatively long lifetimes, thus easing the condition for coherence (see Eqs. (2.64),(2.65)).

### Metal-metal waveguide (T65mm)

Metal-metal waveguide fabrication took place via use of Au-Au thermocompression bonding. Substrate removal was made difficult due to the use of a 200-nm AlAs etch stop layer instead one composed of  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$ . As described in Sec. 5.3.2, such etch stop layers did not survive substrate removal in  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2$  (1:19). Instead, citric acid: $\text{H}_2\text{O}_2$  (1:4) was used, which allowed sufficient yield to obtain a few devices. Ridges were defined by wet etching in  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (10:6:480), and then silicon nitride was deposited using low temperature PECVD in the Plasmaquest. After windows were dry etched on top of the mesas, Ti/Au contacts were evaporated. The metal-metal waveguide was calculated to have  $\alpha_w = 22.7 \text{ cm}^{-1}$  and  $\Gamma = 0.93$ , using  $\tau = 0.5 \text{ ps}$  for the active region,  $\tau = 0.1 \text{ ps}$  for the  $n^{++}$  GaAs,  $\tau = 0.05 \text{ ps}$  for Au.

$I$ - $V$  and  $L$ - $V$  plots for T65mm are shown in Fig. 6-12(a). The  $I$ - $V$  looked qualitatively good, except that the NDR was located at approximately 4.8 V, instead of

2.25 V as for T65. The difference is attributed to poor Schottky contacts due to contamination of the GaAs surface. In hindsight, it would have been better to deposit the Ti/Au in the first photostep immediately after substrate removal. The reduced loss of the metal-metal waveguide allowed the use of edge emitting configuration for spectral measurements, which are shown in Fig. 6-12(b) for a ridge device cleaved along its length to be 1.5-mm wide and approximately 70- $\mu$ m deep. Due to the low signal-to-noise ratio, the spectra were only able to be taken with  $16\text{ cm}^{-1}$  (0.48 THz) resolution, which prevents resolution of the true linewidths. Nevertheless, the edge emitting spectra clearly show the narrowing of the emission peak with bias. This occurs as the diagonal transition is Stark shifted with increasing bias until it joins with the vertical intrawell transition at high bias.

#### 6.4.2 T150C

The T150C design was identical to that of T65, except the 5.9 nm (21 ML) injection barrier was changed to 2.8 nm thick (10 ML). The injection anticrossing gaps were calculated to be  $\Delta_{2'3} = 1.8$  meV and  $\Delta_{1'3} = 2.0$  meV, using 65% band offset. One-hundred fifty modules were grown on an  $n^+$  GaAs substrate with a 500-nm-thick AlAs etch stop layer, and the ridge and grating structures were wet-etched. However, the *I-V* characteristic for this device, shown in Fig. 6-13, displayed a high current density of  $J_0 = 1500\text{ A/cm}^2$ , with only a kink in the *I-V* characteristic at  $\sim 5$  V to indicate the design bias. The likely cause for this was the opening of a parasitic current channel from the upper radiative state  $n = 3$  to the excited state in the next module (see Fig. 6-7). The anticrossing gap for the parasitic channel was 0.9 meV. Collected spectra were overwhelmed by blackbody radiation, and no particular evidence of the intersubband transition was observed. This is not surprising, as the  $n = 3$  state was likely depopulated by the parasitic channel.

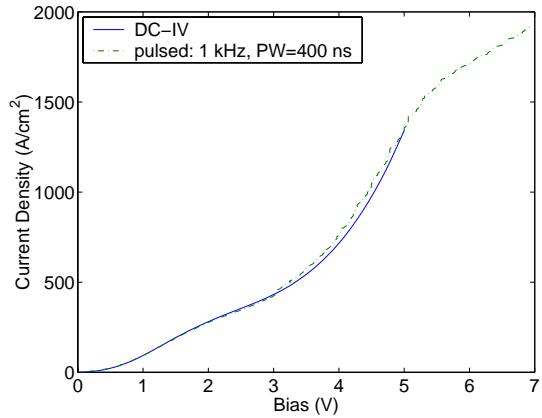


Figure 6-13:  $I$ - $V$  characteristic for T150C  $140 \times 140 \mu\text{m}^2$  diode at 5 K, measured both in dc mode and pulsed mode.

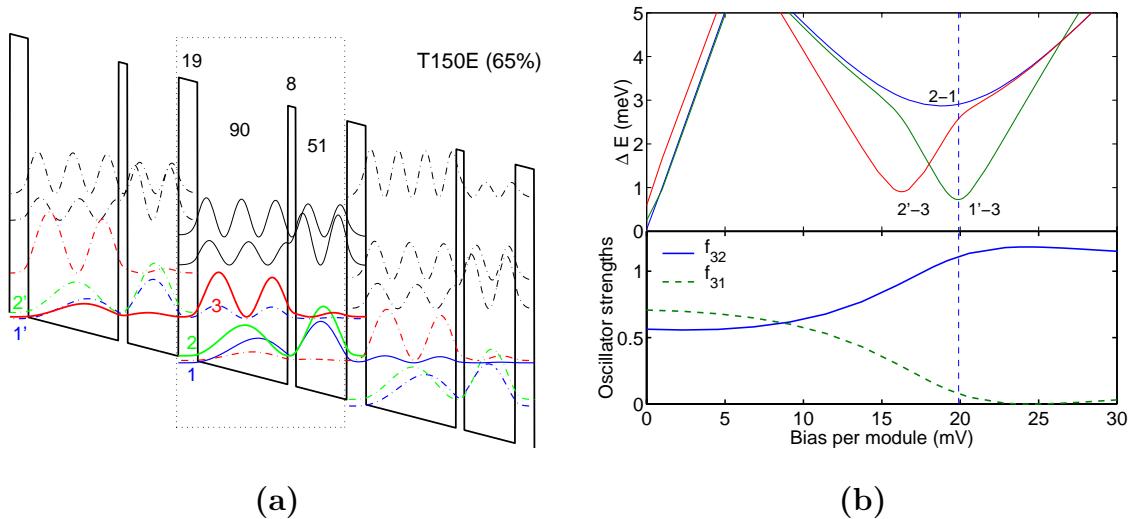


Figure 6-14: (a) Conduction band diagram and anticrossing diagram of T150E/T210E grown in GaAs/Al<sub>0.15</sub>Ga<sub>0.85</sub>As calculated with 65% conduction band offset. Layer thickness are given in monolayers. The 51 ML well was doped at  $1.4 \times 10^{16} \text{ cm}^{-3}$  to yield a sheet density of  $2.0 \times 10^{10} \text{ cm}^{-2}$ . (b) Anticrossing plot and oscillator strengths for T150E/T210E.

Table 6.1: Calculated waveguide parameters for various growths in the T150E/T210E family.

Device	Etch stop	Waveguide	$t_{active}$ ( $\mu\text{m}$ )	$\alpha_w$ ( $\text{cm}^{-1}$ )	$\Gamma$
T150	500 nm AlAs	plasmon	7.12	> 110	0.84
T150E-2mm	500 nm $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$	metal-metal	7.12	11.8	0.97
T150E-3	none	plasmon	7.12	95	0.94
T210E	none	plasmon	9.97	82	0.95

### 6.4.3 T150E, T210E

The T150E series of devices, shown in Fig. 6-14, was designed to have an identical active region as T65, with two modifications. The injector barrier was narrowed from 5.9 nm (21 ML) to 5.4 nm (19 ML) which raised the injection anticrossings to  $\Delta_{2'3} = 0.9$  meV and  $\Delta_{1'3} = 0.73$  meV to allow for higher current. Also, the collector barrier was changed from 2.5 nm (9 ML) to 2.3 nm (8 ML), which raised the anticrossing gap to  $\Delta_{21} = 2.9$  meV. All values were calculated using a conduction band offset of 65%.

There were four separate wafers grown using this active region design. T150E was grown using an AlAs etch stop layer, which proved to be unsuitable for substrate removal. As a result T150E-2 was grown with an 500-nm  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  etch stop layer, but carrier freeze out impeded testing without fabrication into a metal-metal waveguide. Also T150E-3 and T210E were grown without etch stop layers, but instead with 2- $\mu\text{m}$   $n^{++}$  plasmon confinement layers beneath their active regions. The characteristics for their respective waveguides are summarized in Table. 6.1.

#### Plasmon waveguide (T150E-3, T210E)

The  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  characteristics from a typical T150E-3 device (150 cascaded modules) are shown in Fig. 6-15. These characteristics look qualitatively similar to those from T65. At 5 K, the NDR is located at 4.4 V, which corresponds to 29 mV/module. The peak current density was approximately  $J_0 \approx 160 \text{ A/cm}^2$ . Due to the thicker active region, the waveguide loss was smaller than for T65, and edge

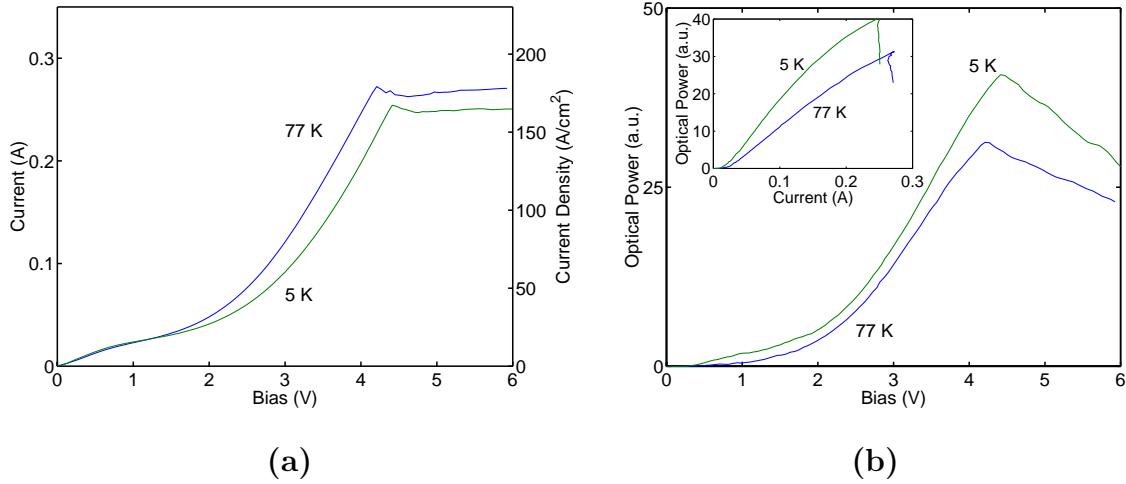


Figure 6-15: (a)  $I$ - $V$  and (b)  $L$ - $V$  and  $L$ - $I$  (inset) for T150E-3 at 5 K and 77 K. 80- $\mu\text{m}$ -wide and 1.9-mm-long ridge structure was tested using photodetector at 50% duty cycle.

emitting spectral measurements were possible. Spectra from edge emitting T150E-3 and T210E devices are shown in Fig. 6-16. These devices are identical except for the number of modules. The edge emission data shows the evolution of the spectra more clearly than the grating data from T65 due to the lack of any grating resonances or artifacts. At low biases, the spectrum is composed of a broad peak at lower energy ( $\sim 16$  meV) that corresponds to the diagonal radiative transition  $3 \rightarrow 2$ , and a narrow peak at a higher energy ( $\sim 22$  meV), that corresponds to the  $3 \rightarrow 1$  transition. At low biases, the two transitions have comparable oscillator strengths (Fig. 6-14). As bias increases, the diagonal transition undergoes a Stark shift and increases in magnitude. At the same time, the vertical transition remains fixed in energy, and decreases in strength relative to the diagonal peak, as indicated by their relative oscillator strengths. Finally at design bias, levels 2 and 1 are at anticrossing, and the electroluminescence is dominated by the  $3 \rightarrow 2$  transition, measured to be  $\sim 20$  meV. This linewidth is relatively narrow, since the transition is now largely vertical, and is less affected by interface roughness scattering. The minimum linewidths measured at 5 K are 2.0 meV for T150E-3, and 3.4 meV for T210E. The difference is likely due to the greater impact of module non-uniformity for the devices with 210 modules. The

spectra are slightly broader at 78 K: 2.6 meV for T150E-3 and 4.7 meV for T210E. These spectra provide qualitative confirmation of the predicted energy separations and oscillator strengths, and are consistent with the observation of the 2-1 anticrossing gap for T65.

### Metal-metal waveguide (T150E-2mm)

The T150E-2mm metal-metal waveguide device was fabricated using Au-Au thermocompression bonding. Ti/Au Schottky contacts were deposited immediately after substrate removal and used as etch masks for mesa etching in the Plasmaquest using  $\text{BCl}_3:\text{N}_2$  (15:5 sccm). The ridge sidewalls and bottom were then protected with silicon oxide. Since the wafer-bonded devices did not survive wire bonding, Au wires were attached to the ridge top contacts with silver epoxy. Unlike T65mm, the *I-V* characteristic (Fig. 6-17(a)) showed no sign of any parasitic beyond a  $\sim 0.4$  V drop that can be attributed to the Schottky diode. However, the NDR appeared at 4.0 V, slightly lower than that of T150E-3. For various other devices tested, the NDR varied between 4.0 and 5.0 V. Furthermore,  $J_0 \approx 110 \text{ A/cm}^2$  for the metal-metal devices at 5 K, compared to  $J_0 \approx 130 \text{ A/cm}^2$  for T150E-2 devices tested without the metal-metal processing. This suggests that the T150E-2mm devices are reaching NDR at an earlier bias point, although the reason for this is not clear.

Spectra taken from a 100- $\mu\text{m}$ -wide and 1.875-mm-long T150E-2mm device are shown in Fig. 6-17(b). While the envelope shapes of the various spectra are similar to those of T150E-3 and T210E, the spectra display deep fringes that subdivide broad peaks into several regularly spaced narrow peaks. The peaks in Fig. 6-17(b) are separated by approximately  $13 \text{ cm}^{-1}$ , and spectra measured from an 80- $\mu\text{m}$ -wide ridge (not shown) display a mode spacing of approximately  $16 \text{ cm}^{-1}$ . These peaks appear to be related to a lateral Fabry-Pérot effect that occurs between the vertical dry-etched sidewalls of the ridge. Discussion of these peaks is deferred until the end of this section.

The peak spontaneous emission power levels for various T150E devices are listed in Table 6.2. The power from the metal-metal waveguide devices was at least a factor

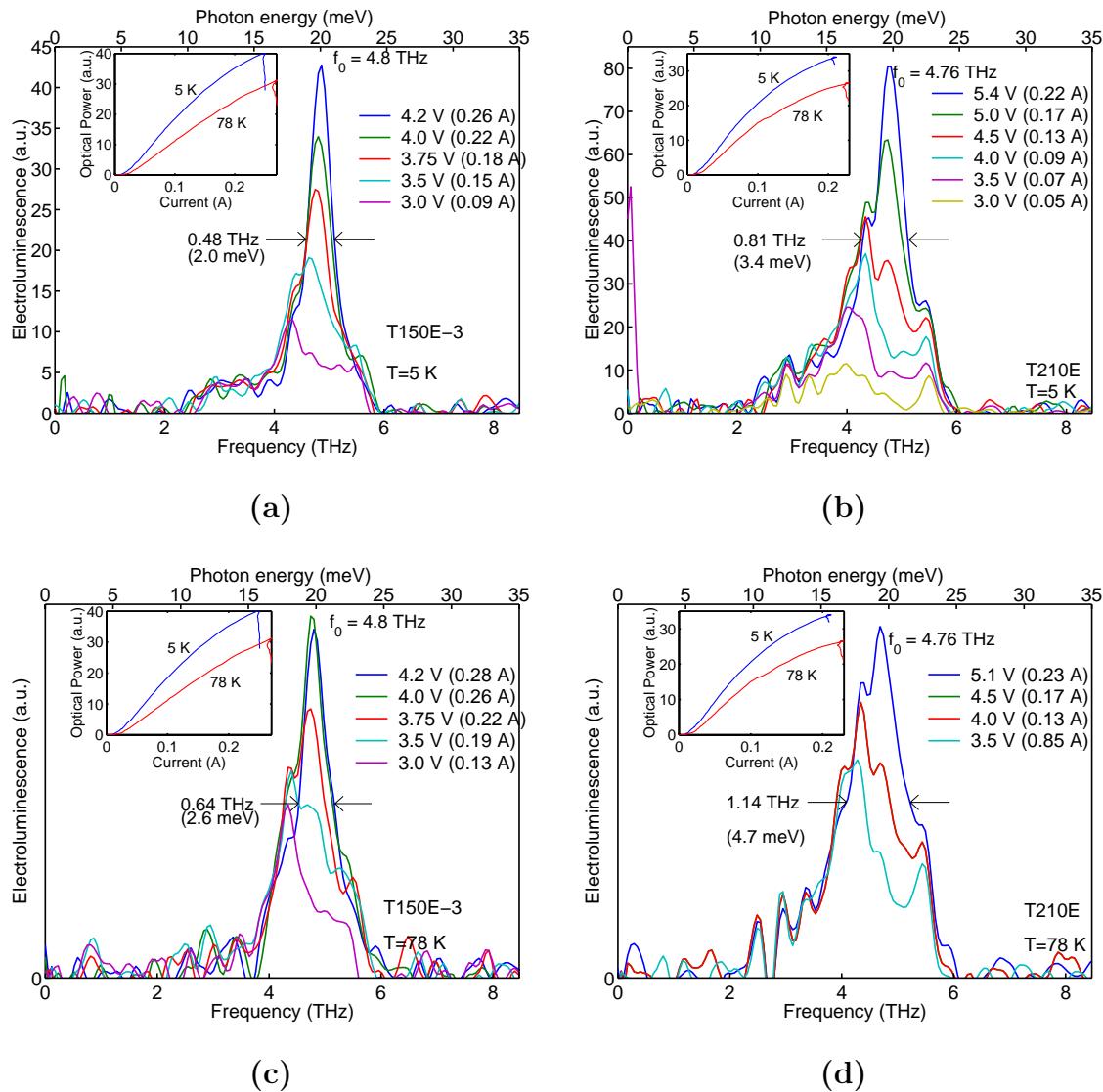


Figure 6-16: (a) (c) Edge emitting spectra and  $L$ - $I$  taken from 80- $\mu\text{m}$ -wide and 1.9-mm-long T150E-3 device at 5 K and 78 K. (b), (d) Edge emitting spectra and  $L$ - $I$  taken from 80- $\mu\text{m}$ -wide and 1.9-mm-long T210E device at 5 K and 78 K. Spectra were taken at 50% duty cycle using the photodetector with a spectral resolution of  $8 \text{ cm}^{-1}$ .

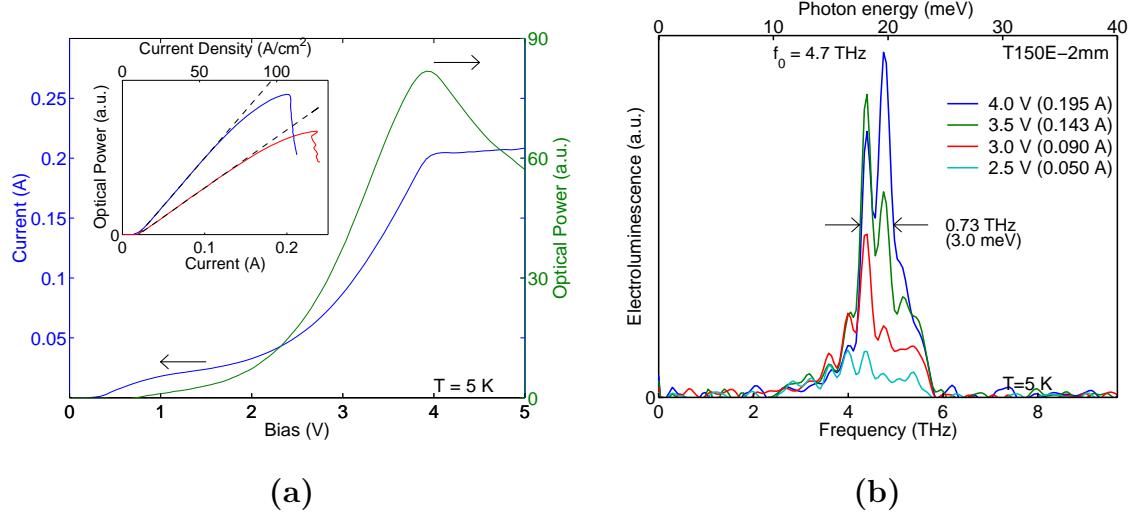


Figure 6-17: (a)  $I$ - $V$  and  $L$ - $V$  plots at 5 K and  $L$ - $I$  plots at 5 K and 77 K for T150E-2mm edge emitting device that is 100- $\mu\text{m}$  wide and 1.875-mm long. (b) Edge emitting spectra from same device taken at 5 K at 50% duty cycle using the photodetector with a spectral resolution of  $8\text{ cm}^{-1}$ .

of two larger than from plasmon waveguide devices of comparable size. Since the out-coupled power is proportional to  $\alpha_m/(\alpha_w + \alpha_m)$ , this is a good confirmation of the lower waveguide loss of the metal waveguide.

#### 6.4.4 T150F

The design of T150F was the same as T150E, except for the use of a thinner 4.5 nm (16 ML) injection barrier. This raised the calculated anticrossings to  $\Delta_{2'3} = 1.3\text{ meV}$  and  $\Delta_{1'3} = 1.1\text{ meV}$ , calculated using 65% conduction band offset. The rest of the module was unchanged, and thus the band diagram and oscillator strengths are essentially the same as shown for T150E in Fig. 6-14(b). The device was MBE grown with a double etch stop layer of  $n^{++}$  200 nm and 300 nm of  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$ , separated by 100 nm of GaAs, over an  $n^{++}$  GaAs buffer on an  $n^+$  GaAs substrate.

#### Plasmon waveguide

Plasmon waveguide devices were fabricated by wet-etching to form ridges approximately 7- $\mu\text{m}$  high. The waveguide loss was calculated to be at least  $\alpha_w = 80\text{ cm}^{-1}$

Table 6.2: Measured spontaneous output powers for edge emitting devices in the T150E family. Power was collected in face-to-face mode using the photodetector at 50% duty cycle at 10 kHz. The units of peak power correspond to the lock-in amplifier voltage (in bandpass mode). Due to optical alignment, there is some uncertainty in the measured power levels.

Device	Size	Peak Current (A)	Peak Power ( $\mu$ V)
T150E-2	80 $\mu$ m $\times$ 1.8 mm	0.18	22
T150E-3	80 $\mu$ m $\times$ 1.9 mm	0.25	40
T210E	80 $\mu$ m $\times$ 1.9 mm	0.21	35
T210E	66 $\mu$ m $\times$ 1.975 mm	0.19	36
T150E-2mm	80 $\mu$ m $\times$ 2.23 mm	0.2	70
T150E-2mm	100 $\mu$ m $\times$ 1.875 mm	0.2	90
T150E-2mm	60 $\mu$ m $\times$ 2.23 mm	0.145	75
T150E-2mm	100 $\mu$ m $\times$ 2.225 mm	0.23	80
T150E-2mm	100 $\mu$ m $\times$ 1.125 mm	0.118	85
T150E-2mm	100 $\mu$ m $\times$ 4.425 mm	0.49	85
T150E-2mm	170 $\mu$ m $\times$ 0.340 mm	0.063	86
T150E-2mm	170 $\mu$ m $\times$ 1.475 mm	0.28	153
T150E-2mm	170 $\mu$ m $\times$ 0.710 mm	0.13	89
T150E-2mm	170 $\mu$ m $\times$ 0.940 mm	0.18	143
T150E-2mm	170 $\mu$ m $\times$ 1.31 mm	0.26	73

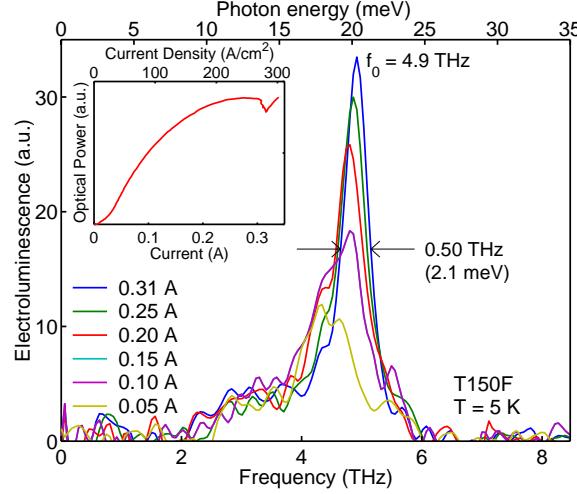


Figure 6-18: Edge emitting spectra and  $L$ - $I$  from a T150F 145- $\mu\text{m}$ -wide and 0.775-mm-long device taken at 5 K at 50% duty cycle using the photodetector with a spectral resolution of  $8 \text{ cm}^{-1}$ .

with  $\Gamma = 0.94$ . The carrier freeze out associated with the  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  layers at 5 K introduced a variable parasitic resistance, which made reliable voltage measurements difficult. The peak current density was  $J_0 \approx 270 \text{ A}/\text{cm}^2$ . Nevertheless, spectral measurements could be performed at 5 K, and are shown in Fig. 6-18. The characteristic Stark shifting and narrowing behavior of the central peak is observed.

### Metal-metal waveguide (T150Fmm)

The metal-metal waveguide T150Fmm device was fabricated via Au-Au thermo-compression bonding, and the devices were dry etched in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm), using the Ti/Au Schottky contacts as self-aligned etch masks. The sidewalls were coated with silicon oxide for protection. The calculated waveguide parameters were  $\alpha_w = 11.9 \text{ cm}^{-1}$  with  $\Gamma = 0.97$ , using  $\tau = 0.5 \text{ ps}$  for the active region,  $\tau = 0.1 \text{ ps}$  for the  $n^{++}$  GaAs,  $\tau = 0.05 \text{ ps}$  for Au. Gold wires were attached to the top of the ridges using conductive epoxy. In some devices, the use of the epoxy appeared to affect the transport, as strange kinks would sometimes appear in the  $I$ - $V$  and  $L$ - $V$  characteristics. This was attributed to a resin in the epoxy that would sometimes seep over the edges of the ridge. Even though the device sidewalls were coated

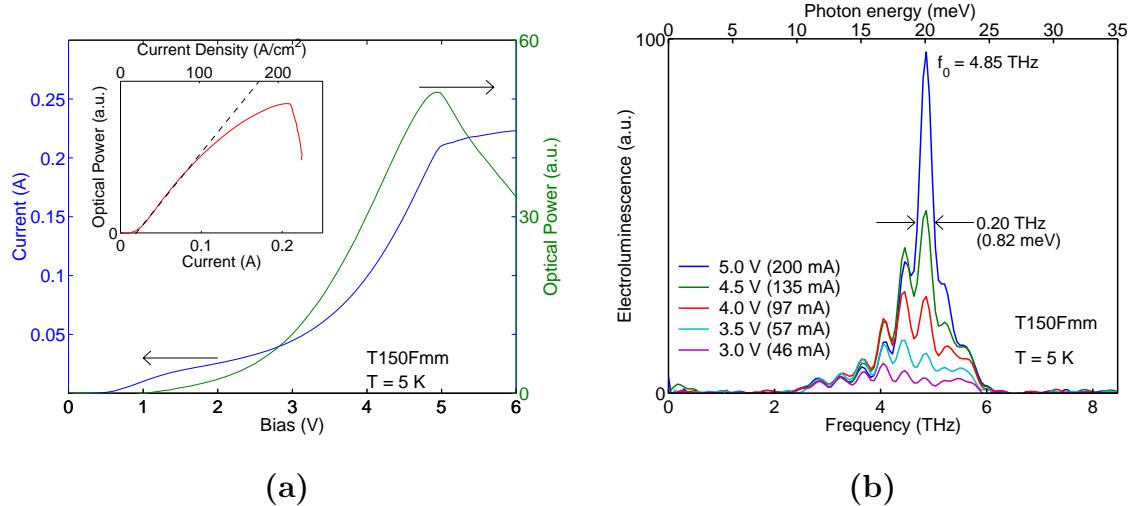


Figure 6-19: (a)  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  plots at 5 K for T150Fmm edge emitting device that was 100- $\mu\text{m}$  wide and 0.975-mm long. (b) Edge emitting spectra from same device taken at 5 K at 50% duty cycle using the photodetector with a spectral resolution of 8  $\text{cm}^{-1}$ .

with  $\text{SiO}_2$  for electrical isolation, it is possible the coverage was not perfect. These problems with wire bonding and epoxy contacts for Au-Au bonded waveguides were one motivation for developing the more robust In-Au bonding technique described in Sec. 5.3

The transport characteristics for the best device (100- $\mu\text{m}$  wide and 0.975-mm long) at 5 K are shown in Fig. 6-19. The NDR is located at 4.9 V with a current density of  $J_0 = 215 \text{ A/cm}^2$ . Just as with T150E-2mm, the peak current density is slightly lower than that obtained without the metal-metal waveguide, which is likely due to the extra heating in the plasmon devices caused by the resistive etch-stop layer. Once again, the spectra display evidence of lateral modes, separated by approximately 13  $\text{cm}^{-1}$ . Spectra taken from a 60- $\mu\text{m}$ -wide ridge display peaks separated by 23  $\text{cm}^{-1}$ . These lateral modes are discussed below in Sec. 6.4.7. If the spectral dips are ignored, the envelope of the spectrum is narrowest at 5.0 V bias, with a FWHM  $\approx 2.5 \text{ meV}$ , which is similar to the 2.1 meV linewidth observed in the plasmon waveguide device (Fig. 6-18). The spontaneous emission power level observed from the metal-metal devices was again larger than that from the plasmon waveguide, often by as much as a factor of four (see Table 6.3).

Table 6.3: Measured output powers for various T150F edge emitting devices collected in face-to-face mode with the photodetector. The units of peak power correspond to the lock-in amplifier voltage (in bandpass mode) for 50% duty cycle at 10 kHz. Due to optical alignment, there is some uncertainty in the measured power levels.

Device	Size	Peak Current (A)	Peak Power ( $\mu$ V)
T150F	145 $\mu$ m $\times$ 0.775 mm	0.3	46
T150Fmm	60 $\mu$ m $\times$ 1.9 mm	0.19	160
T150Fmm	40 $\mu$ m $\times$ 2.55 mm	0.19	80
T150Fmm	100 $\mu$ m $\times$ 0.975 mm	0.21	165
T150Fmm	80 $\mu$ m $\times$ 2.55 mm	0.40	170
T150Fmm	100 $\mu$ m $\times$ 4.675 mm	0.935	160
T150Fmm	100 $\mu$ m $\times$ 0.575 mm	0.10	160

#### 6.4.5 T215G

##### SI-surface-plasmon waveguide

The active region design for T215G was the same as T150E and T150F, except 4.0 nm (14 ML) injection barriers were used. This gave injection anticrossings of  $\Delta_{2'3} = 1.6$  meV and  $\Delta_{1'3} = 1.4$  meV, calculated using 65% conduction band offset. The device was grown with 215 modules, which gave a 9.9- $\mu$ m-thick active region, and used a SI-surface-plasmon waveguide with an 800- $\mu$ m-thick lower contact layer doped at  $n = 2 \times 10^{18}$  cm $^{-3}$ . The waveguide was calculated to have  $\alpha_w = 12.7$  cm $^{-1}$  and  $\Gamma = 0.49$ . Devices were fabricated both using wet etching in NH<sub>4</sub>OH:H<sub>2</sub>O<sub>2</sub>:H<sub>2</sub>O (20:12:480), and also using dry etching in BCl<sub>3</sub>:N<sub>2</sub> (15:5 sccm). For the dry etched devices, after the processing was complete, the devices were placed in a 1 minute wet etch (same as above), to etch  $\sim$ 400 nm so as to cleanup any sidewall contamination. However, it is likely that this etch also undercut the top contact for several microns at the metal/GaAs interface (see Sec. 5.2.1). Lateral alloyed ohmic contacts were deposited  $\sim 10$   $\mu$ m away from the edge of the ridge.

Transport and emission characteristics for T215G devices are shown in Fig. 6-20. The peak current density was  $J_0 \approx 360$  A/cm $^2$  at a bias of  $\sim$ 6.3 V. The spectral behavior is consistent with that seen from previous devices. Although relatively narrow linewidths were obtained (2.5–3.0 meV), no particular evidence of gain narrowing was

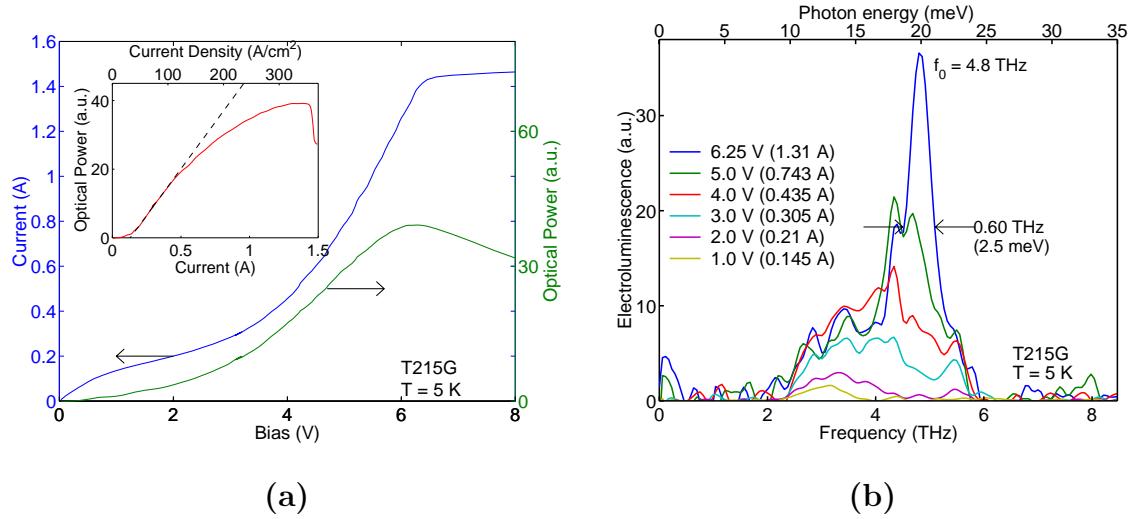


Figure 6-20: (a)  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  plots at 5 K for T215G edge emitting device 150- $\mu$ m wide and 2.70-mm long. (b) Edge emitting spectra from same device taken at 5 K at 5% duty cycle using the photodetector with a spectral resolution of 8  $\text{cm}^{-1}$ .

Table 6.4: Measured spontaneous output powers for various T215G edge emitting devices collected in face-to-face mode with the photodetector. Power was measured at 5% duty cycle (50% duty pulse train at 10 kHz, with 332 ns pulses at 300 kHz) but scaled up by a factor of 10 to be comparable with that listed in Tables 6.2 and 6.3.

Device	Size	Peak Current (A)	Peak Power ( $\mu$ V)
T215G	150 $\mu$ m $\times$ 2.7 mm	1.45	430
T215G	80 $\mu$ m $\times$ 4.025 mm	1.05	210

observed. The output emission powers from various devices are listed in Table 6.4.

#### 6.4.6 T215H

##### SI-surface-plasmon waveguide

The active region design for T215H was the same as T150E, T150F, and T150H, except 3.4 nm (12 ML) injection barriers were used. Injection anticrossings were calculated to be of  $\Delta_{2'3} = 2.0$  meV and  $\Delta_{1'3} = 1.8$  meV, using a 65% conduction band offset. X-ray analysis of this growth revealed an undergrowth by 4.0%. The device was grown to use a SI-surface-plasmon waveguide, using an 500- $\mu$ m-thick lower

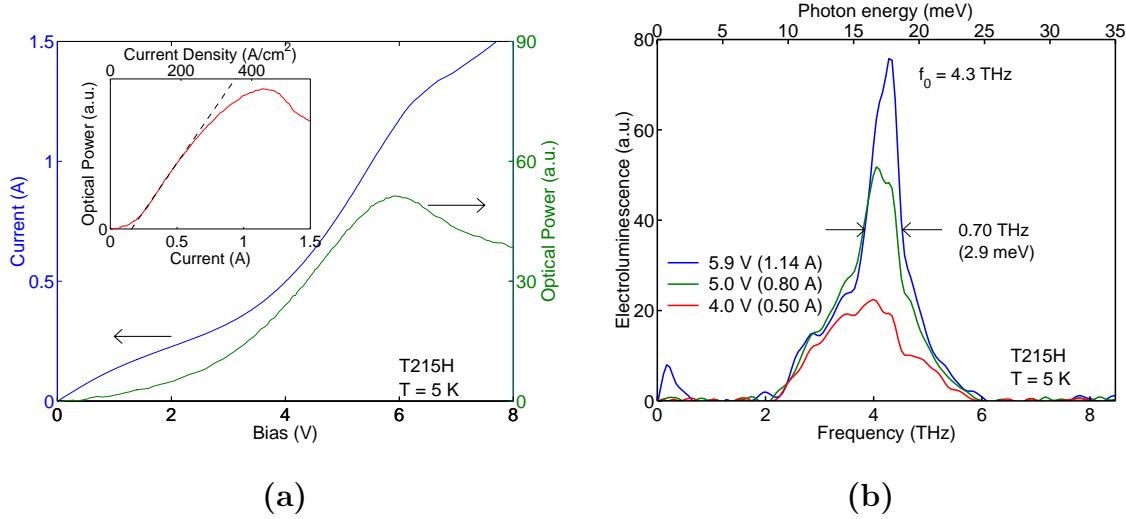


Figure 6-21: (a)  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  plots at 5 K for T215H edge emitting device 100  $\mu\text{m}$  wide and 2.65 mm long. (b) Edge emitting spectra from same device taken at 5 K at 5% duty cycle using the photodetector with a spectral resolution of  $8 \text{ cm}^{-1}$ .

contact layer doped at  $n = 3 \times 10^{18} \text{ cm}^{-3}$ . The waveguide was calculated to have  $\alpha_w = 7.0 \text{ cm}^{-1}$  and  $\Gamma = 0.40$ . Devices were fabricated both using wet etching in  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (20:12:480) and dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm), although only the wet-etched devices are described here. Lateral alloyed ohmic contacts were deposited  $\sim 50 \mu\text{m}$  away from the edge of the ridge.

Transport and emission data from T215H is shown in Fig. 6-21. The design bias point at  $\sim 6 \text{ V}$  is no longer as clear as before, and merely appears as a kink in the  $I$ - $V$ . This suggests that the injection barrier is becoming sufficiently thin that the parasitic coupling between the upper radiative state  $n = 3$  and the excited states in the next module is significant. This is consistent with the behavior observed for T150C, whose injection barrier was slightly thinner at 2.8-nm thick. The peak current at this design point is  $J_0 \approx 500 \text{ A/cm}^2$ . The spectral behavior is consistent with that seen before, although the primary peak is at 18 meV (4.3 THz) instead of at 20 meV for the previous devices. The reason for this change is not clear, although it may be that this device was grown with thicker layers than the others. Since this device was measured to be 4% under-grown, this would suggest that the previous devices were even more severely under-grown. Since there is no x-ray data available for the other

Table 6.5: Measured spontaneous output powers for various T215H edge emitting devices collected in face-to-face mode with the photodetector. Power was measured at 5% duty cycle but scaled up by a factor of 10 to be comparable with that listed in Tables 6.2 and 6.3.

Device	Size	Peak Current (A)	Peak Power (a.u.)
T215H	150 $\mu\text{m}$ $\times$ 2.65 mm	1.72	700
T215H	100 $\mu\text{m}$ $\times$ 2.65 mm	1.2	500
T215H (HR-coated)	150 $\mu\text{m}$ $\times$ 1.98 mm	1.3	1000

devices in the T-series, this speculation cannot be confirmed.

Peak electroluminescence powers from T215H devices are listed in Table 6.5. Although no evidence of gain was observed in this device, it is notable that the device that had the rear facet HR-coated displayed significantly more power than the other devices. Although there is uncertainty due to optical alignment, this suggests that the cavity is relatively transparent over its length ( $\alpha_w - g \leq 5 \text{ cm}^{-1}$ ).

#### 6.4.7 Peaks of lateral modes

In several devices (T150E-2mm, T150Fmm, FL125mm) with metal-metal waveguides and dry-etched vertical sidewalls, the electroluminescence spectrum was observed to break up into several evenly spaced narrow peaks. The dips in the spectrum that separate the peaks do not correspond to atmospheric absorption lines, although the presence of absorption adds some uncertainty to the measured peak positions. Furthermore, the spacing varies in a way that is well correlated with the width of the ridge. For example, for a 100- $\mu\text{m}$ -wide ridge T150E-2mm, the peaks in Fig. 6-17(b) are separated by approximately  $13 \text{ cm}^{-1}$ . This is in reasonably good agreement with a Fabry-Pérot mode spacing of  $\Delta\nu = (2ln_{eff})^{-1} = 12.8 \text{ cm}^{-1}$  if a cavity length of  $l = 100 \mu\text{m}$  and refractive index of  $n_{eff} = 3.9$  is used. This value for  $n_{eff}$  was measured in a metal-metal waveguide laser structure (Sec. 7.3), albeit at a lower frequency of 3.0 THz. Spectra measured from an 80- $\mu\text{m}$ -wide ridge displayed a mode spacing of approximately  $16 \text{ cm}^{-1}$ , which corresponds well with a calculated  $\Delta\nu = 16.0 \text{ cm}^{-1}$ . The same peaks spaced by  $13 \text{ cm}^{-1}$  are observed for a 100- $\mu\text{m}$ -wide ridge T150Fmm

(Fig. 6-19), and a 60- $\mu\text{m}$ -wide ridge displayed peaks separated by 23  $\text{cm}^{-1}$ , in a good agreement with the calculated value of 21.4  $\text{cm}^{-1}$ . Similar results are seen for FL125mm (Sec. 6.5.1). This is puzzling, since light subject to such a Fabry-Pérot cavity between the sidewalls would be traveling in an approximately perpendicular direction to the longitudinal cavity, and thus should not escape the dewar and travel the FTIR beam path to be collected by the detector. One possibility is that the light is traveling at a slight angle to the waveguide, which then refracts upon exit from the cavity to travel nearly parallel to the ridge, where it would be captured by the cone and coupled out of the dewar. For example, light traveling at only an angle of 11° to the lateral inside GaAs will refract to an angle of almost 45° in air. Since the small cone aperture is approximately 1.6 mm in diameter and is positioned less than 1 mm from the ridge, a substantial amount of radiation might be collected.

If these modes are in fact of a result of a lateral Fabry-Pérot effect, this provides confirmation of the smaller loss of the metal-metal waveguides compared to the plasmon waveguides. For a Fabry-Pérot cavity, modes should only be visible if the total loss ( $\alpha_w - g$ ) is less than the cavity facet loss. This is equivalent to a Hakki-Paoli measurement of gain [169]. Since for a 100- $\mu\text{m}$ -long GaAs cavity,  $\alpha_m \approx 114 \text{ cm}^{-1}$ , we cannot obtain very precise information about the loss, except to say that the waveguide loss for plasmon waveguides is on the order of  $\alpha_m$  or greater, which is consistent with Drude model waveguide calculations. While there is no evidence of a population inversion, level  $n = 2$  is not overly full compared to level  $n = 3$ , as only  $n_2 - n_3 \approx 4 \times 10^9 \text{ cm}^{-2}$  is needed to obtain 100  $\text{cm}^{-1}$  of intersubband loss (for  $\Delta\nu = 0.72 \text{ THz}$  (3 meV)). Fringes such as these were never observed for SI-surface-plasmon waveguides despite their low loss. This is attributed to the fact that most of these guides had wet-etched sloped sidewalls. Also, in a SI-surface-plasmon waveguide, the mode extends substantially into the substrate, and thus does not see such a precise reflection at the sidewall/air interface.

#### 6.4.8 T-series summary

A summary of peak current densities and injection anticrossings for T-series devices is listed in Table 6.6. The peak current density increases steadily increases as the injection barrier is made thinner. While this may be partially affected by parasitic state coupling for the devices with the thinnest barriers (T150C, T215H), the dependence of  $J_0$  on anticrossing gap is a clear indication of the incoherent sequential nature of the tunneling injection for nearly all these devices. Thus the coherent model pictured in Fig. 6-6(b) is inappropriate. Since next module parasitic coupling was already overwhelming for the device with the narrowest injection barrier (T150C), it is unlikely that such a coherent model is attainable. We can expect injection to be incoherent for other devices with similar injection anticrossings that have shorter upper state lifetimes. For example, the resonant-phonon devices (FL-series) suffer from a parasitic LO-phonon scattering channel that is not present in the T-series devices; thus all other things being equal, the upper state lifetime in resonant-phonon devices should be shorter.

A clear anticrossing gap was observed between states  $n = 2$  and  $1$  via magneto-tunneling spectroscopy for T65, which provides evidence for coherent interaction between the lower states. For edge emitting structures, coarse evidence of the 2-1 anticrossing was also observed in the emission spectra, as the evolution of the peaks with bias corresponded qualitatively well to the predicted behavior of the energy separations and oscillator strengths. Optical measurements yielded very narrow intersubband emission linewidths, both at 5 K and at 77 K. The narrowest were  $\sim 0.7$  meV, measured from T65 at 5 K from surface-emitting grating structures. For thicker plasmon waveguides,  $\alpha_w$  was lower and edge emission was feasible, and linewidths as narrow as 2.0 meV at 5 K and 2.6 meV at 78 K (T150E-3) were observed. The difference in linewidths was likely due to the larger number of modules, as well as due to the effects of the grating. No clear evidence of gain was observed, either through linewidth narrowing or superlinear  $L$ - $I$  characteristics. However, low loss cavities were obtained using metal-metal and SI-surface-plasmon waveguides as confirmed

Table 6.6: Injection barrier thicknesses  $t_{inj}$  along with injection anticrossings (calculated for 65% band offset) and peak injection current density  $J_0$  for T-series devices.

Device	$t_{inj}$ (nm)	$\Delta_{2'3}$ (meV)	$\Delta_{1'3}$ (meV)	$J_0$ (A/cm <sup>2</sup> )
T65	5.9	0.75	0.5	120
T150C	2.8	1.8	2.0	1500
T150E	5.4	0.7	0.9	160
T150E-2	5.4	0.7	0.9	130
T150E-3	5.4	0.7	0.9	160
T210E	5.4	0.7	0.9	140
T150F	4.5	1.3	1.1	250
T215G	4.0	1.6	1.4	370
T215H	3.4	2.0	1.8	500

by measurements of total collected power and the emergence of lateral Fabry-Pérot modes in the spectra for metal-metal devices. For later generation devices such as T215G and T215H, there was likely not an overwhelming intersubband absorption at design bias, and thus the devices may not have been far from transparency.

Examination of  $L$ - $I$  characteristics revealed clear sublinear behavior or saturation of the upper state population  $n_3$  with current. This is a clear indication the reduction of the lifetime  $\tau_3$  at high injection levels, which is due to a combination of e-e scattering, thermally activated LO-phonon scattering, or the alignment of the upper state with a parasitic level at design bias. In Ref. [61], Rochat *et al.* fit their  $L$ - $I$  from a similar structure with a square root behavior and concluded that this indicated an e-e scattering limited lifetime. The  $L$ - $Is$  in the T-series structures are poorly fit by such a square root behavior, which suggests a different limiting mechanism.

Since gain in this structure depends on a cold electron distribution within the 2-1 doublet, lasing will probably be difficult to achieve unless the anticrossing  $\Delta_{21}$  is made much larger. As mentioned in Sec. 2.4, even at liquid helium lattice temperatures, during operation the electron gas will become significantly heated. This is necessary to allow LO-phonon emission to cool the electron distribution and balance the dissipated power [83]. This is confirmed by Monte Carlo simulations of the T65 device performed by Hans Callebaut, where the electron temperature ranges from 70–100 K [94]. As

such, a cold electron distribution should not be expected, and sufficient gain for lasing will be difficult to ever obtain without using  $\Delta_{21} \geq k_b T_e$ . However, the use of a larger energy separation will hinder fast e-e scattering between 2 and 1, which is necessary to obtain a thermal distribution. Alternatively, more wells and more lower states could be added to the structure, effectively obtaining a wider miniband. Following this strategy to its logical conclusion however will lead to a chirped-superlattice or bound-to-continuum design, where miniband depopulation is used.

## 6.5 FL series (resonant phonon)

### 6.5.1 FL125

FL125 was the first in the series of resonant-phonon depopulation devices, which would eventually result in a terahertz laser. The design philosophy was to use the fast depopulation resulting from intrawell phonon scattering combined with the selective depopulation of a diagonal radiative transition (M series) and the larger oscillator strength of a vertical radiative transition (L series). The conduction band and anticrossing diagrams are shown for this device in Fig. 6-22, calculated using both 65% and 80% conduction band offsets. The lower radiative state ( $n = 4$ ) was selectively delocalized by anticrossing it with the excited state in the wide quantum well ( $n = 3$ ), resulting in fast scattering down to the injector states. LO-phonon scattering from  $5 \rightarrow (2, 1)$  is suppressed due to the small wavefunction overlap. This method for depopulation was discussed in earlier work by Stroscio *et al.* [170] and Callebaut [171]. At the 4-3 anticrossing at 49 mV/module, the anticrossing gap is  $\Delta_{43} = 2.7$  meV,  $\tau_4 \approx \tau_3 \approx 0.4$  ps, and  $\tau_{5 \rightarrow (2,1)} \approx 30$  ps (calculated with 80% conduction band offset).

The design for FL125 was originally performed with 65% conduction band offset (Fig. 6-22(a-b)), but due to the transport problems discussed below, the barrier height was reconsidered and the calculations were repeated with 80% offset (Fig. 6-22(c-d)). For the 80% offset case, the 4-3 anticrossing occurs at 49 mV/module, when in fact it ought to occur closer to 1'-5 anticrossing at 58 mV/module, i.e. the design bias. As

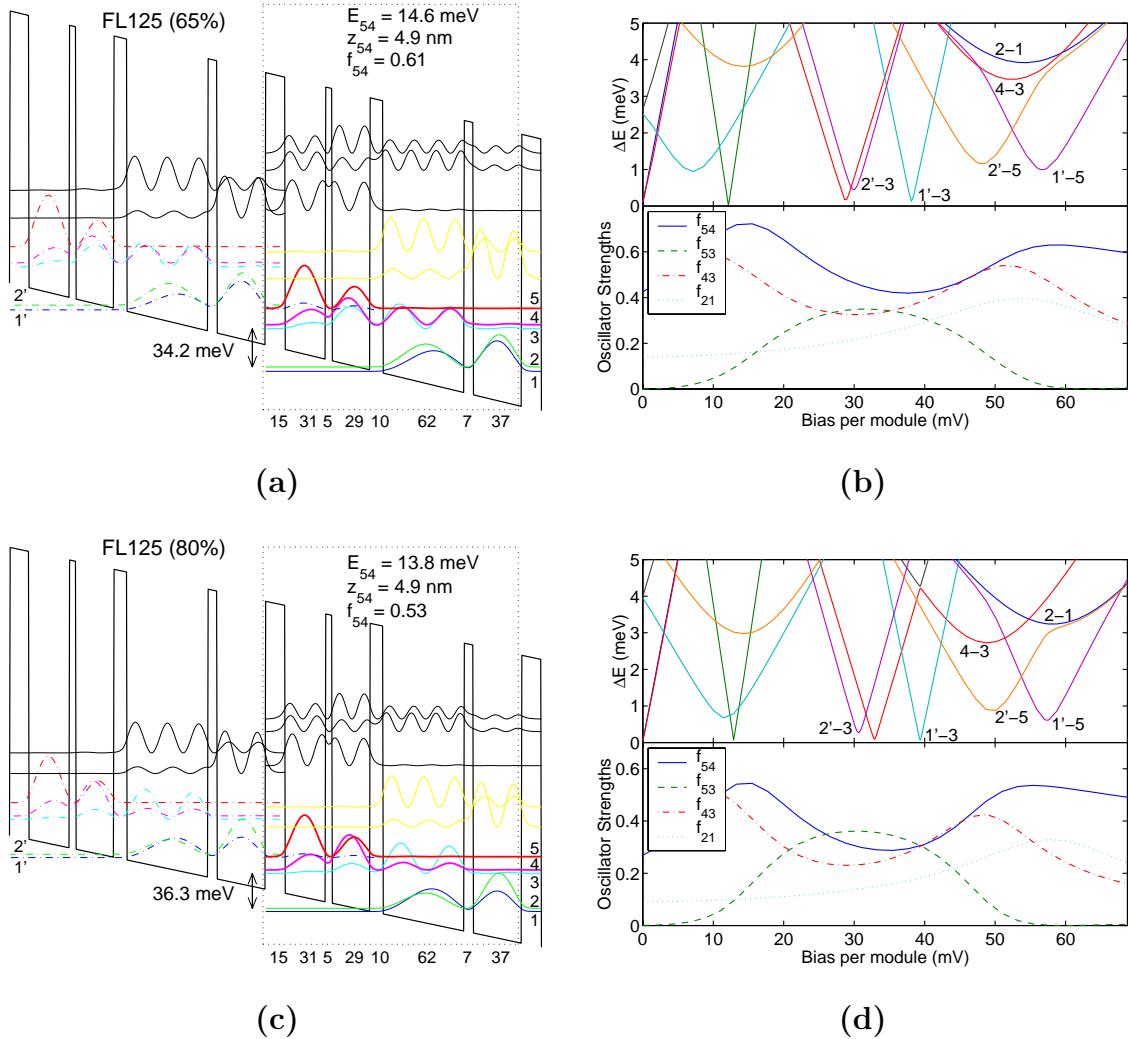


Figure 6-22: Conduction band diagram and anticrossing diagram of FL125 grown in GaAs/Al<sub>0.3</sub>Ga<sub>0.7</sub>As calculated with (a), (b) 65% conduction band offset, and (c), (d) 80% offset. Layer thickness are given in monolayers. The 62 ML well was doped at  $2.3 \times 10^{16} \text{ cm}^{-3}$  to yield a sheet density of  $4.0 \times 10^{10} \text{ cm}^{-2}$ , however  $C-V$  measurements indicate a doping of  $3.4 \times 10^{10} \text{ cm}^{-2}$  per module.

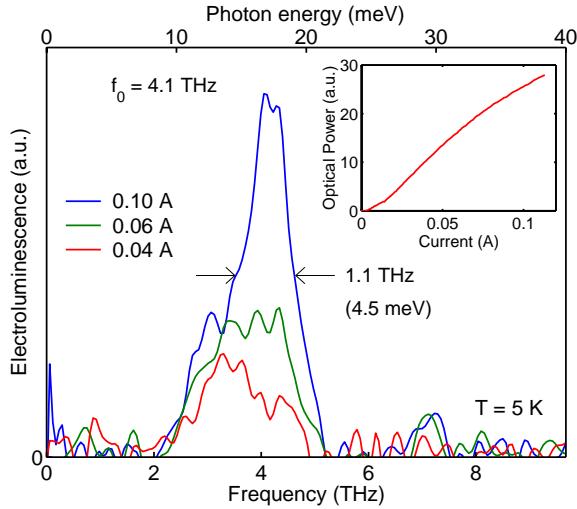


Figure 6-23: Spectra and  $L$ - $I$  from 145- $\mu\text{m}$ -wide, 260- $\mu\text{m}$ -long edge emitting FL125 device at  $T = 5 \text{ K}$  measured with the photodetector at  $8 \text{ cm}^{-1}$ .

discussed below, this simulation provided a physical picture that was more consistent with our experimental results, and the 80% barrier height was adopted for future designs. However, no x-ray diffraction data was taken for this sample, so it is difficult to rule out over- or under-growth.

## Plasmon waveguide

125 modules were grown on an  $n^+$  GaAs substrate with  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  etch stop layers. Fabrication into ridge structures approximately 7  $\mu\text{m}$  high took place via wet etching. As a result of the etch stop layers, at 5 K substantial carrier freeze out occurred, leading to a variable parasitic resistance. For this reason, transport data is deferred until the discussion on metal-metal devices below. However, some electroluminescence data was obtained and is shown in Fig. 6-23. Immediately before NDR (0.10 A or 265  $\text{A}/\text{cm}^2$ ), a relatively broad peak (FWHM=4.5 meV) was observed centered at approximately at 17 meV (4.1 THz), somewhat higher than the calculated value of 13.8 meV. Comparison of spectra at lower bias reveals a slight Stark shift and narrowing of the peak.

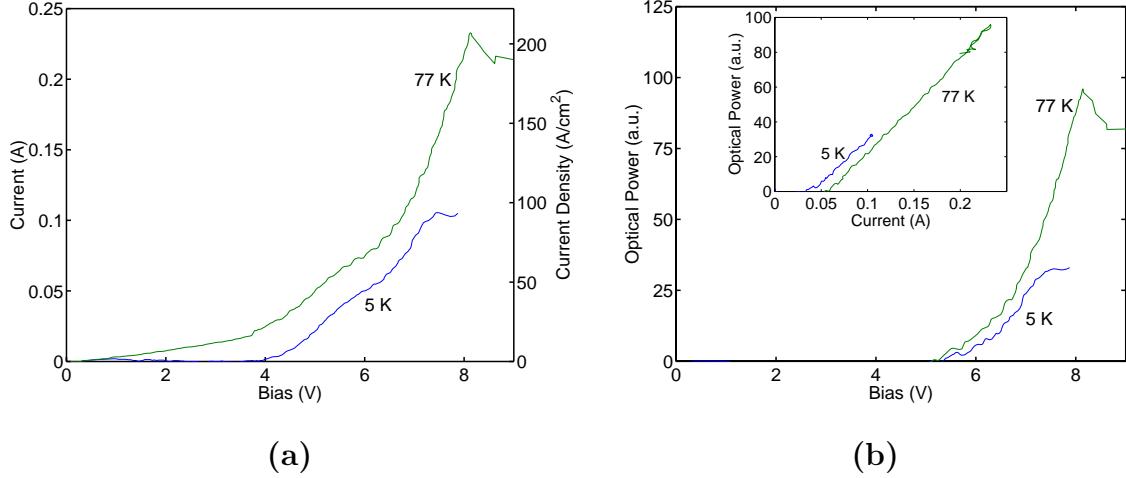


Figure 6-24: (a) *I-V* and (b) *L-V* and *L-I* (inset) for 100- $\mu\text{m}$ -wide, 1.125-mm-long FL125mm edge emitting device with metal-metal waveguide taken with 5% duty cycle.

### Metal-metal waveguide (FL125mm)

The FL125mm structure was fabricated with a metal-metal waveguide using Au-Au thermocompression bonding discussed in Sec. 5.3. Upper Schottky contacts were formed by evaporating Ti/Au (200/4000 Å) layers, and these contacts were used as etch masks for dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm). Silicon oxide was deposited using low-temperature PECVD to cover the mesa sidewalls. The calculated waveguide loss was  $\alpha_w = 13.9 \text{ cm}^{-1}$  with a confinement  $\Gamma = 0.96$ . No particular fabrication problems were noted except for some adhesion problems at the bonding layers for some devices. Gold wires were attached to the top of the mesas using silver epoxy.

Transport through the FL125mm devices exhibited an unusual temperature dependence that was indicative of level misalignment. *I-V* and *L-V* characteristics taken at 5 K and 77 K are displayed in Fig. 6-24. The peak current density  $J_{max}$  increased strongly with temperature. In and of itself, this would not be notable, as this phenomenon would typically be due to the reduction in the radiative state lifetime  $\tau_5$  due to thermally activated LO-phonon scattering. However, in this case, the increase in  $J_{max}$  was accompanied by a concomitant increase in the emitted electroluminescence as well as a slight increase in the NDR voltage. Since the optical power

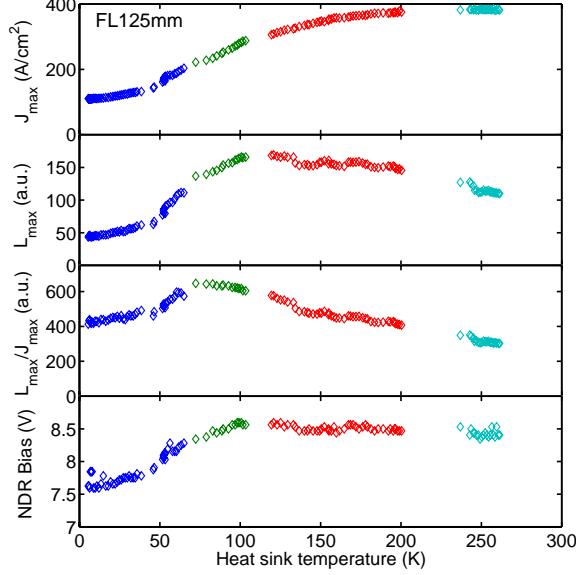


Figure 6-25: Current and optical power at NDR, and NDR bias, for 170- $\mu\text{m}$ -wide, 0.76-mm-long edge emitting FL125mm device taken at 5% duty cycle versus heat sink temperature.

is proportional to the upper state population, this reflects an increase in upper state injection current with temperature, rather than a decrease in upper state lifetime. This is confirmed by the very similar slope efficiencies of the  $L$ - $I$  characteristics at 5 K and 77 K shown in the inset of Fig. 6-24(b). This is shown to a greater extent in Fig. 6-25, where the current density and optical power at NDR is plotted versus heat sink temperature. Above 100 K, the NDR voltage remains fixed, and the injection ceases to improve. Further increase in current is a result of the reduction in  $\tau_5$ , which causes the optical power to drop. It should be noted that contamination of the intersubband emission with blackbody radiation at elevated temperatures is also possible, and may suppress the roll-off in  $L_{max}$  above 100 K.

Spectra at 5 K and 77 K heat sink temperatures are shown in Fig. 6-26. Because of the additional heating caused by the 50% duty cycle operation, the devices were able to be biased at higher bias points than indicated by Fig. 6-24. The spectra are composed of many narrow peaks. These narrow peaks should not be confused with the actual spontaneous emission spectra, as they are associated with the dry etched vertical sidewalls in a metal waveguide and are likely due to a Fabry-Pérot

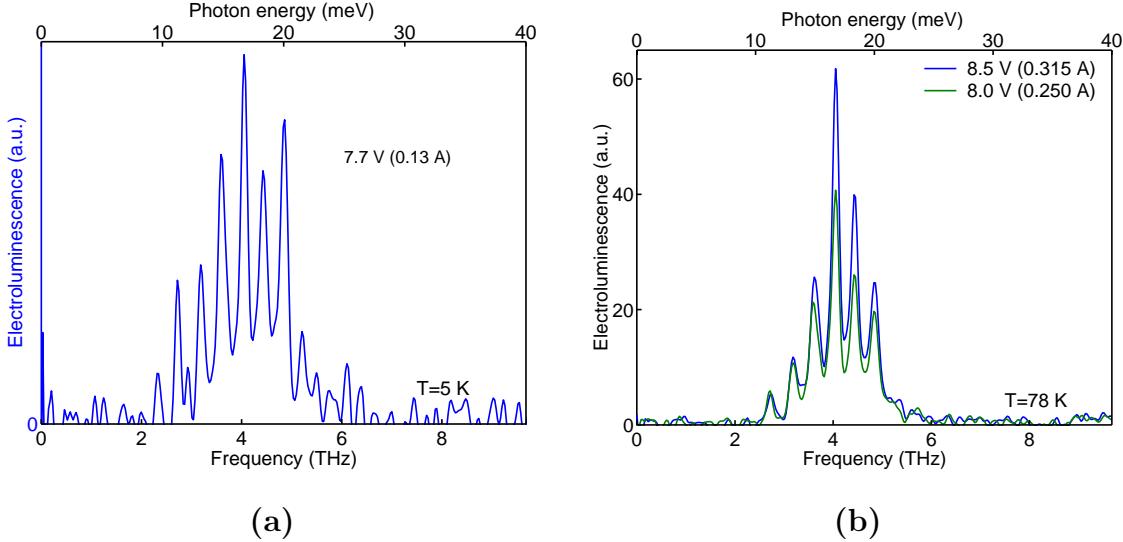


Figure 6-26: Spectra for 100- $\mu\text{m}$ -wide, 1.125-mm-long FL125mm edge emitting device with metal-metal waveguide taken with 50% duty cycle at (a) 5 K and (b) 78 K heat sink temperature. Spectra were taken with photodetector with  $4 \text{ cm}^{-1}$  resolution.

effect (see Sec. 6.4.7). The true spontaneous emission spectrum can be estimated by the overall envelope of the narrow peaks, which is in fact comparable with the spectra in Fig. 6-23 from the plasmon waveguide. At 5 K, the spectrum is quite broad, with an envelope FWHM  $\approx 6\text{--}7 \text{ meV}$ . At 77 K, the device can be biased higher, and the spectral envelope becomes narrower (FWHM  $\approx 4 \text{ meV}$ ), which is slightly narrower than in the plasmon waveguide. This indicates that the increase in NDR bias with temperature is in fact real, and not due to a parasitic resistance outside the active region.

One likely explanation for the unusual transport behavior is that there is some subband misalignment in the injector regions (states  $n = 1, 2$ ). If the ground state associated with the wide well is lower than expected, then at NDR a delocalized doublet will not have formed in the injector, but rather many of the carriers will be trapped in the wide well. Furthermore, due to the thick injection barrier the injection anticrossing gaps are small ( $\Delta_{2'5} = 0.8 \text{ meV}$  and  $\Delta_{1'5} = 0.6 \text{ meV}$  calculated with 80% barrier heights). It is likely that the NDR occurs at the 2'-5 anticrossing, before injection from  $n = 1'$  can take place. As the electron temperature rises, carriers will be thermally excited into the injector state  $n = 2'$  in the narrower well, thus increasing

current density. Furthermore, as the temperature rises the level broadening will increase, and the structure will sustain higher biases before breaking and entering an NDR region, potentially even allowing injection from the lower  $n = 1'$  injector state. This scenario explains the experimentally observed increase of NDR voltage with temperature.

The simulation with 80% barrier height produces results that correspond more closely to this scenario than does the 65% simulation. However, the possibility of other growth problems such over- or under-growth of the epitaxial layers cannot be ruled out. The design remedy is to thin the widest well to raise both the ground and excited states. This will move the 4-3 anticrossing closer to the design bias, i.e. the 1'-5 anticrossing. It will also reduce any misalignment between states  $n = 2$  and 1. In the revision of this design FL175B, this step was taken, and the 10 ML and 7 ML barriers were thinned as well to increase coupling. Aside from the transport difficulties, this design was considered promising, particularly because of the linear  $L$ - $I$  characteristic which suggested a constant upper state lifetime, even at high injection levels.

### 6.5.2 FL175B

As discussed in the previous section, FL175B was a redesign of FL125, in which the widest well was thinned from 62 to 60 ML, and the 10 ML and 7 ML barriers were reduced to 9 ML and 6 ML respectively. The intention was to keep the radiative transition and injector barrier the same, while removing the transport bottleneck associated with the previous device. The conduction band diagram and anticrossing diagram, calculated using a 80% conduction band offset, is shown in Fig. 6-27. The 4-3 anticrossing is now in better alignment with the 1'-5 design bias,  $\Delta_{43}$  increased from 2.7 meV to 3.6 meV, and  $\Delta_{21}$  increased from 3.2 meV to 4.0 meV. The injection anticrossing was  $\Delta_{1'5} = 0.75$  meV.

FL175B was fabricated using a SI-surface plasmon waveguide, with a 0.8- $\mu\text{m}$ -thick contact layer doped to  $n = 2 \times 10^{18} \text{ cm}^{-3}$ . 175 modules were grown to form an active region approximately 9.5  $\mu\text{m}$  thick. Wet etching in  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  was used to

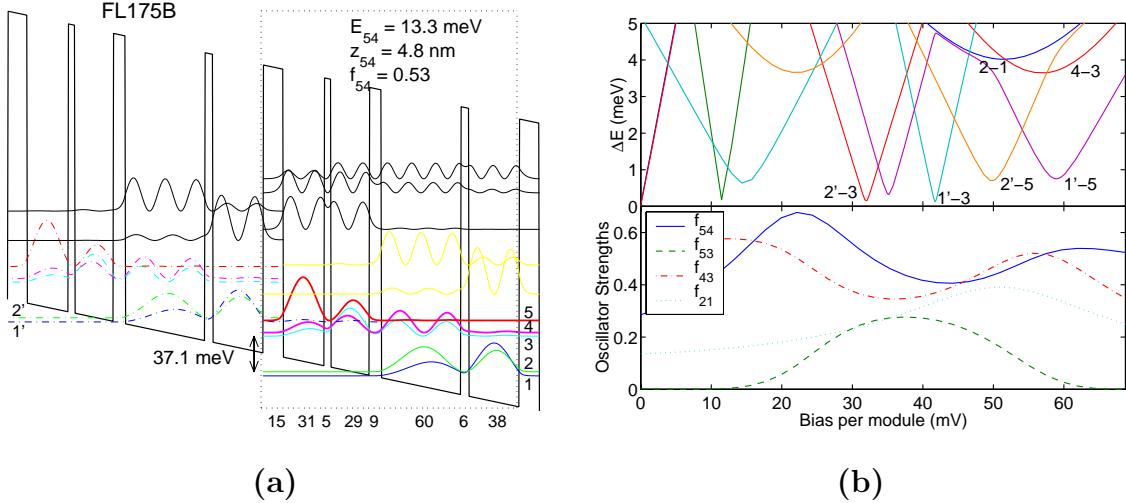


Figure 6-27: (a) Conduction band diagram of FL175B grown in GaAs/Al<sub>0.3</sub>Ga<sub>0.7</sub>As calculated with 80% offset. Layer thicknesses are given in monolayers. The 60 ML well was doped at  $2.3 \times 10^{16}$  cm<sup>-3</sup> to yield a sheet density of  $3.9 \times 10^{10}$  cm<sup>-2</sup> per module. (b) Anticrossing diagram and oscillator strengths versus bias.

fabricate ridge structures, and lateral contact were deposited with approximately a 10-μm separation from the ridge. The calculated waveguide loss was  $\alpha_w = 11.9$  cm<sup>-1</sup> with a confinement  $\Gamma = 0.47$ .

Examination of the transport characteristics indicate that the design changes were successful.  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  characteristics are shown in Fig. 6-28 at 5 K and 77 K. The peak current density at 5 K is now 320 A/cm<sup>2</sup> compared to 95 A/cm<sup>2</sup> for FL125mm. The increase in current with temperature was very mild, and it was accompanied by a decrease in emitted optical power. This is consistent with constant injection and an upper state lifetime that decreases with temperature. The current density and optical power at the NDR point are plotted versus temperature in Fig. 6-29.

Electroluminescence was observed in an edge emitting structure, and is shown in Fig. 6-30(a). While the spectra were similar to those taken from FL125, the linewidths measured were extremely broad, 7.4 meV just below the NDR point. This broadening was likely inhomogeneous, perhaps caused by module-to-module inhomogeneity. The emission frequency is largely determined by the anticrossing gap between states

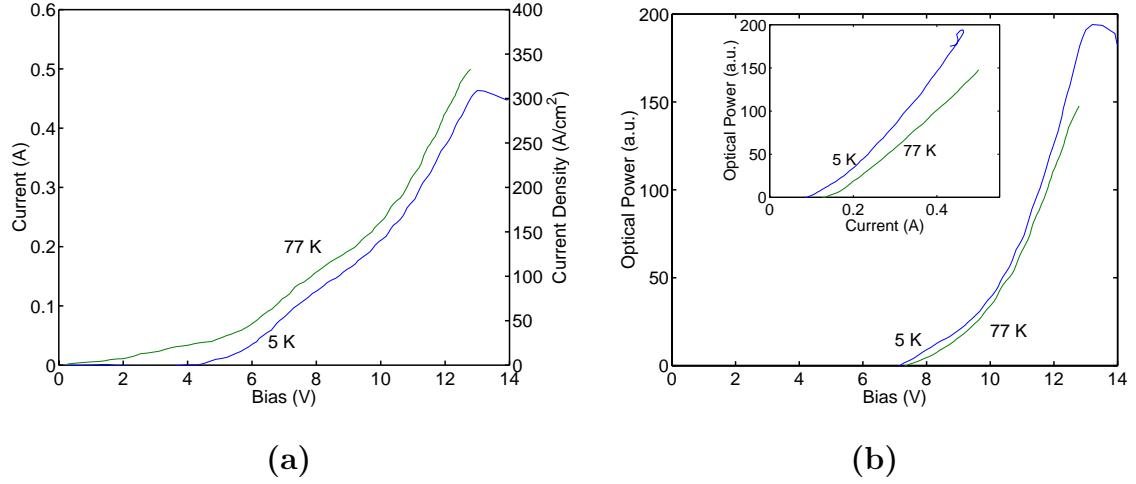


Figure 6-28: (a)  $I$ - $V$  and (b)  $L$ - $V$  and  $L$ - $I$  (inset) for 200- $\mu\text{m}$ -wide, 0.75-mm-long FL175B edge emitting device with SI-surface-plasmon waveguide taken with 50% duty cycle.

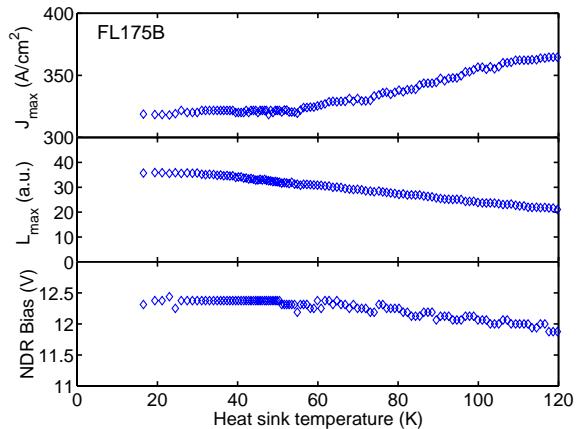


Figure 6-29: Current and optical power at NDR, and NDR bias, for 150- $\mu\text{m}$ -wide, 2.4-mm-long edge emitting FL175B device taken at 5% duty cycle versus heat sink temperature.

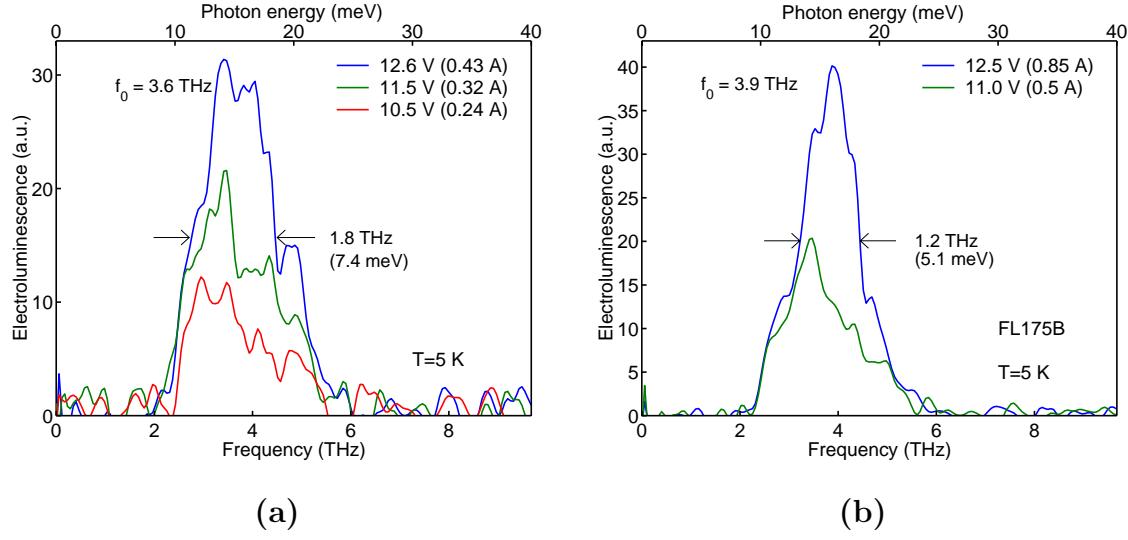


Figure 6-30: (a) Spectra from 200- $\mu\text{m}$ -wide, 0.75-mm-long edge emitting FL175B device with 10- $\mu\text{m}$  lateral contact separation and (b) spectra from 150- $\mu\text{m}$ -wide, 2.0-mm-long, rear HR coated edge emitting FL175B device with 50- $\mu\text{m}$  lateral contact separation. Data was at 5 K taken with photodetector at  $8 \text{ cm}^{-1}$  resolution.

$n = 5$  and 4, which is set by the 5 ML barrier. Monolayer fluctuations in this barrier thickness can cause shifts in the radiative energy, which produces a large spontaneous emission linewidth. Since the transition is slightly diagonal as well, bias inhomogeneity between modules may also contribute to broadening. However, because of the large spectral width, blackbody contamination of the spectra cannot be definitively ruled out. Due to the SI-surface plasmon waveguide, and the sloped sidewalls, no lateral mode structure was observed in the spectra as for FL125mm.

This wafer was more recently reprocessed using wet etching in  $\text{H}_3\text{PO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (10:10:250), and a new mask was used so that the lateral contacts were positioned 50  $\mu\text{m}$  away from the ridge. This was expected to reduce loss caused by modal coupling to lossy surface plasmons associated with the lateral contacts. The rear facet was HR coated with Ti/Au evaporated over  $\text{Al}_2\text{O}_3$ . Lasing was not seen in this device either, although a narrower spontaneous emission linewidth of 5.1 meV was observed (Fig. 6-30(b)). This may indicate a slight linewidth narrowing due to gain.

Other than the broad linewidth, this design was considered to be a good candidate for a laser. The  $L$ - $I$  characteristic showed no signs of sublinearity. No concrete

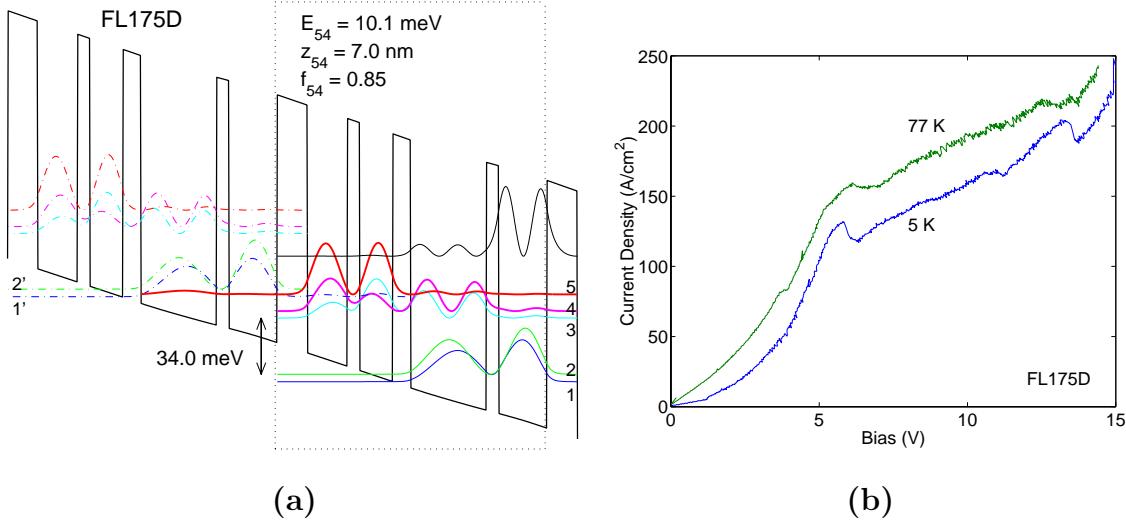


Figure 6-31: (a) Self-consistent conduction band profile of FL175D calculated using 80% conduction band offset. The device is grown in the GaAs/Al<sub>0.15</sub>Ga<sub>0.85</sub>As material system. Beginning with the left injection barrier, the layer thicknesses in Å are 65/86/27/71/38/163/27/104 (as grown 4.7% above design). The 163 Å well is doped at  $n = 1.9 \times 10^{16}$  cm<sup>-3</sup>, which yields a sheet density of  $3.1 \times 10^{10}$  cm<sup>-2</sup> per module. (b)  $I$ - $V$  taken at 5 K and 77 K of a  $200 \times 200 \mu\text{m}^2$  FL175D diode taken with 400-ns pulses at a 1-kHz pulse repetition frequency.

evidence for gain was seen, although some evidence of linewidth narrowing was seen when the lateral contact separation from the ridge was widened from 10  $\mu\text{m}$  to 50  $\mu\text{m}$ . Two-dimensional mode simulations have shown that when the lateral contacts are too close, the mode can couple to lossy surface plasmons associated with the contacts, which reduces the confinement factor and increases the loss [172]. The next modification of this device was FL175C, in which the aluminum content of the barriers was reduced to 15%. This allowed the use of thicker barriers to achieve the same anticrossing gaps, which reduces the effect of monolayer fluctuations. The device FL175C was the first terahertz QCL at MIT to successfully lase, and is described in detail in Sec. 7.2.

### 6.5.3 FL175D

The FL175D design (shown in Fig. 6-31(a)) was designed to be identical to FL175C, except with a 22 ML (62 Å) injection barrier instead of a 20 ML (57 Å) injection

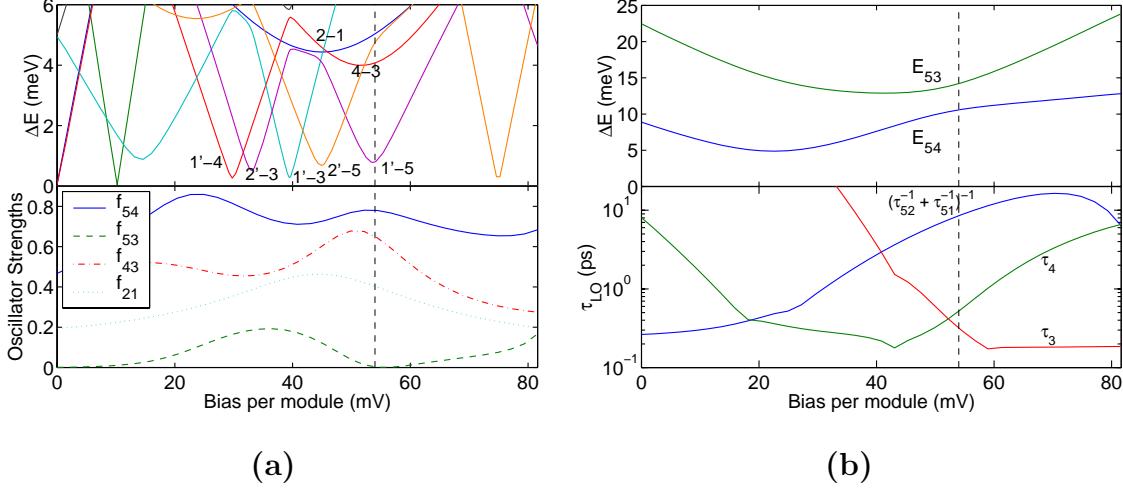


Figure 6-32: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL175D. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction.

barrier. However, while the FL175C device was grown 4.2% thinner, FL175D was grown 4.7% thicker, which resulted in a structure that behaved very differently. For example, the injection anticrossing gap changed from  $\Delta_{1'5} = 1.8$  meV for FL175C to  $\Delta_{1'5} = 0.8$  meV for FL175D. The calculated anticrossing gaps, oscillator strengths and LO-phonon scattering times are shown in Fig. 6-32.

The device was processed into a semi-insulating waveguide with a  $0.8\text{-}\mu\text{m}$ -thick contact layer doped to  $n = 2 \times 10^{18} \text{ cm}^{-3}$ . Self aligned dry etching using the Ti/Au (200/4000 Å) top contacts in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm) was used to define the mesas, followed by a short cleanup etch in  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (10:6:480). A pulsed  $I$ - $V$  from a small  $200 \times 200 \mu\text{m}^2$  device is shown in Fig. 6-31(b). An NDR occurs at  $\sim 6$  V whereupon high-field domains developed. Beyond this NDR, which likely corresponds to the  $1' - 3$  parasitic channel, the  $I$ - $V$  characteristics tended to vary from device to device, and even run to run. No spectra were taken.

#### 6.5.4 FL179E, FL162G, and FL137H

The devices FL179E, FL162G, FL137H, and FL152F were the next generation of devices designed after FL175C. They were designed to operate at approximately

2.5 THz. Only FL152F successfully lased (see Sec. 7.3). Device FL179E was most like FL175C, and had a two-well active region and a two-well injector. FL152F had a two-well active region and a three-well injector, FL162G had a three-well active region and a two-well injector, and FL137H had a three-well active region and a three-well injector. As part of the design philosophy, most of the injector barriers were substantially thinned, which increased the injector and parasitic injector anticrossing gaps, as shown below. This design choice proved to be misguided, because all of these devices turned out to have very high parasitic channel currents.

These devices were grown on semi-insulating substrates with 500-nm  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  etch-stop layers to accommodate processing for either SI-surface-plasmon or metal-metal waveguides. The lower contact layer was 600-nm thick and doped at  $2 \times 10^{18} \text{ cm}^{-3}$ . However, during the processing for the SI-surface-plasmon devices, the active region was incompletely etched before depositing the lateral ohmic contacts. As a result, these devices had large parasitic resistances that made it very difficult to reach design bias. Therefore, most of the data is shown for metal-metal waveguide devices, which exhibited very small parasitics. The exception is FL179E, which was accidentally grown with a 5-nm thick etch-stop layer, and was not able to be processed into a metal-metal waveguide.

## FL179E

The conduction band profile for FL179E is shown in Fig. 6-33(a). The injector barrier is 41-Å thick, compared to with 54-Å thick for FL175C. As a result, the injection anticrossing gaps are much larger:  $\Delta_{1'5} = 2.5 \text{ meV}$  and  $\Delta_{2'5} = 2.5 \text{ meV}$ . The parasitic anticrossing gap is  $\Delta_{1'3} = 1.3 \text{ meV}$ . The anticrossing gaps, oscillator strengths, and LO-phonon scattering times are shown in Fig. 6-34.

The  $I$ - $V$  for a  $200 \times 200 \mu\text{m}^2$  SI-surface-plasmon waveguide device is shown in Fig. 6-33(b). The NDR that occurs at 10 V is due to the parasitic current channel, which has a value of  $J_{para} \approx 1400 \text{ A/cm}^2$ . This occurs at a relatively large voltage due to the parasitic resistance associated with the incomplete etching for the lateral contacts. The device did not lase. This large resistance made it impossible to bias to

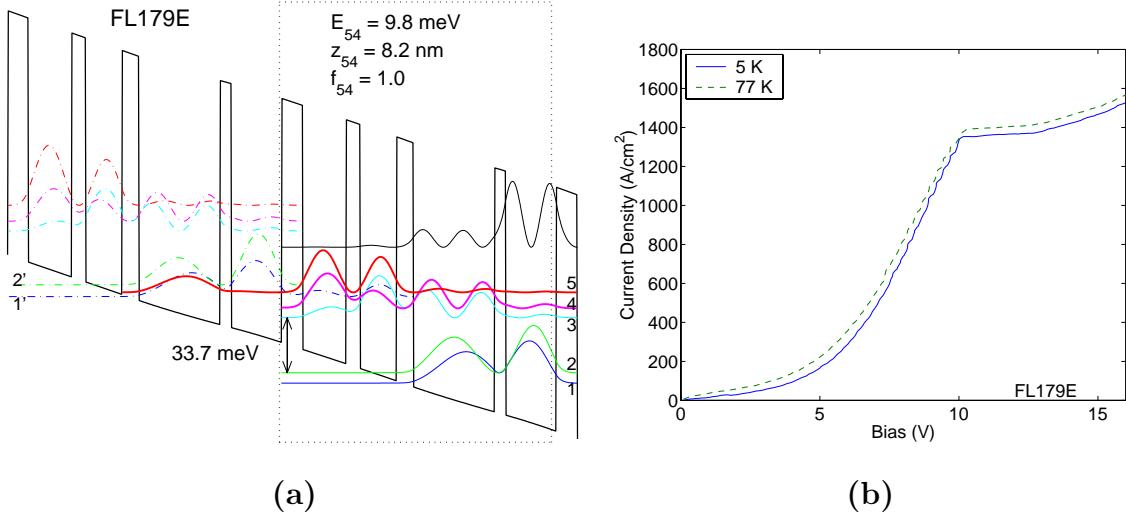


Figure 6-33: (a) Self-consistent conduction band profile of FL179E calculated using 80% conduction band offset. The device is grown in the GaAs/Al<sub>0.15</sub>Ga<sub>0.85</sub>As material system. Beginning with the left injection barrier, the layer thicknesses in Å are 41/85/27/71/33/158/22/98 (as grown 3.4% below design). The 158 Å well is doped at  $n = 1.8 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $2.8 \times 10^{10} \text{ cm}^{-2}$  per module. (b)  $I$ - $V$  taken at 5 K and 77 K of FL175E  $200 \times 200 \mu\text{m}^2$  diode taken with 400-ns pulses at a 1-kHz pulse repetition frequency.

the design point with a long enough duty cycle so that spectra could be taken.

## FL162Gmm

The conduction band profile for FL162G is shown in Fig. 6-35(a). This device has a three-well active region and a two-well injector. The injector barrier is 41-Å thick, which gives injection anticrossing gaps of  $\Delta_{1'5} = 2.3$  meV and  $\Delta_{2'5} = 2.2$  meV. The parasitic anticrossing gap is  $\Delta_{1'3} = 1.0$  meV. The anticrossing gaps, oscillator strengths, and LO-phonon scattering times are shown in Fig. 6-36.

Metal-metal waveguide devices were processed using the In-Au reactive bonding process. Ti/Au (200/4000 Å) Schottky contacts were evaporated on top of the contact layer (thinned to 200 nm), and ridges were wet etched using H<sub>3</sub>PO<sub>4</sub>:H<sub>2</sub>O<sub>2</sub>:H<sub>2</sub>O (10:10:250).  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  characteristics for a 100-μm-wide, 0.86-mm-long device are shown in Fig. 6-35(b). At 8 V a NDR occurs due to the parasitic channel ( $J_{para} \approx 1150 \text{ A/cm}^2$ ), and high-field domains develop beyond this point. No good

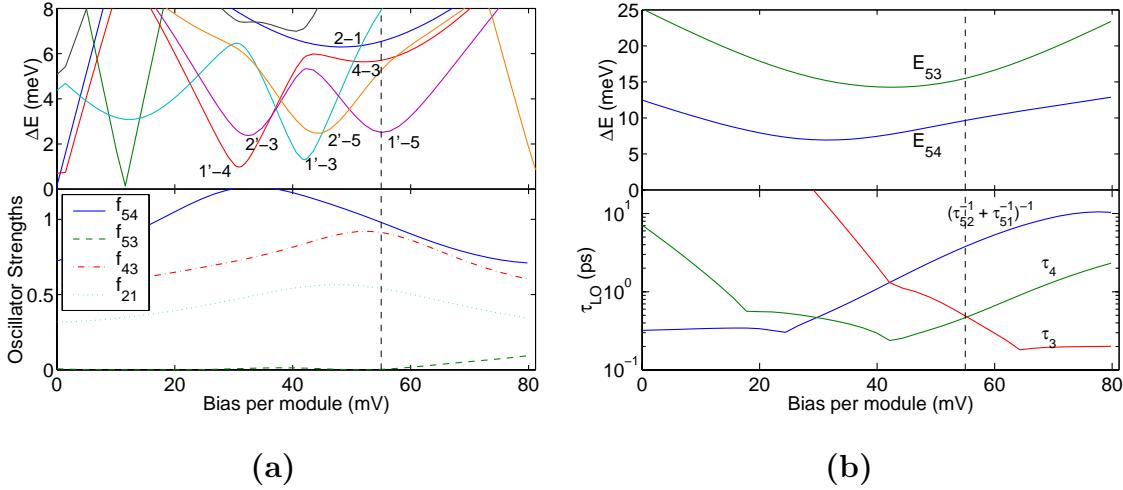


Figure 6-34: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL179E. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction .

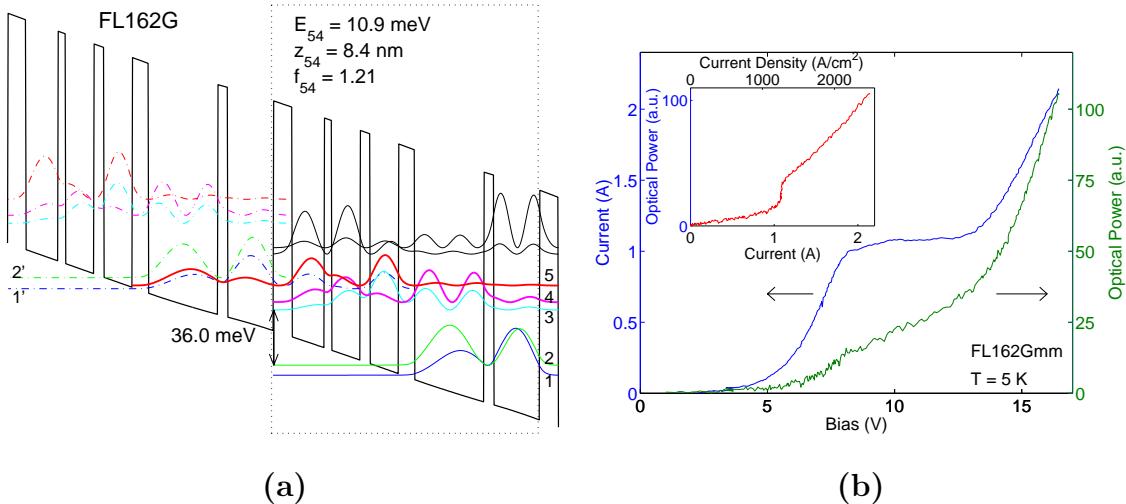


Figure 6-35: (a) Self-consistent conduction band profile of FL162G calculated using 80% conduction band offset. The device is grown in the GaAs/Al<sub>0.15</sub>Ga<sub>0.85</sub>As material system. Beginning with the left injection barrier, the layer thicknesses in Å are 41/71/16/63/22/63/35/152/22/101 (as grown 3.7% below design). The 152 Å well is doped at  $n = 1.9 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $2.9 \times 10^{10} \text{ cm}^{-2}$  per module. (b)  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  taken at 5 K at 5% duty cycle with bolometer with filter 2 from a 100-μm-wide, 0.86-mm-long wet-etched FL162Gmm device.

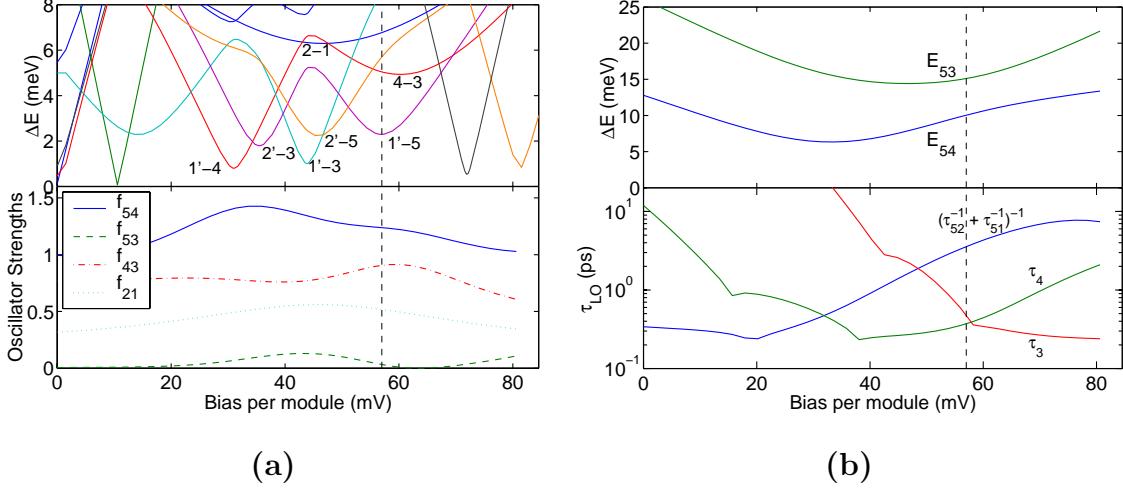


Figure 6-36: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL162G. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction .

spectra were obtained from these devices, even beyond this parasitic point. It is likely that most modules remain severely misaligned once high-field domains develop.

### FL137HMM

The conduction band profile for FL137H is shown in Fig. 6-37. This device has a three-well active region and a three-well injector. X-ray diffraction revealed that this device was grown 5.6% thinner than designed. The injector barrier is 35-Å thick, which gives injection anticrossing gaps of  $\Delta_{1'6} = 2.4$  meV and  $\Delta_{2'6} = 2.3$  meV. The parasitic anticrossing gap is  $\Delta_{1'4} = 0.8$  meV. The anticrossing gaps, oscillator strengths, and LO-phonon scattering times are shown in Fig. 6-38 (calculated for 80% conduction band offset).

Metal-metal waveguide devices were processed using In-Au reactive bonding. The contact layer was thinned to 200 nm, and Ti/Au (200/4000 Å) Schottky contacts were evaporated. Ridges were defined by dry etching using the metallization as a self-aligned etch mask. Three different etchants were used. The first devices were wet etched using  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (10:6:480) (PA solution). This solution performed poorly when etching the thin film from the backside, and widely flared sidewalls were

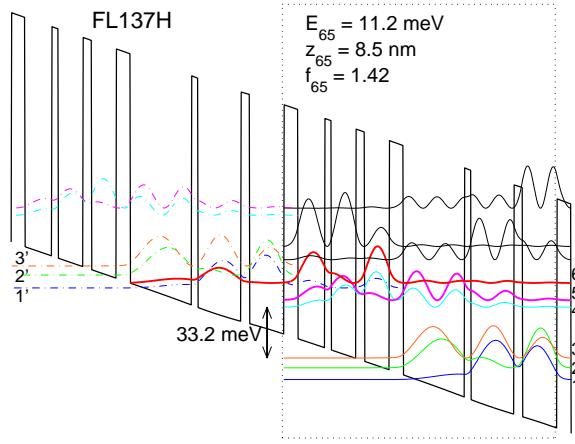


Figure 6-37: Self-consistent conduction band profile of FL137H calculated using 80% conduction band offset. The device is grown in the GaAs/Al<sub>0.15</sub>Ga<sub>0.85</sub>As material system. Beginning with the left injection barrier, the layer thicknesses in Å are 35/67/16/61/21/61/35/152/16/107/21/85 (as grown 5.6% below design). The 107 Å well is doped at  $n = 2.7 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $2.9 \times 10^{10} \text{ cm}^{-2}$  per module.

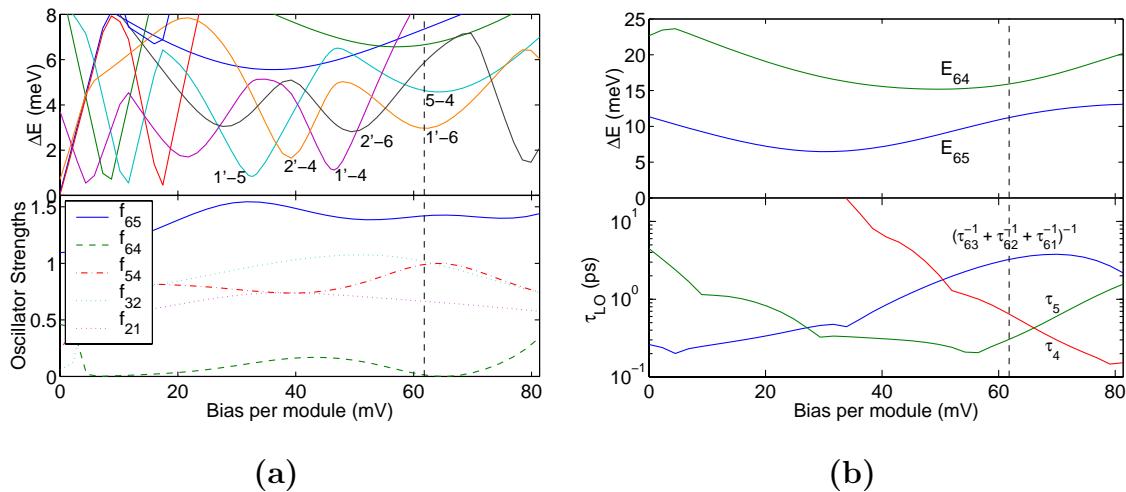


Figure 6-38: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL137H. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction .

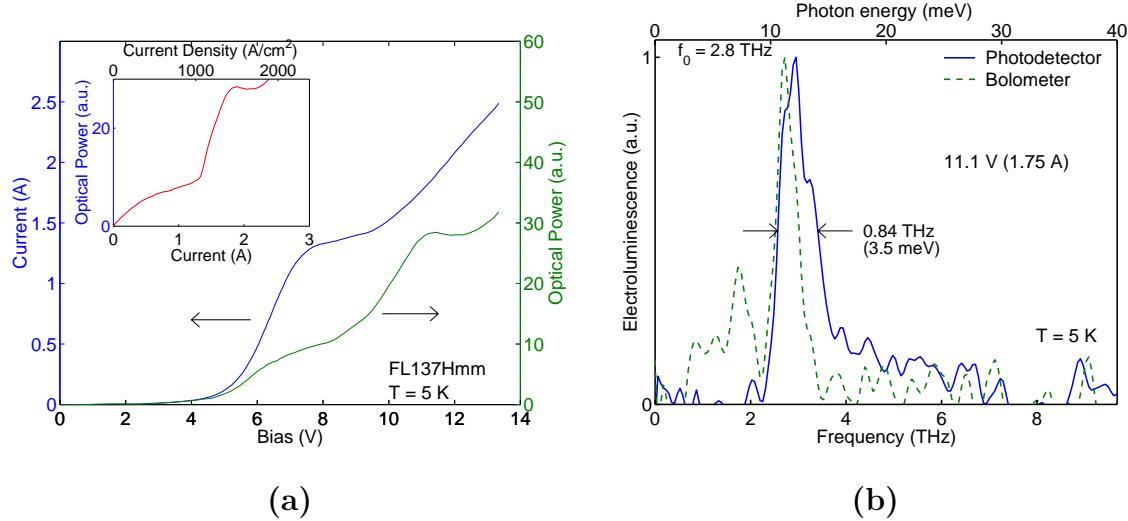


Figure 6-39: (a)  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  taken at 5% duty cycle with photodetector and (b) spectra taken at 5% duty cycle at 5 K from a rear facet HR coated 150- $\mu\text{m}$ -wide, 0.84-mm-long wet-etched FL137HMM device with photodetector and bolometer at 8  $\text{cm}^{-1}$  resolution.

obtained ( $30^\circ$  slope). A second set was wet etched in  $\text{H}_3\text{PO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (10:10:250), which provided better sidewalls with a  $\sim 48^\circ$  slope. Also, a set of devices was dry etched using the top contacts as self aligned etch masks.

None of these devices lased. The most promising characteristics were observed with a device that was wet-etched in the PA solution. An  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  from such a rear facet HR coated 150- $\mu\text{m}$ -wide, 0.84-mm-long FL137HMM device are shown in Fig. 6-39(a). At 7.5 V, an NDR occurs due to the parasitic channel with a current density  $J_{para} = 1030 \text{ A}/\text{cm}^2$ . Although no evidence of reaching design bias was seen in the  $I$ - $V$  the local maxima in the  $L$ - $V$  at 11 V was a clear indication of intersubband emission. Relatively narrow linewidth spectra were observed at this point, and are displayed in Fig. 6-39(b), taken with both the bolometer and the photodetector. The narrowest linewidth measured was approximately 2.1 meV (0.5 THz). The narrow linewidth may indicate that this device was close to lasing threshold. It is not clear why this did not lase, although module misalignment following the development of high-field domains is the most likely explanation.

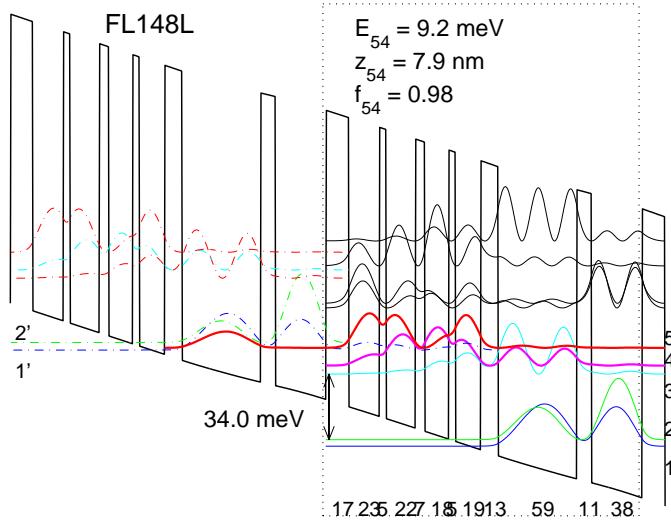


Figure 6-40: Self-consistent conduction band profile of FL148L calculated using 80% conduction band offset (at design thickness). Layer thicknesses are given in monolayers. The 59 ML well is doped at  $n = 1.8 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $3.0 \times 10^{10} \text{ cm}^{-2}$  per module.

### 6.5.5 FL148L

The design of FL148L was an extension of the FL-series resonant phonon design based around a four-well active region and a two-well injector. The conduction band profile is shown in Fig. 6-40. The four-well active region was designed to enhance the oscillator strength for a given transition by adding additional wells, as described in Sec. 7.8. X-ray diffraction analysis reveals that this device was over-grown by only 0.8%, therefore the calculations in this section are calculated for the design thicknesses. The injector barrier is 48-Å (17 ML) thick, which gives injection anticrossing gaps of  $\Delta_{1'5} = 1.25 \text{ meV}$  and  $\Delta_{2'5} = 1.35 \text{ meV}$ . The parasitic anticrossing gap is  $\Delta_{1'3} = 0.35 \text{ meV}$ . The anticrossing gaps, oscillator strengths, and LO-phonon scattering times are shown in Fig. 6-41 (calculated for 80% conduction band offset).

### FL148Lmm

Metal-metal waveguide devices were obtained using the In-Au reactive bonding process. Ti/Au (200/4000 Å) Schottky contacts were evaporated on top of the contact

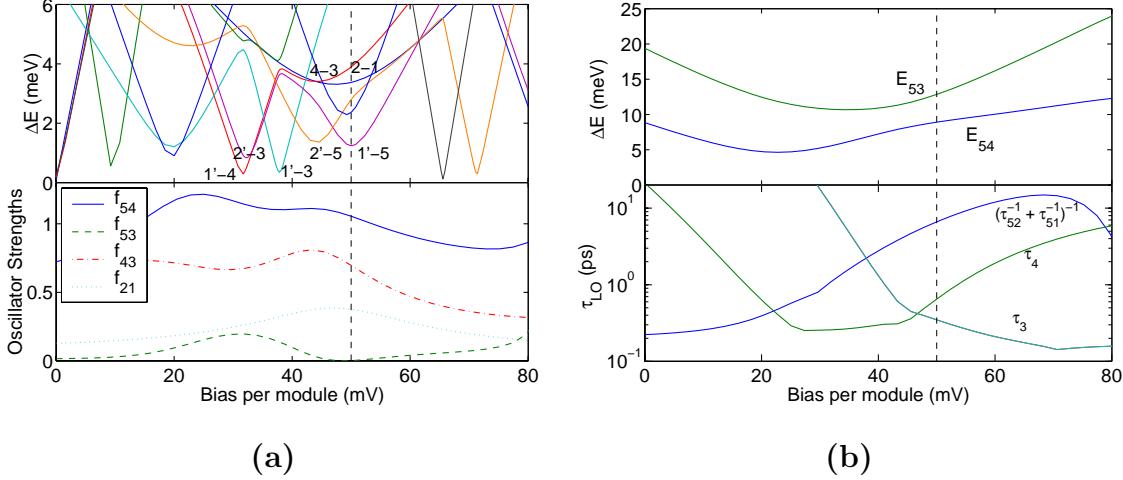


Figure 6-41: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL148L. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as designed .

layer doped at  $2 \times 10^{18} \text{ cm}^{-3}$  (thinned to 200 nm), and were used as self-aligned masks for dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm). The rear facets of a few devices were HR coated with  $\text{Al}_2\text{O}_3/\text{Ti}/\text{Au}$ , although most devices were tested without coatings. The calculated waveguide loss (at 3.4 THz) was  $\alpha_w = 16.0 \text{ cm}^{-1}$ , with a confinement factor of  $\Gamma = 0.98$ .  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  characteristics for a 150- $\mu\text{m}$ -wide, 1.60-mm-long, uncoated ridge structure are shown in Fig. 6-42(a). Design bias occurs at approximately 12 V, with a somewhat low peak current density of  $J_0 = 140 \text{ A/cm}^2$ . The parasitic current channel displayed a magnitude of approximately 75  $\text{A/cm}^2$ .

No lasing was seen, even though uncoated devices as large as 200- $\mu\text{m}$  wide and 3.64-mm long were tested. However, rather narrow electroluminescence was seen (see Fig. 6-42(b)). For example, in a 150- $\mu\text{m}$ -wide, 1.60-mm-long device, an emission peak was observed at 3.4 THz (14 meV) with a linewidth of 0.42 THz (1.7 meV), when biased at 11.9 V (just before NDR). This emission frequency is much larger than the calculated value of  $E_{54} = 9.2 \text{ meV}$ . The reason for this is not clear, but it seems likely that design bias occurred at a much larger bias than intended, by which point the radiative transition had become significantly Stark shifted, which caused an increase in  $E_{54}$ . Evidence of this Stark shift was observed in the lower energy of the peak at a bias of 10.3 V, where the center frequency was 3.1 THz (12.8 meV).

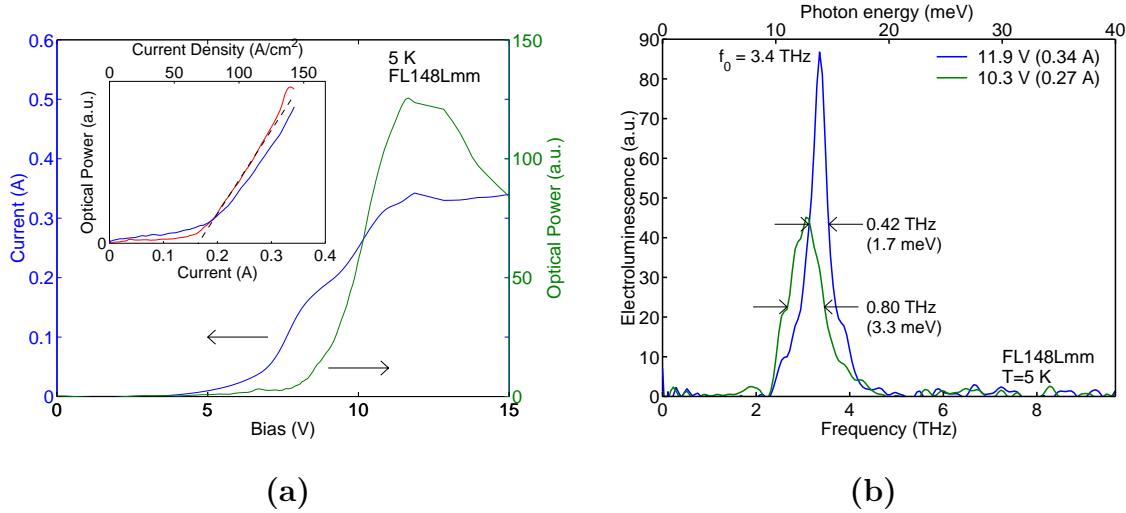


Figure 6-42: (a)  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  taken at 5% duty cycle and (b) spectra taken at 20% duty cycle at 5 K from an uncoated 150- $\mu\text{m}$ -wide, 1.60-mm-long FL148Lmm device. Spectra were taken using photodetector and with  $8 \text{ cm}^{-1}$  resolution.

This explanation is consistent with the fact that the NDR occurs at approximately 12 V, which corresponds to a bias of roughly 78 mV/module after a 0.5 V diode drop is subtracted. This value is much larger than the 50 mV/module design bias, which begs the question of where the extra bias has gone? The observation of a narrow emission linewidth suggests that gain was present and these devices may have been close to lasing threshold, but that the NDR was reached before sufficient current was injected. This is supported by the appearance of some slight superlinearity in the  $L$ - $I$  characteristic taken 5 K as shown in Fig. 6-42(a)). These devices are therefore very promising, and might lase if they are redesigned to increase the current density.

However, it must be mentioned that for an unexplained reason, approximately half of the tested FL148Lmm devices displayed completely different transport characteristics. In these devices measured  $I$ - $V$ s showed no sign of reaching design bias, but rather apparently got stuck at a parasitic current density of  $\sim 400 \text{ A}/\text{cm}^2$  at approximately 6–10 V. No good emission was seen. As of this writing, the schizophrenic behavior of these nominally identical devices from the same wafer remains a mystery.

## 6.6 FFL140

The FFL140 design was an evolution of the FL series of devices, although it was designed before lasing was achieved in FL175C. It is based around the resonant phonon depopulation concept, but instead of a two-well active region, the radiative transition takes place between the middle two states of a four-well miniband. Using this transition has the advantage of a larger oscillator strength (see Sec. 7.8). The larger number of wells has the potential to isolate the phonon well from the injector and reduce the parasitic current channel. The conduction band diagram for this device is shown in Fig. 6-43. The x-ray diffraction data confirmed that the growth period matched the design. The radiative transition is from state  $n = 6$  to  $n = 5$ , which has  $f_{65} = 1.97$  compared to  $f = 0.96$  for FL175C. The lower radiative state  $n = 5$  as well as the lowest miniband state  $n = 3$  are anticrossed with the excited state in the wide phonon well. Ideally, these states form a triplet, each state of which has a short lifetime due to resonant LO-phonon scattering. The challenge with this design is to align the excited phonon state  $n = 4$  with both states at the same time. However, the triplet states cannot be bundled too tightly without bringing the next highest state  $n = 7$  too close in energy to the upper radiative state. Also, the collector barrier cannot be made too thin without reducing the lifetime of  $n = 6$  via LO-phonon scattering to the injector. The anticrossing plot, oscillator strengths, energy separations and LO-phonon scattering rates are all shown in Fig. 6-44. The injector barrier is 45.2-Å-thick (16 ML) compared to 54.1 Å ( $\sim$ 19 ML) for FL175C (as grown). However the injection anticrossing gaps are very similar:  $\Delta_{1'6} = 1.7$  meV for FFL140 and 1.8 meV for FL175C, since the state  $n = 6$  in FFL140 is more delocalized. The same is true for the 10 ML collector barrier. At design bias, the lifetimes of lower triplet state were calculated to be  $\tau_5 = 0.8$  ps,  $\tau_4 = 0.3$  ps, and  $\tau_3 = 1.6$  ps. The upper state parasitic scattering time was calculated to be  $\tau_{6 \rightarrow (2,1)} = 6.5$  ps.

FFL140 was fabricated with a SI-surface-plasmon waveguide, using a 0.6- $\mu\text{m}$ -thick contact layer doped at  $3 \times 10^{18}$  cm $^{-3}$ . The waveguide is calculated to have  $\alpha_w = 5.6$  cm $^{-1}$  and  $\Gamma = 0.31$ , using  $\tau = 0.1$  ps and 0.5 ps for the heavily doped

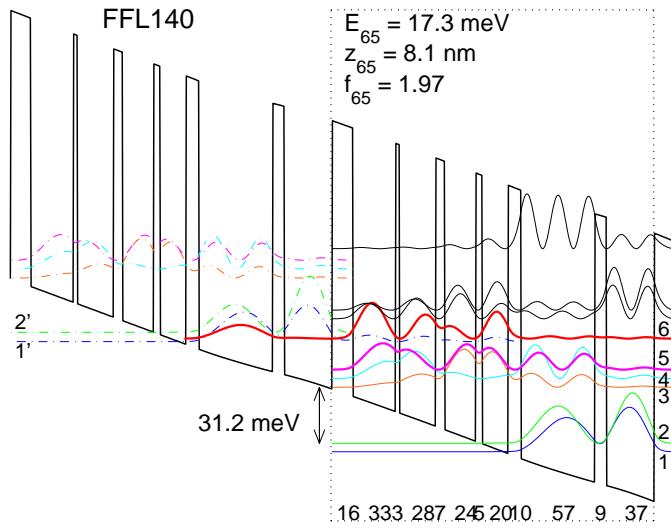


Figure 6-43: Self-consistent conduction band profile of FFL140 calculated using 80% conduction band offset (as grown at design thickness). Layer thicknesses are given in monolayers. The 57 ML well is doped at  $n = 2.4 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $3.9 \times 10^{10} \text{ cm}^{-2}$  per module.

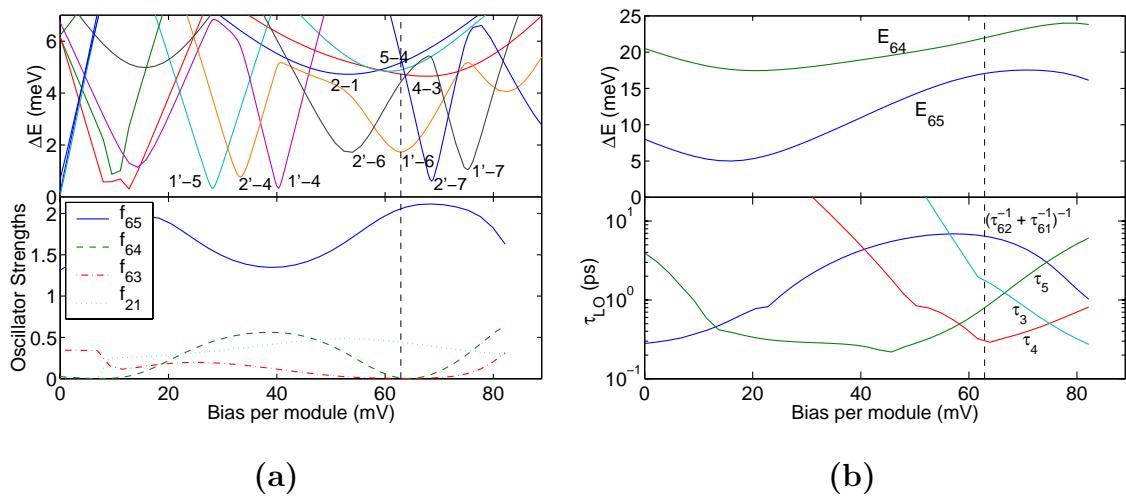


Figure 6-44: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FFL140. All data was calculated using 80% band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction .

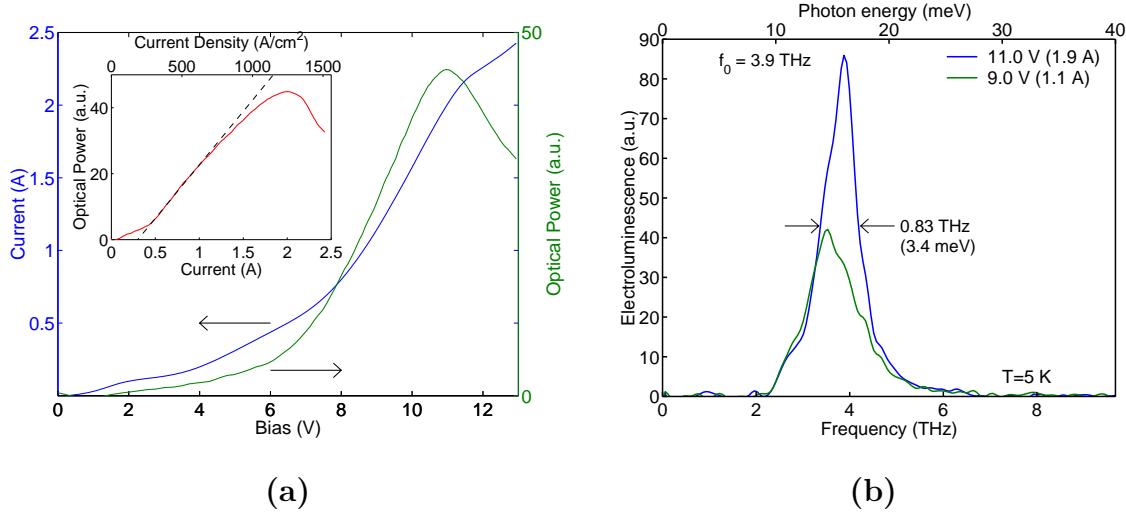


Figure 6-45: (a)  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  taken at 2.5% duty cycle and (b) spectra taken at 5% duty cycle at 5 K from a rear facet HR coated 150- $\mu\text{m}$ -wide, 1.07-mm-long FFL140 device. Spectra were taken using photodetector and with  $8 \text{ cm}^{-1}$  resolution.

and active regions. Wet etching in  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  was used to fabricate ridge structures, and the mask used had a 50- $\mu\text{m}$  separation between the lateral contacts and the ridge. The rear facets were HR coated with evaporated  $\text{Al}_2\text{O}_3/\text{Ti}/\text{Au}$  layers. Additional devices that were fabricated using dry etching were also tested, but a short cleanup wet etch was performed afterward that accidentally undercut the metal contacts by several microns on each side.

Lasing was not obtained in any FFL140 devices despite promising design qualities and experimental results. The  $I$ - $V$ ,  $L$ - $V$ , and  $L$ - $I$  characteristics as well as typical spectra are plotted in Fig. 6-45. Design bias occurred at 11 V; there was not a sharp NDR in the  $I$ - $V$  but the  $L$ - $V$  had a clear peak at that point. This corresponds to  $\sim 78 \text{ mV/module}$ , which is somewhat larger than the design point of  $63 \text{ mV/module}$ . The parasitic current channel from the injector to the excited LO-phonon well state is relatively small,  $\sim 250 \text{ A/cm}^2$  compared to  $J_0 = 1300 \text{ A/cm}^2$  (at the peak of emission power). This is not surprising, since the parasitic anticrossing gap is  $\Delta_{1'4} = 0.3 \text{ meV}$ , which is relatively small. At design bias, a single peak centered at 16 meV (3.9 THz) was observed with a linewidth of 3.4 meV. Although this is relatively narrow, considering the many-well nature of the radiative transition and the thin

barriers (3 ML) in the active region, there was no clear evidence of gain.

The reasons for FFL140's failure to lase was likely twofold. First, there was almost certainly a failure in depopulation. The calculations of LO-phonon scattering times in Fig. 6-44(b) illustrate the difficulty in keeping the lifetimes  $\tau_5$ ,  $\tau_4$ , and  $\tau_3$  sufficiently short over a significant bias range. It is likely that electrons were stuck in one or more of those lower levels. The solution would be to thin the collector barrier, although this would come at the expense of the upper radiative state lifetime. Second, the state  $n = 7$  is only 11 meV above the upper radiative state. It is likely that this state opened up a parasitic current channel which reduced the injection efficiency near design bias. Electrons injected into  $n = 7$  would quickly scatter down to the lower radiative state and not contribute to gain. Examination of the anticrossing plot in Fig. 6-44(a) reveals the close proximity in bias of the  $2'-7$  and  $1'-7$  anticrossings to the design bias. This is supported experimentally by the sublinear  $L-I$  and the lack of a sharp NDR in the  $I-V$ . Unfortunately, to raise the subband separation  $E_{76}$  would also separate the two lower miniband states, which would in turn slow depopulation. However, this problem might be alleviated by reducing the  $\Delta_{21}$  injector anticrossing to improve injection efficiency.

Nonetheless, the FFL140 device provided confirmation of a relatively narrow emission linewidth with strong electroluminescence from a four-well miniband active region with a large number of interfaces. Also, this device displayed a relatively small parasitic current injection channel, which was suppressed by the large spatial and energy separation between the injector states and the wide well excited state. These characteristics point the way to improvements in active region design for future laser devices.



# Chapter 7

## Terahertz quantum cascade lasers

### 7.1 Introduction

We obtained lasing in a terahertz quantum cascade device for the first time on November 18, 2002. The design, known as FL175C, was an evolution of the FL125 and FL175B designs discussed in Sec. 6.5. Unlike the first terahertz QCLs, which were based on chirped superlattice designs, this design was based on the resonant phonon depopulation concept, in which a combination of resonant tunneling and direct LO-phonon scattering is used to quickly remove electrons from the lower state. This method, first proposed by Stroscio *et al.* [170] and Callebaut [171], offers two distinct advantages. First, when a collector state is separated from the lower radiative state by at least  $E_{LO}$ , depopulation can be extremely fast and it does not depend much on temperature or the electron distribution. Second, the large energy separation provides intrinsic protection against thermal backfilling of the lower radiative state ( $E_{LO}/k_B \sim 415$  K). Both properties are important in allowing higher temperature operation of lasers at longer wavelengths. Since then, we have demonstrated lasing in five different designs in total. Four of these are based on the resonant phonon concept, and the fifth, known as CSL104, is a copy of the original chirped superlattice terahertz QCL developed by Köhler *et al.* [64].

## 7.2 FL175C

FL175C was designed in the GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As material system with  $x = 0.15$ , as compared to  $x = 0.3$  for FL175B. This reduction in barrier height was intended to reduce the effect of barrier width fluctuations on the spontaneous emission linewidth and other transport parameters. Since thicker barriers are needed to compensate for the lower barrier height, monolayer growth fluctuations and drift should have proportionally less effect. Also, several barriers were adjusted to give larger anticrossings to enhance coupling between various states. The anticrossing gaps were further increased by the fact that this device was under-grown by 4.2%, as revealed by x-ray rocking diffraction analysis. All parameters in this section have been calculated using the thinner layers unless explicitly stated.

The conduction band diagram of FL175C is shown in Fig. 7-1(a). The four quantum well module is based around a vertical transition between levels  $n=5$  and  $n=4$  (highlighted with thicker lines), with a oscillator strength of  $f_{54} = 0.96$ . At design bias ( $\sim 64$  mV/module), levels  $n=4$  and  $n=3$  are brought into resonance with an anticrossing gap of 5.3 meV, enabling fast depopulation via LO-phonon scattering into the levels  $n=2$  and  $n=1$  ( $\tau_{4 \rightarrow (2,1)} = 0.55$  ps). This allows the barrier to be kept relatively thick (38 Å), which limits parasitic scattering out of the upper state ( $\tau_{5 \rightarrow (2,1)} = 7.1$  ps). Assuming a fully coherent tunneling process between levels  $n=3$  and 4 [168], additional depopulation of  $n=4$  can take place via electron-electron scattering into  $n=3$ , which also has a short lifetime of  $\tau_{3 \rightarrow (2,1)} = 0.46$  ps. Electrons collect in the anticrossed 2-1 doublet ( $E_{21} = 6.6$  meV) where they are injected into the  $n=5$  state in the next module. While LO-phonon emission between radiative states is suppressed at low electron temperatures, the raw scattering times for electrons that are thermally excited sufficiently high in the subband to emit LO-phonons electrons are  $\tau_{54}^{(hot)} \approx \tau_{53}^{(hot)} \approx 0.5$  ps. As described in Eq. (2.49), this leads to a total LO-phonon scattering rate of  $\tau_{54}^{-1} = \tau_{54}^{(hot)-1} \exp((E_{54} - E_{LO})/k_B T_e)$ , and similarly for  $\tau_{53}$ . Compared to FL175B, the per module doping was reduced from  $4 \times 10^{10}$  cm<sup>-2</sup> to  $2.8 \times 10^{10}$  cm<sup>-2</sup>. The calculated anticrossings, oscillator strengths, energy separations,

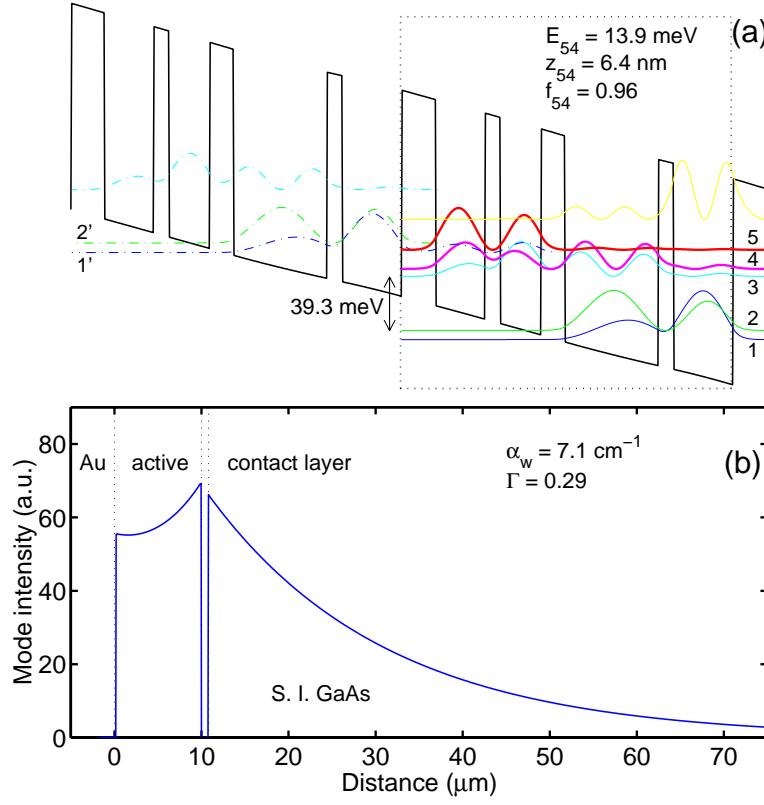


Figure 7-1: (a) Self consistent conduction band profile of FL175C (calculated using 80% band offset) biased at 64 mV/module, which corresponds to a field of 12.2 kV/cm. The device is grown in the GaAs/Al<sub>0.15</sub>Ga<sub>0.85</sub>As material system. The four well module is outlined. Beginning with the left injection barrier, the layer thickness in Å are 54/78/24/64/38/148/24/94 (as grown 4.2% below design). The 148 Å well is doped at  $n=1.9 \times 10^{16} \text{ cm}^{-3}$ , yielding a sheet density of  $2.8 \times 10^{10} \text{ cm}^{-2}$ . (b) Optical mode profile of SI-surface-plasmon waveguide.

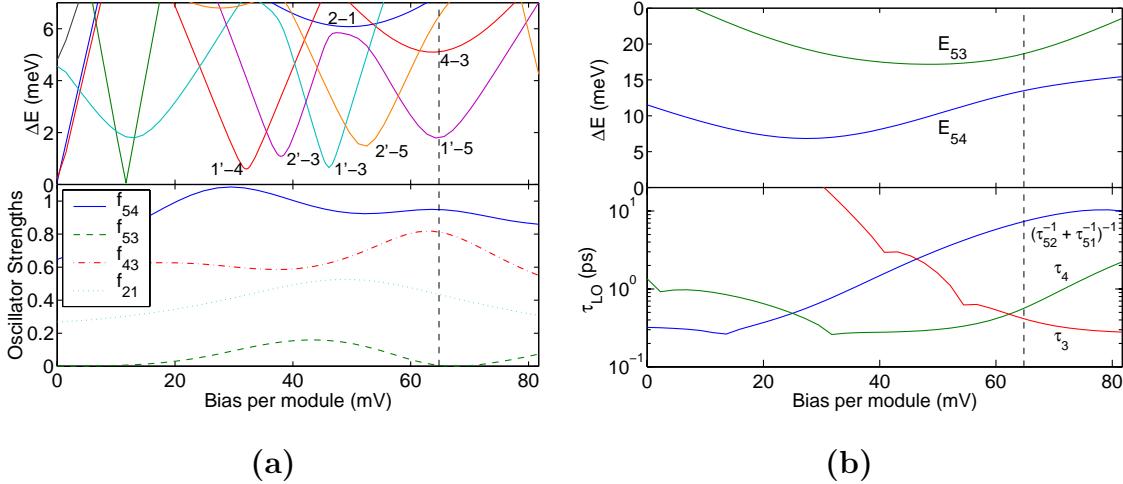


Figure 7-2: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL175C. All data was calculated using an 80% conduction band offset and layer thicknesses as measured by x-ray diffraction.

and LO-phonon scattering times are shown in Fig. 7-2.

Molecular beam epitaxy was used to grow 175 cascaded modules, sandwiched between a 60-nm upper contact layer (doped to  $n=5 \times 10^{18} \text{ cm}^{-3}$ ) and a bottom 800-nm contact layer ( $n=3 \times 10^{18} \text{ cm}^{-3}$ ) on a semi-insulating GaAs substrate. Special care was taken to limit the drift in the growth rate to less than 0.7% over the entire growth, to ensure the accuracy of the doping concentration in the active region within 10%, and the Al fraction within 14.5%–15.5%. A top non-alloyed ohmic contact was made by evaporating Ti/Au layers on low temperature grown GaAs. Wet etching in  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (20:12:480) was used to define ridge structures of various widths, and Ni/Ge/Au alloyed contacts were made to the lower contact layer; the metallization ended 50  $\mu\text{m}$  from the bottom of the ridge. The mode profile of the SI-surface-plasmon waveguide is shown in Fig. 7-1(b). An scanning electron micrograph of a typical device (not FL175C) is shown in Sec. 5.2.1 in Fig. 5-7.

## Subthreshold

The first devices tested did not lase. Typical  $I$ - $V$ ,  $L$ - $V$ ,  $L$ - $I$ , and electroluminescence spectra from non-lasing devices are shown in Fig. 7-3. The  $I$ - $V$  characteristic is

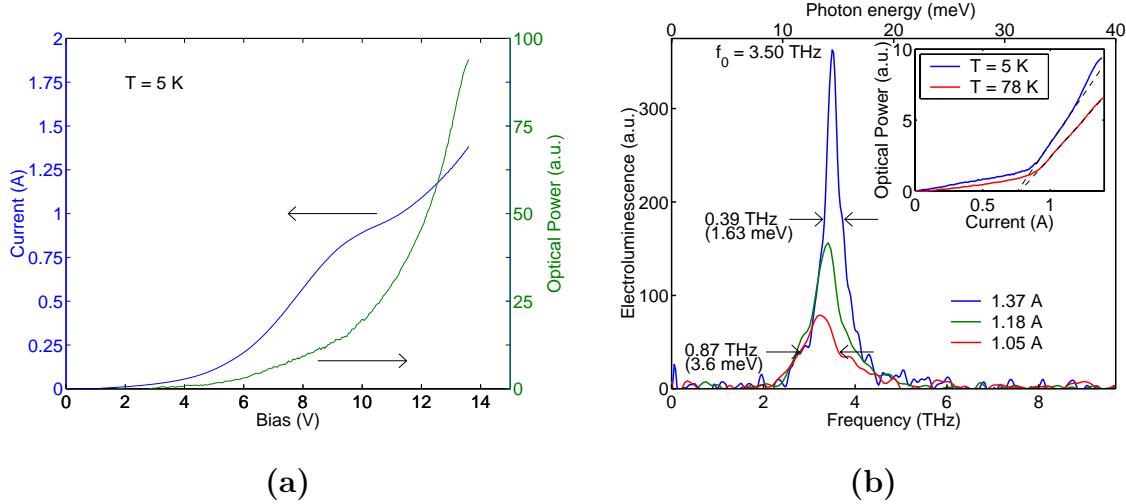


Figure 7-3: (a)  $I$ - $V$  and  $L$ - $V$  and (b) subthreshold spectra taken at 5 K and  $L$ - $I$  curves from a uncoated 150- $\mu\text{m}$ -wide, 0.95-mm-long FL175C device at 5% duty cycle. Data ends immediately before onset of negative differential resistance (NDR). Spectra were taken using photodetector and with  $8 \text{ cm}^{-1}$  resolution.

similar to that seen with FL125 and FL175B, except the peak current density at 5 K is much higher ( $\sim 1000 \text{ A/cm}^2$ ), apparently as a result of the large injection anticrossing ( $\Delta_{1'5} = 1.8 \text{ meV}$ ). The bump in the  $I$ - $V$  at  $\sim 9 \text{ V}$  is more prominent as well. This feature corresponds to a parasitic current path from the injector to the excited state in the wide well ( $1' \rightarrow 3$ ). Only beyond this bias point is substantial current injected into  $n = 5$ ; this is indicated by the sudden increase in slope efficiency in the subthreshold  $L$ - $I$ .

Upon testing a 150- $\mu\text{m}$ -wide, 0.95-mm-long device, indications emerged that this device was close to threshold. The  $L$ - $I$  characteristic reveals a clear superlinear behavior at high injection currents, and the spectra concurrently display a dramatic narrowing, from 3.6 meV to 1.6 meV. Taken together, these indicated the presence of gain in the structure, which had never before been conclusively observed. It appeared that gain was present, but was insufficient to overcome the total losses.

### Coated laser device

To decrease the facet loss, the back facet was high reflectivity (HR) coated by evaporating Ti/Au over silicon nitride, and a 150- $\mu\text{m}$ -wide, 1.19-mm-long ridge was tested

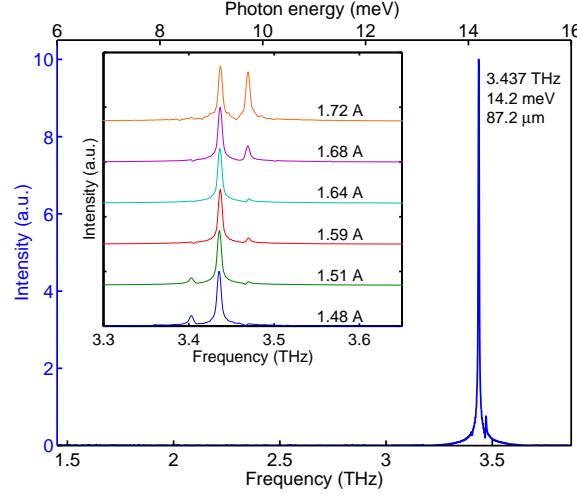


Figure 7-4: Emission spectrum above threshold for 150- $\mu\text{m}$ -wide, 1.19-mm-long FL175C device with rear HR coating. Device is biased at 1.64 A with 100-ns pulses repeated at 10 kHz, at  $T = 5$  K. Spectra were measured with the Ge:Ga photodetector. The inset shows an expanded view of spectra at various bias points, offset for clarity. The spectra were collected in linear-scan mode, and the measured linewidth is limited by the instrumental resolution of  $0.125 \text{ cm}^{-1}$  (3.75 GHz).

(see Ref. [66]). Lasing at 3.437 THz ( $\lambda = 87.2 \mu\text{m}$ ) was immediately obtained in this device at a threshold current density of  $840 \text{ A/cm}^2$  at 5 K. Current pulses of 100-ns duration were applied with a pulse repetition frequency of 2 kHz to minimize device heating. Typical emission spectra above threshold are shown in Fig. 7-4. The emission frequency corresponds to an energy of 14.2 meV, close to the calculated value of 13.9 meV. For much of the bias range, the emission is dominated by a single mode, although the spectrum shifts toward higher energy modes with increasing bias (inset, Fig. 7-4). This is due to the Stark shift of the intersubband gain curve versus applied voltage, which is expected in structures based on a non-intrawell transition (see  $E_{54}$  in Fig. 7-2(b)). The longitudinal modes are spaced by  $1.1 \text{ cm}^{-1}$ , which indicates an effective modal index of  $n_{eff} = 3.81$ .

Measured optical power versus current ( $L-I$ ) curves at low duty cycle are plotted in Fig. 7-5(a). Lasing is observed in pulsed mode up to 64 K with a peak power level of  $100 \mu\text{W}$ , compared to the 10 mW observed at 5 K. The inset shows the increase in threshold current density ( $J_{th}$ ) with temperature. This is mostly caused by the re-

duction of the upper state lifetime due to thermally activated LO-phonon scattering, rather than thermal backfilling of the lower state, as in the case of chirped superlattice THz lasers [65]. The LO-phonon based depopulation mechanism is relatively temperature insensitive, and thermal backfilling is expected to be minimal due to the large energy separation between  $n=4$  and the collector states 2 and 1, where most electrons reside. Often, a phenomenological function  $J_{th} = J_0 + J_1 \exp(T/T_0)$  is used to characterize the temperature dependence in terms of a figure of merit  $T_0$ . A rough fit is shown in the inset for  $J_0 = 840$  A/cm<sup>2</sup>,  $J_1 = 4.25$  A/cm<sup>2</sup>, and  $T_0 = 20$  K.

Fig. 7-5(b) displays the voltage versus current, as well as several  $L$ - $I$  curves taken for pulses of increasing width repeated at 1 kHz. The peak power remains largely the same for duty cycles below 1%, but for longer pulses, the power drops due to heating. Although 2 mW peak power (1 mW average power) was observed at 50% duty cycle, the device burned out in continuous wave (cw) mode when the silicon nitride layer on the coated facet failed due to thermal stress at the large DC power dissipation level ( $\sim 20$  W). Other devices tested also failed to lase cw, and the highest achieved duty cycle was 60%. In this device, at low duty cycle, the lasing threshold occurs at a bias of 12 V, and NDR is reached slightly above 13 V, which correspond roughly 69 mV/module and 74 mV/module respectively. This is consistent with, although somewhat higher than the design value of 64 mV/module. The difference is likely due to parasitic resistance associated with the lateral contacts.

The lasing threshold is reached when the modal gain  $\Gamma g$  becomes equal to the sum of the waveguide losses  $\alpha_w$  and the mirror losses  $\alpha_m$ . Based on the Drude conductivity model ( $\tau = 0.5$  ps in the active region and 0.1 ps in the contact layers),  $\alpha_w$  is calculated to be 7.1 cm<sup>-1</sup>, with a confinement factor of  $\Gamma = 0.29$ . For this rear facet coated structure,  $\alpha_m = 4.9$  cm<sup>-1</sup>, which requires a threshold bulk gain  $g = 40.9$  cm<sup>-1</sup>. Using the expression for peak gain (Eq. (2.37)), we find this requires a population inversion of  $\Delta N = 3.2 \times 10^9$  cm<sup>-2</sup>, which is approximately 10% of the total sheet density. The true spontaneous emission linewidth was taken to be  $\Delta\nu = 1.0$  THz ( $\sim 4$  meV). This was measured from a device with a short cavity length (180  $\mu$ m) to prevent significant spectral narrowing due to gain. A spectrum with a Lorentzian fit is shown in Fig. 7-6.

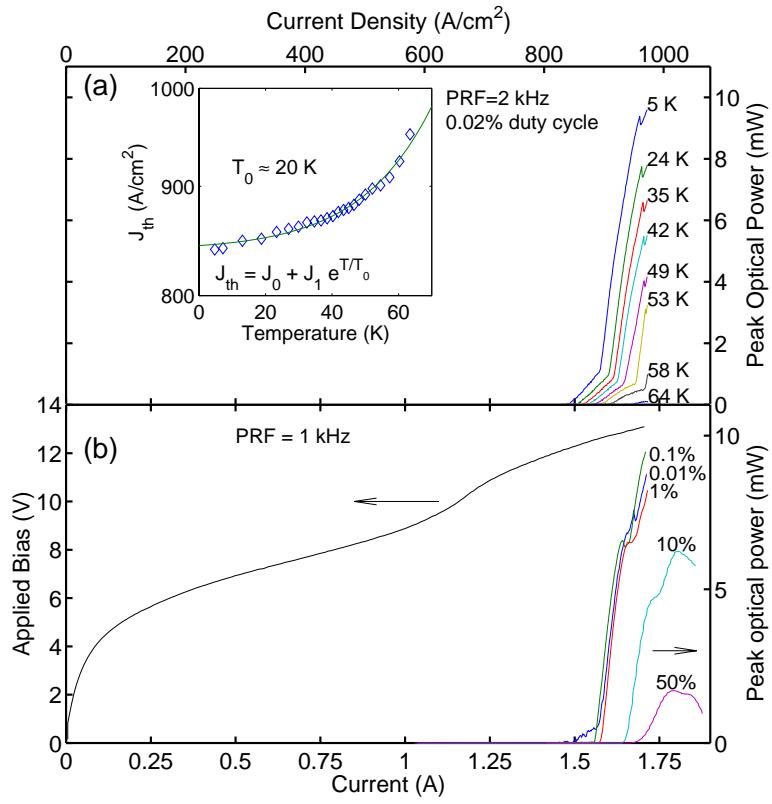


Figure 7-5: (a) Collected light versus current at various temperatures for 150- $\mu\text{m}$ -wide, 1.19-mm-long FL175C device. Data is taken using 100-ns pulses repeated at 2 kHz with the Ge:Ga photodetector. The inset is a semi-log plot of the threshold current density  $J_{th}$  versus temperature. (b) Applied bias and peak optical power versus current, collected at various duty cycles with a pulse repetition frequency of 1 kHz. The data was taken with heat sink temperatures of 5 K, except at 10% and 50% duty cycle, which had  $T = 8 \text{ K}$  and 15 K, due to the large power dissipation. The pyroelectric detector was used and the collected power level was calibrated using the thermopile.

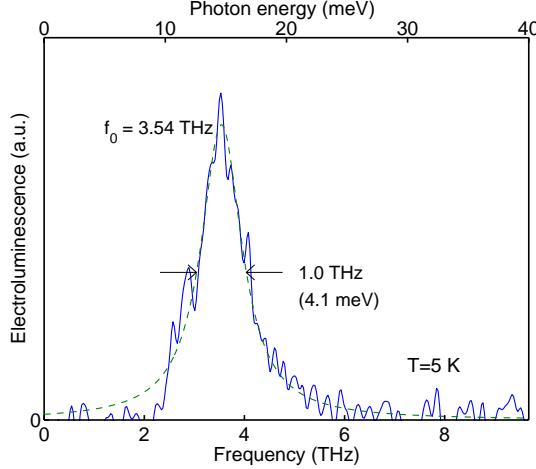


Figure 7-6: Edge emission electroluminescence spectrum taken from lateral edge of longitudinally cleaved  $\sim 180\text{-}\mu\text{m}$ -long, 1.03-mm-wide FL175C device at  $\sim 860 \text{ A}/\text{cm}^2$ . Lorentzian fit to spectrum is shown.

Earlier tests on a 0.95-mm-long uncoated ridge structure ( $\alpha_m = 12.0 \text{ cm}^{-1}$ ) failed to produce lasing, so we may assume that the necessary bulk gain  $g = 65.3 \text{ cm}^{-1}$  was unattainable.

### Long device

The highest temperature and highest power performance was observed from a 200- $\mu\text{m}$ -wide, 2.56-mm-long FL175C ridge with rear HR coating (see Ref. [67]). As shown in Fig. 7-7, lasing in pulsed mode was observed up to 87 K, with  $\sim 14$  mW of peak power collected at 5 K, and  $\sim 4$  mW at 77 K. The power was collected by a cone onto a 2-mm diameter pyroelectric detector, which was then scaled up by a factor of four to account for calibration by a broad area (25 mm diameter) thermopile detector. The actual emitted power may even be somewhat higher, as the collection efficiency is less than unity. Threshold currents were  $J_{th} = 806 \text{ A}/\text{cm}^2$  and  $J_{th} = 904 \text{ A}/\text{cm}^2$  at 5 K and 77 K respectively. The displayed spectrum was taken with the device cooled by liquid nitrogen. At 78 K, the emission is largely single mode for lower injection currents in the range where the slope of the  $L$ - $I$  relation is positive ( $I \leq 4.8 \text{ A}$ ). Beyond this range, the emission power decreases with an increasing current, as a result of the injector becoming misaligned with the upper state. Consequently the emission

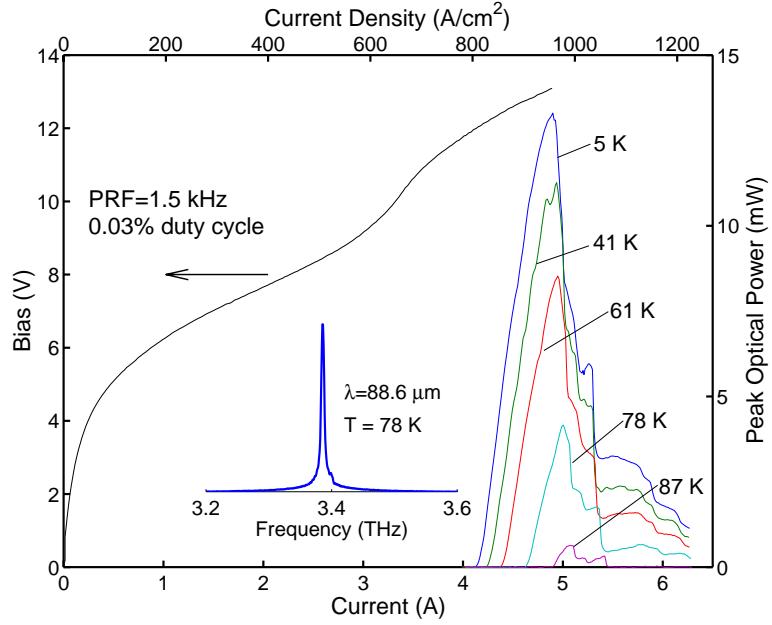


Figure 7-7:  $L$ - $I$  taken with 200-ns pulses repeated at 1.5 kHz from 200- $\mu\text{m}$ -wide, 2.56-mm-long FL175C ridge with rear HR coating. The  $V$ - $I$  curve is from a different device with a smaller area, and only the current density is directly comparable. The 78 K spectrum was taken with 200-ns pulses repeated at 10 kHz.

spectra become increasingly multi-mode and shift to higher frequencies at higher currents. The measured mode spacing is approximately  $0.51 \text{ cm}^{-1}$  at 5 K, which corresponds to an effective mode index of  $n_{eff} = 3.8 \pm 0.1$ . The individual modes are continuously red-shifted by roughly  $0.16 \text{ cm}^{-1}$  (4.8 GHz) as the temperature is increased from 5 K to 78 K. A fit of  $J_{th}$  versus temperature (not shown) to the function  $J_{th} = J_0 + J_1 \exp(T/T_0)$  gave the values  $J_0 = 802 \text{ A/cm}^2$ ,  $J_1 = 5.2 \text{ A/cm}^2$ , and  $T_0 = 25.5 \text{ K}$ .

The improvement in temperature performance compared to the smaller device is due to both the increase in cavity length and width. Increasing the cavity length from 1.19 mm to 2.56 mm reduced the mirror losses from  $\alpha_m = 4.9 \text{ cm}^{-1}$  to  $\alpha_m = 2.2 \text{ cm}^{-1}$ . Given the threshold condition Eq. (4.1), and using the calculated values of  $\Gamma = 0.29$  and  $\alpha_w = 7.1 \text{ cm}^{-1}$ , this increase in cavity length therefore reduces the threshold material gain  $g_{th}$  from  $41 \text{ cm}^{-1}$  to  $32 \text{ cm}^{-1}$ . The wider ridge also improves the device performance, primarily by providing a more uniform current distribution. The wet

etch processing produces sloped ridge sidewalls, and hence different modules in the active region are effectively biased at different points, which can reduce the peak available gain. This effect is exacerbated by the presence of the parasitic current channel, so that current does not contribute to gain until beyond the bias point of  $\sim 10$  V ( $650$  A/cm $^2$ ) (see Fig. 7-7). As a result, the gain per module is a strong function of current, which results in an increased sensitivity of the total gain to current non-uniformity. The wider ridge may also slightly reduce the waveguide loss by reducing the lateral mode overlap with the alloyed ohmic contacts [138].

## Slope Efficiency

The expression for slope efficiency in a QCL is given by

$$\frac{dL}{dI} = \frac{\hbar\omega}{e} N_{mod} \frac{\alpha_m}{\alpha_m + \alpha_w} \eta_i, \quad (7.1)$$

where  $N_{mod}$  is the number of cascade modules. The internal quantum efficiency  $\eta_i$  can be reduced both by a non-zero population in the lower radiative state, and non-unity injection efficiency, i.e. the partial injection of carriers into the lower radiative state. The complete derivation of the term  $\eta_i$  is given in Appendix C (see Eq. (C.19)). A key feature of a cascade laser is the capability for overall quantum efficiencies greater than unity, i.e. an injected electron can produce more than one photon. The maximum measured slope efficiency is approximately 96 mW/A, measured from the 150- $\mu$ m-wide, 1.19-mm-long device (Fig. 7-5). This is substantially smaller than the ideal value of 1.5 W/A for  $\eta_i = 1$ . Aside from imperfect optical collection efficiency, the discrepancy arises since  $\eta_i$  in our structure is certainly less than unity. For example, the feature at  $\sim 10$  V in the  $V$ - $I$  curve seen in Fig. 7-5(b) corresponds to the injection of current from  $n=1'$  into  $n=4$ . The proximity of this feature to the threshold bias of 12 V indicates that some current is likely still being injected into the lower state, reducing efficiency and population inversion. Indeed our measured slope efficiency should be interpreted with care, since the value of  $\eta_i$  itself increases as the bias voltage approaches the NDR point. The slope efficiency for the longer (200- $\mu$ m-wide 2.56-

mm-long) device is substantially less (24 mW/A). This is partly due to the reduction of the factor  $\frac{\alpha_m}{\alpha_m + \alpha_w}$  from 0.40 to 0.24 and but also reflects a different optical collection efficiency.

The shape of the  $L$ - $I$  curves for the two devices are also somewhat different at low duty cycle. A two-slope behavior is seen in the shorter device, but not the longer one. The double slope feature disappears for longer pulse widths, since lasing in the region of lower slope efficiency occurs only at the beginning of the pulse, before the active region temperature rises significantly. This is likely due to the alignment of the upper injector state  $n=2'$  with  $n=5$ , while most of the electrons remain in  $n=1'$ , allowing them to be parasitically injected into  $n=4$ . At higher biases,  $n=1'$  is aligned with  $n=5$ , which results in a higher slope efficiency. Another possibility is that there is some bias non-uniformity in the structure, and not all of the modules are turning on simultaneously, causing regions of different slope.

## Thermal Analysis and Power Dissipation

Despite attaining lasing up to 87 K in FL175C, operation was only achieved up to 60% duty cycle. This is a result of the relatively high power dissipation at dc bias. For example, for the 150- $\mu$ m-wide 1.19-mm-long device described above, operation just before NDR consumes approximately 23 W. The first chirped superlattice structures that obtained cw operation had roughly a factor of 10 less power dissipation. There are two reasons for the difference in our resonant phonon structure. First, the voltage drop per module for a resonant phonon design will always be larger than a chirped superlattice design, due to the explicit inclusion of an LO-phonon emission channel. However, this is not necessarily a disadvantage, since the large energy separation  $E_{LO}$  can suppress thermal backfilling and improve high temperature performance. Second, the threshold current density  $J_{th}$  in this device is much higher (800 A/cm<sup>2</sup>) than is typical for chirped superlattice designs where  $J_{th} \sim 100\text{--}300$  A/cm<sup>2</sup> (pulsed at 5 K). Hence, reducing  $J_{th}$  by suppressing the parasitic current channel is a priority for improving the performance of these devices.

## Coherence and resonant phonon depopulation

The lifetimes of levels  $n = 4$  and  $n = 3$  were calculated to be approximately  $\tau_4 \approx \tau_5 \approx 0.5$  ps at design bias. These rates were calculated using a semiclassical Fermi's golden rule approach where it was assumed that both wavefunctions were delocalized. This is equivalent to the coherent limit using the tight-binding resonant tunneling formalism from Sec. 2.5.2, where the relaxation time is  $\tau = 0.25$  ps. Using the calculated anticrossing gap of  $\Delta_{43} \simeq 5$  meV, and assuming a pure dephasing time of  $T_2^* = 0.33$  ps, Fig. 2-9(b) gives an effective tunneling time of 0.75 ps. In other words, dephasing has slowed depopulation by 50%. Although the accuracy of this number depends strongly on the value for  $T_2^*$ , which is highly uncertain, this example illustrates the potential for degradation of dephasing. Also, fast dephasing can broaden the current versus detuning lineshape (see Eq. (2.63)), which will increase coupling of  $n = 3$  with the upper radiative state  $n = 5$  and reduce  $\tau_5$ . If dephasing becomes severe, levels 4 and 3 will become localized on either side of the collector barrier, which will reduce the oscillator strength of the radiative transition, which is calculated assuming wavefunction coherence through the collector barrier.

The detrimental effects of dephasing on resonant phonon depopulation will dictate the minimum 4-3 anticrossing, which must be considered when designing lower frequency lasers. In order to selectively depopulate the lower radiative state while preserving a long upper state lifetime, when  $E_{54}$  is reduced the 4-3 anticrossing gap must be reduced accordingly. Since dephasing scattering will increase with temperature, it may eventually limit higher temperature operation. In such a scenario, designs in which depopulation is insensitive to dephasing will prove advantageous.

## 7.3 FL152Fmm

The FL152F design, shown in Fig. 7-8, was one of the successor designs to FL175C, and had a similar structure, save for the addition of an extra well to the injector [70]. It was intended to operate at roughly 2.5 THz, although due to design approximations and a 2.5% MBE undergrowth, the calculated radiative energy separation was

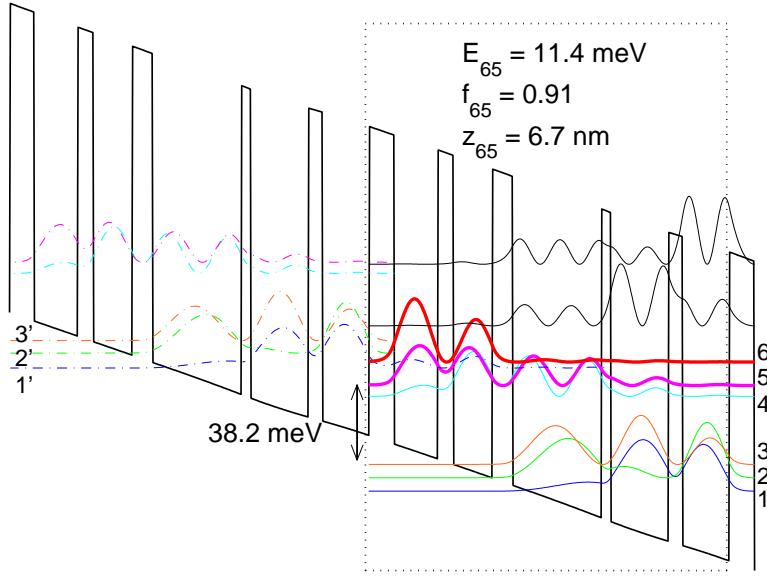


Figure 7-8: Self-consistent conduction band profile for FL152F at a field of 9.5 kV/cm (60.5 mV/module) calculated using 80% conduction band offset. Beginning with the left injection barrier, the layer thicknesses in Å are 44/77/28/69/36/157/17/102/25/83 (as grown 2.5% below design). The 102 Å well is doped at  $n = 2.9 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $3.0 \times 10^{10} \text{ cm}^{-2}$  per module.

$E_{65} = 11.4 \text{ meV}$ . The extra injector well was designed to reduce the parasitic current channel present in FL175C and allow lower threshold densities. The states in the injector would only line up and form a miniband at design bias; below that point, the states would be localized, and the addition of an extra well would reduce the coupling into the next module. As will be shown below, this approach was generally unsuccessful, largely because many of the other barriers were thinned. At the design bias ( $\sim 61 \text{ mV/module}$ ), levels  $n = 5$  and  $n = 4$  are anticrossed ( $E_{54} = 5.2 \text{ meV}$ ) and both levels are quickly depopulated through sub-picosecond LO-phonon scattering ( $\tau_5 \approx \tau_4 \approx 0.4 \text{ ps}$ ) into the injector states ( $n = 3, 2, 1$ ). The small overlap between  $n = 6$  and the injector states limits the calculated parasitic scattering time to  $\tau_{6 \rightarrow (1,2,3)} \approx 6.0 \text{ ps}$ .

The cladding and contact layers were chosen so that the device could be processed using either a SI-surface-plasmon or metal-metal waveguide. FL152F was grown on

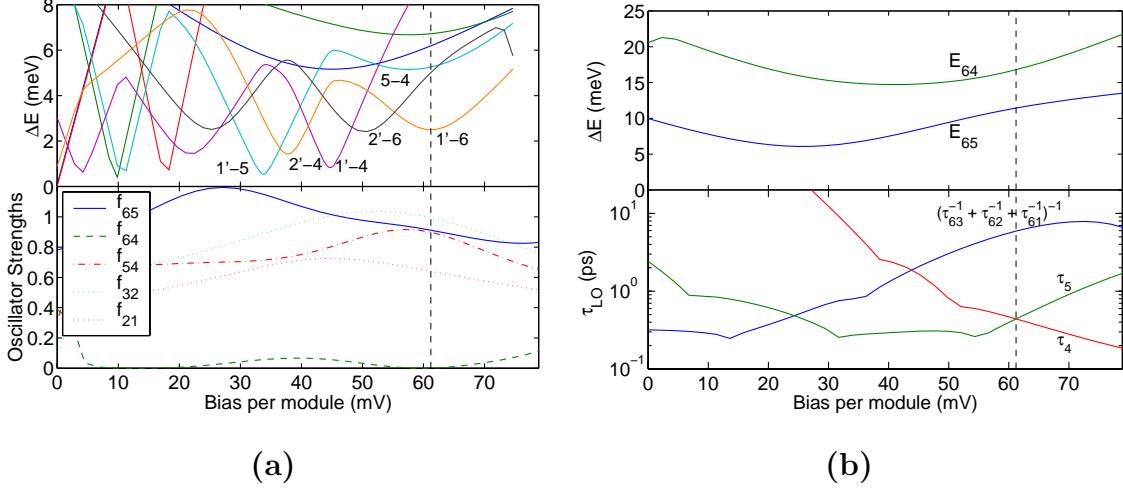


Figure 7-9: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL152F. All data was calculated using 80% band offset and layer thicknesses as measured by x-ray diffraction.

a SI GaAs substrate; first a 0.5- $\mu$ m undoped Al<sub>0.5</sub>Ga<sub>0.5</sub>As etch-stop layer was grown, then a 0.6- $\mu$ m GaAs contact layer doped at  $n = 2 \times 10^{18}$  cm<sup>-3</sup>. Then 152 cascaded modules were grown to form the 10- $\mu$ m-thick active region. A 60-nm GaAs contact layer ( $n = 5 \times 10^{18}$  cm<sup>-3</sup>) was grown above the active region, followed by a thin low-temperature-grown GaAs cap layer to allow for the use of a non-alloyed ohmic contact. Fabrication of the SI-surface-plasmon waveguide structure took place via wet etching in NH<sub>4</sub>OH:H<sub>2</sub>O<sub>2</sub>:H<sub>2</sub>O (20:12:480) as described in Sec. 5.2. For FL152Fmm, the metal-metal waveguide was fabricated using the In-Au reactive wafer bonding process described in Sec. 5.3. Ti/Au (200/4000 Å) Schottky contacts were evaporated and used as self-aligned etch masks for ridge structures, and ECR-RIE using BCl<sub>3</sub>:N<sub>2</sub> (15:5 sccm) was used to dry etch ridges. An additional metal-metal processing run was also performed where the ridges were wet etched using H<sub>3</sub>PO<sub>4</sub>:H<sub>2</sub>O<sub>2</sub>:H<sub>2</sub>O (1:1:25), which resulted in devices with sloped sidewalls. For some devices, Al<sub>2</sub>O<sub>3</sub>/Ti/Au layers (approximately 3000/150/1500 Å) were evaporated on the back facets to form a HR coating. The resulting metal waveguide structures were sufficiently mechanically strong to allow multiple wire bonds to be made directly to the top of the ridge. Scanning electron micrographs of various metal-metal waveguides are displayed in

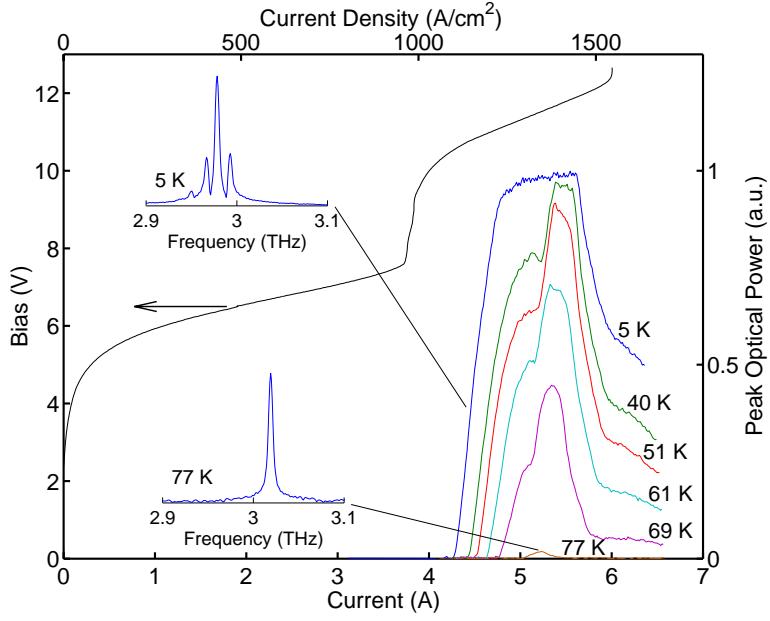


Figure 7-10:  $L$ - $I$  and  $V$ - $I$  taken in pulsed mode for 150- $\mu\text{m}$ -wide, 2.59-mm-long FL152Fmm ridge with metal-metal waveguide and rear HR coating. Data was measured using 100-ns pulses repeated at 1 kHz. Note that the current-voltage characteristic was measured using a similar, but smaller device, so only the current density scale is applicable. Spectra taken at 5 K and 77 K (with liquid nitrogen cooling) using 100-ns pulses repeated at 10 kHz are also shown. Emitted light was measured using a Ge:Ga photodetector, and the spectra were collected in linear-scan mode with a resolution of 0.125 cm<sup>-1</sup>.

Chap. 5 in Sec. 5.3.3.

The SI-surface-plasmon waveguide suffered from a large parasitic resistance that was likely due to a failure to completely etch down to the  $n^+$  contact layer. As a result of the large current density at NDR (1500 A/cm<sup>2</sup>), reaching the design bias was extremely difficult, and required using a high voltage pulser with a 50- $\Omega$  impedance-matching resistor for laser device tests. This precluded measuring standard  $I$ - $V$  and  $L$ - $V$  characteristics, due to the lack of a voltage sense readout and the small duty cycle. However, lasing was not observed in any SI-surface-plasmon devices, with the longest tested being 2.70-mm long and 150- $\mu\text{m}$  wide with a HR coating ( $\alpha_m = 2.1 \text{ cm}^{-1}$ ).

Electrical characteristics of the metal-metal device FL152Fmm were good, with only a small voltage offset ( $\sim 0.6$  V) added to the  $V$ - $I$  characteristic (Fig. 7-10) due to the forward biased Schottky top contact. No large parasitic resistances were observed.

Lasing was observed in several structures, the best being a 150- $\mu\text{m}$ -wide, 2.59-mm-long device with a rear HR coating, which lased in pulsed mode up to a heat sink temperature of 77 K. Light versus current relations and spectra for this device are shown in Fig. 7-10. At 5 K, emission takes place from 2.94 THz ( $\lambda = 102 \mu\text{m}$ ) to 3.06 THz ( $\lambda = 98 \mu\text{m}$ ) as the bias is increased. At 77 K emission is single mode at 3.02 THz, and the spectrum was taken with the laser cooled by liquid nitrogen. The emission corresponds to a transition energy of 12.4 meV, slightly larger than the calculated value of 11.4 meV. The threshold current density at 5 K is  $J_{th} = 1100 \text{ A/cm}^2$ , and the peak power is estimated to be roughly 100  $\mu\text{W}$ . The high value of  $J_{th}$  is a result of a parasitic current channel from  $n = 1', 2'$  into  $n = 4$  that dominates transport for  $J \leq 1000 \text{ A/cm}^2$ , which gives rise to the shoulder-like feature in the  $V$ - $I$  at a bias of  $\sim 7 \text{ V}$ . Not until beyond this bias point does the injector line up with the upper radiative state  $n = 6$  to give gain. In fact, this shoulder is much more prominent than in FL175C, which is likely a result of the larger parasitic anticrossings:  $\Delta_{1'4} = 0.84 \text{ meV}$ ,  $\Delta_{2'4} = 1.44 \text{ meV}$ . In fact, due to the large injection anticrossing ( $\Delta_{1'6} = 2.5 \text{ meV}$ ), the injection efficiency is likely rather poor, and many carriers are injected into the lower radiative state  $n = 5$ , even at design bias. Also, the sharp NDR  $V$ - $I$  at  $V \approx 7 \text{ V}$  suggests the development of high-field domains inside the structure. Even though normal transport is obtained again at a bias of  $\sim 10 \text{ V}$  and a smooth  $V$ - $I$  is regained, it is not clear that all of the modules are correctly aligned. Both poor injection efficiency and module misalignment would cause low slope efficiency, which would explain the low peak output powers.

The maximum duty cycle was limited, as lasing occurred only over the first 2  $\mu\text{s}$  of an applied pulse. This effect is likely related to device heating during operation, especially since the threshold current density is high. Sensitivity due to heating is partly due to the small amount of gain present in this device, due to the poor slope efficiency. However, lasing was observed up to a heat sink temperature of 77 K, which is not that much lower than FL175C where  $T_{max} = 87 \text{ K}$ . This effect was exacerbated by an increased thermal resistance at the bonding interface, which slowed heat removal from the active region. Aside from the short pulse width, this conclusion is supported

Table 7.1: Characteristics of FL152Fmm devices that successfully lased. All listed devices had rear HR coatings. Devices marked with “\*” were cleaved from edge of wafer piece, where the bonding quality was poor.

Device size	Etching	$J_{th}$ (A/cm <sup>2</sup> )	$T_{max}$ (K)	$PW_{max}$ ( $\mu$ s)
150 $\mu$ m $\times$ 1.07 mm	wet	1175	46	2
150 $\mu$ m $\times$ 1.27 mm	wet	1190	37	0.5
100 $\mu$ m $\times$ 1.55 mm	dry	1118	70	
* 200 $\mu$ m $\times$ 2.05 mm	dry	1320	12	0.08
150 $\mu$ m $\times$ 2.59 mm	dry	1100	77	
* 150 $\mu$ m $\times$ 2.60 mm	dry	1230	41	0.15
200 $\mu$ m $\times$ 1.55 mm	dry	1150	64	

by the fact that the maximum lasing pulse widths decrease rapidly and  $J_{th}$  increases rapidly for devices away from the wafer piece center. This is well correlated with the observation that the bonding quality is the best in a  $\sim$ 5-mm diameter region in the center of the piece, and degrades further outside. This is confirmed by SEM micrographs of bonded devices, which reveal the presence of voids in the bonding layer (see Sec. 5.3.3). Such a phenomenon has been reported in Ref. [137] for a mid-infrared QCL in a metal waveguide. Modifications to the processing to improve the bonding interface quality should alleviate this problem.

The characteristics of the FL152Fmm devices that lased are summarized in Table 7.1. Lasing was also obtained in two wet-etched FL152Fmm devices, although in general they had higher  $J_{th}$  and smaller peak output powers than dry-etched devices. As discussed for FL175C, the current spreading associated with the sloped sidewalls of the wet-etched devices reduces the available gain (see Fig. 5-13), and limits the minimum ridge width. Indeed, a dry-etched 100- $\mu$ m-wide ridge lased, whereas the narrowest wet-etched ridge (FL152Fmm or FL175C) that successfully lased was 150- $\mu$ m wide. The effective index was obtained from the mode spacings in the two wet etched devices, giving  $n_{eff} = 3.9 \pm 0.1$ . The dry etched devices tended to have multiple lateral modes present, which made measurement of the mode spacing difficult.

The observation of the first terahertz lasing using a metal waveguide confirms its effectiveness for long wavelength operation. This is corroborated by the fact that

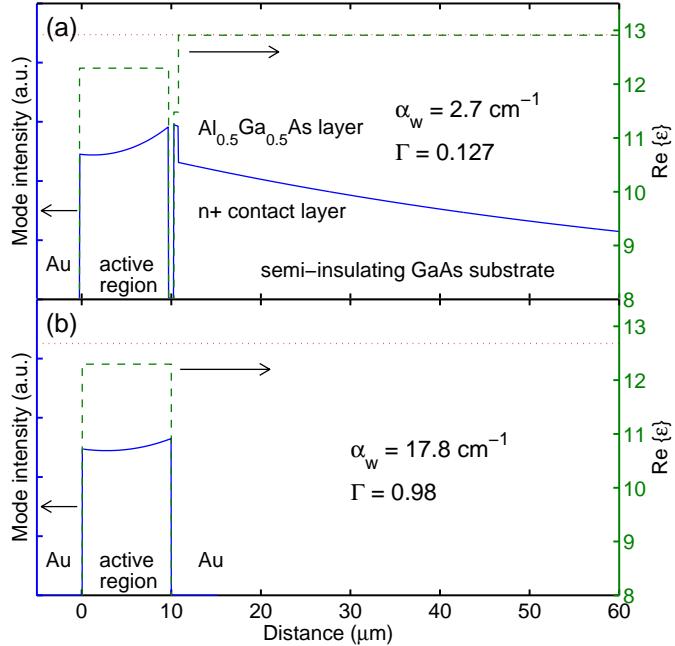


Figure 7-11: Mode intensities (solid lines) and the real part of the dielectric constant  $\epsilon(\omega)$  (dashed lines) for the (a) SI-surface-plasmon waveguide and (b) metal-metal waveguide. The dotted line represents the square of the modal effective index.

devices of similar dimensions (150- $\mu\text{m}$  wide and 2.7-mm long, HR coated) that were fabricated using the SI-surface-plasmon waveguide failed to lase. The one-dimensional mode patterns for the two waveguides were calculated using a Drude model solver and are shown in Fig. 7-11. Drude scattering times of  $\tau = 0.1 \text{ ps}$  and  $0.5 \text{ ps}$  were used for the heavily doped and lightly doped semiconductor regions respectively, and  $\tau = 0.05 \text{ ps}$  was used for gold [131]. The SI-surface-plasmon waveguide was calculated to have a waveguide loss of  $\alpha_w = 2.7 \text{ cm}^{-1}$  and a confinement factor of  $\Gamma = 0.127$ . The  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  layer, which is not normally present in SI-surface-plasmon waveguides, causes a slight reduction in both confinement and loss. The metal waveguide was calculated to have  $\alpha_w = 17.8 \text{ cm}^{-1}$  and  $\Gamma = 0.98$ . Only about  $6 \text{ cm}^{-1}$  of that loss is due to the metal and contact layers and the remainder is due to free carrier absorption in the active region.

Given that the threshold material gain  $g_{th}$  is

$$g_{th} = \frac{\alpha_w}{\Gamma} + \frac{\alpha_m}{\Gamma}, \quad (7.2)$$

and  $\alpha_m$  typically ranges from 2–10 cm<sup>-1</sup>, it is clear that the large confinement factor is a major advantage. While the term  $\alpha_w/\Gamma$  is of similar magnitude for both waveguides, the term  $\alpha_m/\Gamma$  can become quite large for the SI-surface-plasmon waveguide. As discussed in Chap. 4, metal waveguides become increasingly advantageous for longer wavelengths where the confinement factor  $\Gamma$  for the SI-surface-plasmon waveguide decreases even further. The reduction in  $\Gamma$  is due both to a geometric effect as the wavelength grows relative to the active region thickness, and to the  $1 - \omega_p^2/\omega^2$  reduction of the dielectric constant in the active region. As a result of this reduction in  $\Re\{\epsilon(\omega)\}$  by the free carriers, the mode extends further into the substrate where the dielectric constant is higher (Fig. 7-5(a)), or even becomes unbound below some cutoff frequency (2.8 THz at our doping level in the active region). If the mode becomes unbound, it becomes necessary to thin the substrate and use the backside metal contact for confinement. The reduction in  $\Gamma$  may be mitigated by increasing the thickness of the  $n^+$  contact layer, or by decreasing the active region doping, but this comes at the cost of increased loss or reduced design flexibility. On the other hand,  $\Gamma$  for the double-sided metal waveguide is always close to unity regardless of the doping concentration in the active region. This important advantage will be valuable or even crucial in the effort to extend quantum cascade laser operation to even longer wavelengths.

## 7.4 FL148I-M1

The design of FL148I-M1 is a modification of the original laser FL175C; like FL152F, an extra well was added to the injector region to suppress the parasitic current channel. In theory, at design bias the injector states will align to form a conductive miniband that injects carriers into the upper state. Below the design bias, the states

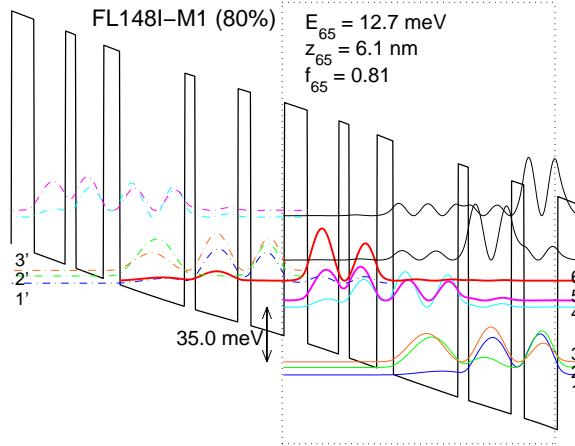


Figure 7-12: Self-consistent conduction band profile for FL148I-M1 calculated using 80% band offset. Beginning with the left injection barrier, the layer thicknesses in Å are 58/78/26/69/40/161/26/106/32/83 (as grown 1.8% above design). The 161/26/106-Å-thick layers are doped at  $n = 1.0 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $2.9 \times 10^{10} \text{ cm}^{-2}$  per module.

will be localized in their respective wells. The addition of extra wells to the injector increases the spatial separation and reduces the parasitic coupling of the lowest, most populated, injector state at the lower biases. The conduction band diagram is shown in Fig. 7-12. X-ray diffraction analysis indicated a 1.8% over-growth for the layers. The calculated anticrossings, oscillator strengths, and bulk LO-phonon scattering times are shown in Fig. 7-13. All parameters were calculated using 80% conduction band offset. The injection anticrossings are  $\Delta_{1'6} = 1.4 \text{ meV}$  and  $\Delta_{2'6} = 1.2 \text{ meV}$ . The parasitic anticrossing is  $\Delta_{1'4} = 0.3 \text{ meV}$ , which is considerably smaller than 0.65 meV for FL175C.

FL148I-M1 was processed into metal-metal structures using the In-Au wafer bonding method. After the substrate and 300-nm-thick  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  etch-stop layer was removed, the  $n^+$  contact layer was thinned down to approximately 200-nm thick. Ti/Au (200/4000 Å) contacts were evaporated and used as etch masks for dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm). Wire bonds were made directly to the top of the devices. The calculated waveguide loss was  $\alpha_w = 16.2 \text{ cm}^{-1}$  with a confinement factor of  $\Gamma = 0.98$ .

This device met with mixed results. Lasing was achieved in pulsed mode despite

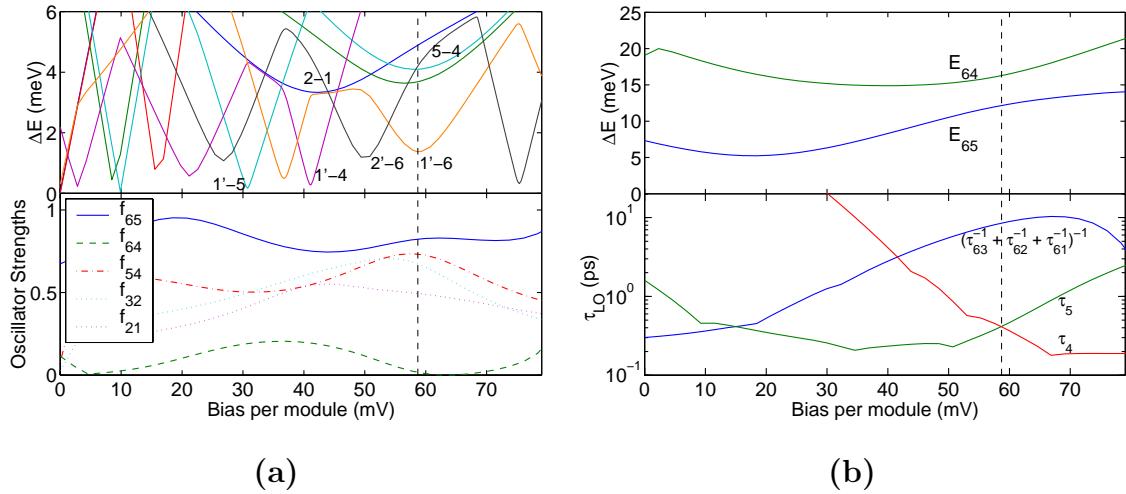


Figure 7-13: (a) Anticrossing plot and oscillator strengths and (b) energy differences and LO-phonon scattering times for FL148I-M1. All data was calculated using an 80% conduction band offset, no self-consistency, and layer thicknesses as measured by x-ray diffraction.

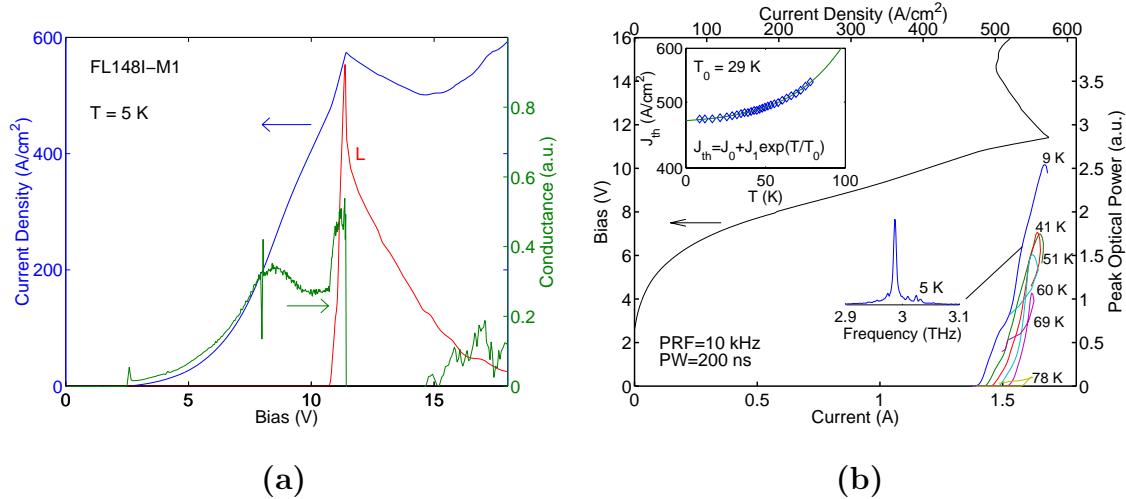


Figure 7-14: (a)  $I-V$ ,  $L-V$ , and conductance (digitally taken  $dI/dV$ ) versus voltage at 5 K in pulsed mode for a 200- $\mu$ m-wide, 1.47-mm-long FL148I-M1mm ridge with metal-metal waveguide (uncoated). (b)  $L-Is$  versus temperature and  $V-I$  at 5 K for the same device. Data was measured using 200-ns pulses repeated at 10 kHz with the photodetector. An inset shows  $J_{th}$  versus heat sink temperature. A spectrum taken at 5 K for a bias of 1.58 A is also shown.

the fact that the transport characteristics were non-ideal. A pulsed  $I$ - $V$  and  $L$ - $V$  are shown in Fig. 7-14(a) along with the conductance versus voltage, as calculated digitally from the  $I$ - $V$  curve. The parasitic channel is observed at approximately 9 V, with  $J_{para} \approx 300$  A/cm<sup>2</sup>. Lasing begins after the device is biased beyond the parasitic channel, at 11 V with a threshold current density of  $J_{th} = 475$  A/cm<sup>2</sup>. However, the NDR occurs very soon afterward at 11.5 V with  $J_0 = 575$  A/cm<sup>2</sup>. As a result, the bias range over which lasing occurs is quite limited. However, even with this limitation, lasing was observed up to 78 K in pulsed mode.  $L$ - $I$  characteristics and  $J_{th}$  at different temperatures are shown in Fig. 7-14(b), along with a typical spectrum taken at 5 K. A fit of  $J_{th}$  versus temperature yielded  $T_0 = 29$  K. The emission frequency is approximately 3.0 THz, which corresponds to a wavelength of  $\lambda \approx 100$   $\mu$ m, or an energy of 12.4 meV. This is close to the calculated value of 12.7 meV. The device only lases for the first 4  $\mu$ s of an applied pulse; this is attributed to the poor heat sinking associated with the metal-metal bonding, as described for FL152Fmm in Sec. 7.3. The discontinuity in the conductance curve at 11.0 V reflects the clamping of the population inversion at threshold, accompanied by at least a partial clamping of the upper state population. The clear visibility of the discontinuity indicates the lower state population is relatively empty and depopulation is efficient [104].

The early onset of NDR is attributed to a mis-design of the injector region that causes the injector miniband to line up too early. The injector region was originally designed to be slightly diagonal at design bias, so that the states would form a funnel to selectively inject carriers into the upper radiative state, as is common in mid-infrared QCL designs [173]. This is useful for mid-IR designs, because the injector miniband width is typically several tens of meV, and there is a risk of injecting carriers high into the upper radiative state where they heat the distribution, or into even higher bound or continuum states where they do not contribute to gain. However, in most of our terahertz designs, the injector miniband width is less than 10 meV, and there are no states close in energy to the upper radiative state. Thus there is no overwhelming advantage to using a funnel injector. Indeed, a miniband injector that is

diagonal at design bias will be flat at a lower bias, which prevents good suppression of the parasitic channel. If this scenario is correct, the observed NDR would be caused by the breaking of the injector miniband rather than misalignment of the injector states with the upper radiative state. The fact that the calculated 2-1 anticrossing occurs at the same bias as the parasitic 1'-4 anticrossing (Fig. 7-13(a)) is consistent with this view. In the future, injector design will be better aimed at the upper state.

Nonetheless, the device FL148I-M1 was relatively successful. The parasitic current channel was reduced to  $\sim 300$ – $400$  A/cm<sup>2</sup>, compared to  $\sim 600$  A/cm<sup>2</sup> for FL175C. This allowed threshold current densities as low as 475 A/cm<sup>2</sup> to be obtained, the lowest yet for resonant phonon terahertz QCLs. The minimum value for  $J_{th}$  appears to be still limited by the parasitic channel, so there is every reason to believe that threshold densities can be reduced much further. Despite the limited current bias range, 78 K operation was obtained, which suggests that much higher temperatures are attainable with improved injection.

## 7.5 CSL104

The CSL (chirped superlattice) structure is a copy of the active region of the first terahertz QCL, described in Köhler *et al.* [64]. These devices were grown in order to reproduce the first lasing structure as a test of both our growth and processing capabilities. A self-consistent conduction band profile for the CSL structures (as designed) is shown in Fig. 7-15, calculated with an 80% band offset.

Four CSL structures were grown (see Table 7.2), but only the last structure lased. The first structure, known as CSL65, was grown with sixty-five modules and a 7- $\mu\text{m}$ -thick active region. This device was fabricated in a metal-metal waveguide using Au-Au thermal compression bonding ( $\alpha_w = 12.4$  cm<sup>−1</sup>,  $\Gamma = 0.97$ ). No lasing was observed, although relatively narrow spontaneous emission was observed at 20 meV (4.9 THz) with a linewidth of 3.4 meV. The next three devices were grown with 104 modules (10.9- $\mu\text{m}$  thick) and were fabricated to have identical SI-surface-plasmon waveguides as the original device [64]. The lower contact layer was 800-nm thick and

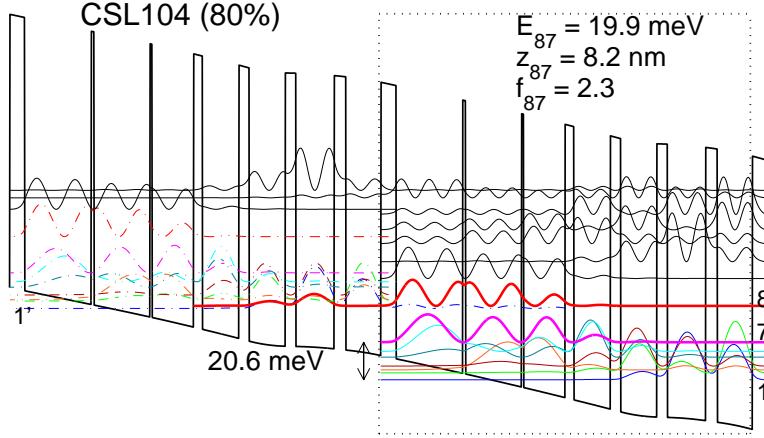


Figure 7-15: Self-consistent conduction band profile for CSL104 calculated using 80% band offset. Beginning with the left injection barrier, the layer thicknesses in Å are 43/188/8/158/6/117/25/103/29/102/30/108/33/99. The 102 Å well is doped at  $n = 4 \times 10^{16} \text{ cm}^{-3}$ , which yields a sheet density of  $4.1 \times 10^{10} \text{ cm}^{-2}$  per module. These layer thicknesses are the design values, as opposed to the actual growth values.

Table 7.2: Characteristics of four chirped superlattice (CSL) growths. Only CSL104-3 lased.

Device	Waveguide	Etching	Top Contact	Growth	Comments
CSL65mm	metal-metal	dry	non-alloyed		
CSL104	SI-SP	wet, dry	non-alloyed		
CSL104-2	SI-SP	wet	alloyed	-5.9%	0.1°C/hour ramp
CSL104-3	SI-SP	dry	alloyed	+4.7%	faster shutter, 0.1°C/hour ramp

was doped at  $2 \times 10^{18} \text{ cm}^{-3}$ . The waveguide was calculated to have loss  $\alpha_w = 14.5$  and a confinement factor of  $\Gamma = 0.475$  at 4.45 THz. CSL104 was fabricated with a non-alloyed ohmic contact on top and using lateral contacts that came within 10  $\mu\text{m}$  of the ridge. Emission spectra revealed extremely broad emission, with approximately 10 meV FWHM. Because of the evidence of growth rate drift for previous long growths, the wafer CSL104-2 was grown with a 0.1°C/hour ramp rate for the MBE source cells. CSL104-2 was fabricated using alloyed NiGeAu ohmic contacts on top of the ridges and using lateral contacts with a 50- $\mu\text{m}$  ridge separation. With this device as well, only very broad emission was observed.

Lasing was achieved in the final device in this series: CSL104-3. The design for this device was nominally identical to CSL104 and CSL104-2. Once again, the source cell temperatures were ramped up during the growth in order prevent growth rate drift and module-to-module non-uniformity. X-ray diffraction revealed that CSL104-3 was over-grown by 4.7%. Additionally, the MBE shutter speed was increased in order to obtain sharper interfaces. Poorly defined interfaces were postulated to have been a possible failure mode in previous growths, especially since the design contains layers as thin as 6 Å. This device was fabricated using alloyed NiGeAu ohmic contacts on top of the ridge and using lateral contacts with a 50-μm ridge separation. Dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm) was used to define ridges, and a short cleanup etch in  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  (10:6:480) was used to remove ~400 nm of GaAs on the sidewalls and bottoms before the lateral contacts were deposited. No significant undercutting at the top metal-semiconductor interface was caused by the cleanup etch, unlike some other devices (see Sec. 5.2.1). This effect may have been suppressed by the use of an annealed alloyed contact, rather than a non-alloyed contact. The rear facets were HR coated by depositing photoresist on the wafer, cleaving it, evaporating  $\text{Al}_2\text{O}_3/\text{Ti}/\text{Au}$  (approximately 3000/150/1500 Å), and then removing the photoresist. However, this method was found to partially short out the devices because of the floppy metallic coating that was left on the edge of the cleave after the photoresist was removed. Some of these devices with these coatings still lased, but with a high parasitic current, and afterward SEM pictures revealed that the corners of the ridge were effectively melted. As a result, the best results were obtained from uncoated devices.

Lasing was seen at approximately 3.9–4.0 THz, which corresponds to a wavelength of  $\lambda = 77 \mu\text{m}$ , or a transition energy of 16.5 meV. This is much lower than the 4.4 THz frequency reported by Köhler in Ref. [64]. The difference is likely due to the 4.7% over-growth of the structure, which results in a reduction in the transition energy. Typical spectra taken in pulsed mode are shown in Fig. 7-16(a) for a 100-μm-wide, 2.97-mm-long ridge structure. Pulsed  $L$ - $I$  and  $V$ - $I$  curves are shown in Fig. 7-16(b) for one of the best performing devices, a 200-μm-wide, 2.97-mm-long ridge structure. In this device,  $J_{th} = 150 \text{ A/cm}^2$  in pulsed mode, and lasing was observed up to 51 K.

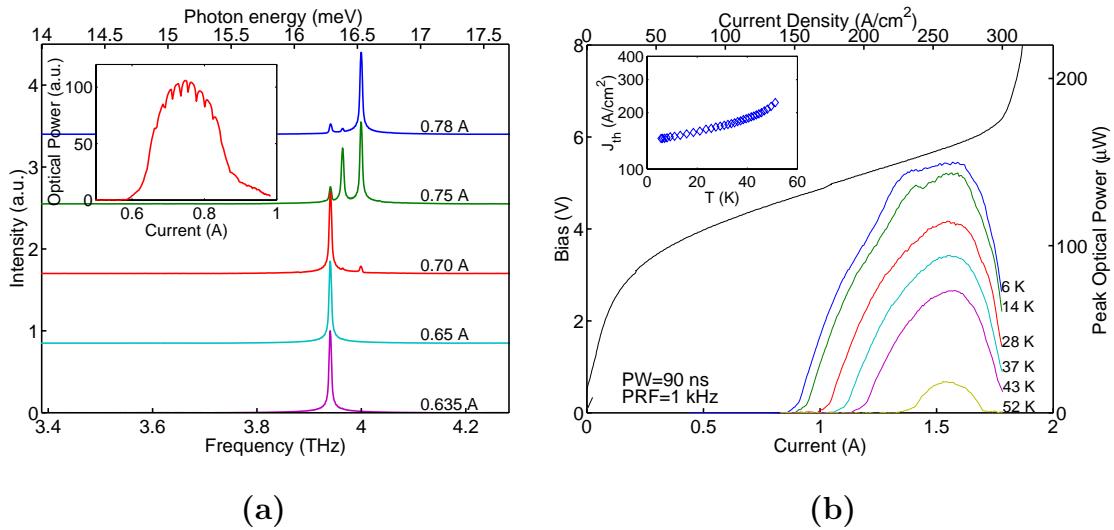


Figure 7-16: (a) Spectra at 5 K taken from 100- $\mu$ m-wide, 2.97-mm-long CSL104-3 ridge using 100-ns pulses repeated at 10 kHz ( $L$ - $I$ s in inset). (b)  $L$ - $I$ s versus temperature for 200- $\mu$ m-wide, 2.97-mm-long uncoated ridge. A  $V$ - $I$  at 5 K is plotted from a smaller device (100- $\mu$ m-wide, 2.97-mm-long), so only the current density should be compared. Data was measured using 90-ns pulses repeated at 1 kHz with the photodetector. An inset shows  $J_{th}$  versus heat sink temperature.

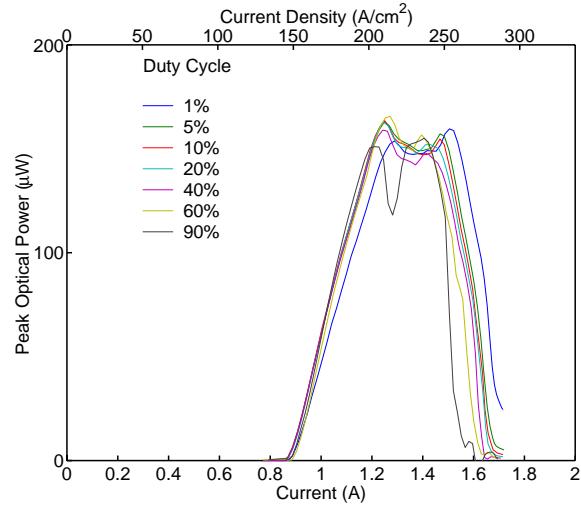


Figure 7-17:  $L$ - $I$ s at various duty cycles for 200- $\mu$ m-wide, 2.97-mm-long uncoated ridge, measured at 5 K with pyroelectric detector. The pulse width was varied for a constant pulse repetition frequency of 1 kHz.

Table 7.3: Characteristics of dry-etched, uncoated CSL104-3 lasing devices. The last two devices were longitudinally cleaved for lateral emission.

Device size	$J_{th}$ (A/cm <sup>2</sup> )	$T_{max}$ (K)	Peak power ( $\mu$ W)	$T_{max,CW}$
60 $\mu$ m × 3.05 $\mu$ m	220	30		24
100 $\mu$ m × 2.97 $\mu$ m	200	19	13.5	13
200 $\mu$ m × 2.97 $\mu$ m	150	51	150	28
95 $\mu$ m × 1.8 $\mu$ m	220	30		24
125 $\mu$ m × 3.12 $\mu$ m	155	51		

Up to approximately 150  $\mu$ W of peak power was collected. Lasing in CW mode was obtained up to a temperature of 28 K.  $L$ - $I$ s taken at various duty cycles at 5 K are shown in Fig. 7-17. Due to the extremely low power dissipation, there is almost no change in  $J_{th}$  or peak power as the duty cycle is changed from 1% to 90%.

Characteristics for a number of CSL104-3 lasers are listed in Table 7.3. The performance of these devices varied from device to device and showed a strong dependence on the width of the ridge structure. For example, the 100- $\mu$ m-wide device had a much higher  $J_{th}$  and a lower  $T_{max}$  than a 200- $\mu$ m-wide ridge adjacent to it. Since the sidewalls are vertical, current spreading is not an issue. This behavior is more likely due to increased loss caused by coupling of the mode with the lossy surface plasmons associated with the lateral contacts [138]. Sidewall roughness may also play an increased role for narrower ridges, but this is unlikely, since the length scale of the roughness ( $< 2 \mu$ m) is smaller than  $\lambda/10$ . In contrast, dry-etched metal-metal devices have shown very little dependence on ridge width, and excellent performance has been obtained in devices as narrow as 60- $\mu$ m wide (see FL178C-M1mm, especially Table 7.5).

The successful operation of CSL104-3 confirmed our ability to grow and fabricate chirped superlattice designs. Our initial difficulties are testament to the importance of extremely high-quality MBE growth. Special effort was taken to ensure sharp interfaces, accurate doping levels, accurate barrier heights, and high uniformity and low drift over the course of the growth. The fact that cw operation was obtained up to temperatures not much lower than maximum pulsed temperatures reflects the

benefits of low threshold currents and voltages. In our lab, this was the first device that lased using dry etching, and the first device that lased using narrow  $60\text{-}\mu\text{m}$ -wide ridges. This indicates that sidewall roughness from dry etching appears not to be a major factor for these devices. However, the degradation of performance as narrower ridges were used even without current spreading suggests mode coupling with lateral contacts may be a hazard in SI-surface-plasmon waveguides. This device is also notable as it is the only laser we have measured that used alloyed ohmic top contacts. These are expected to be detrimental due to the uncontrolled free carrier profile at the surface. It is currently standard when fabricating terahertz QCLs to either use non-alloyed ohmic contacts, or to use alloyed contacts only at the edges of the ridge [65]. Despite the low threshold densities, this chirped superlattice design displayed disappointing temperature behavior. The highest operating temperature was only 51 K, less than that of FL175C (87 K), which had approximately a factor of 10 higher power dissipation. This is likely a result of thermal backfilling of the lower radiative state from the injector miniband in the CSL structure ( $\Delta E \sim 20$  meV), which is expected to be less of an issue for resonant phonon designs ( $\Delta E \sim 36$  meV).

## 7.6 FL178C-M1

The most recent laser tested is FL178C-M1 [174]. The active region for this device is the same as our first laser device FL175C (see Figs. 7-1, 7-2), except that the 24-Å ( $\sim 9$  ML) intra-injector barrier was widened to 30 Å ( $\sim 11$  ML). The purpose of this change was to tighten the injection doublet for better injection efficiency, and to reduce the parasitic current channel. X-ray diffraction revealed that this device was under-grown by 4.4%; this is similar to FL175C which was under-grown by 4.2%. As a result of the thicker barrier, the 2-1 anticrossing gap is smaller:  $\Delta_{21} = 4.6$  meV, compared to  $\Delta_{21} = 6.1$  meV for FL175C. The calculated injection anticrossings are  $\Delta_{1'5} = 1.9$  meV and  $\Delta_{2'5} = 1.3$  meV and the parasitic anticrossing is  $\Delta_{1'3} = 0.67$  meV—which is strangely not any smaller than for FL175C, despite the thicker barrier.

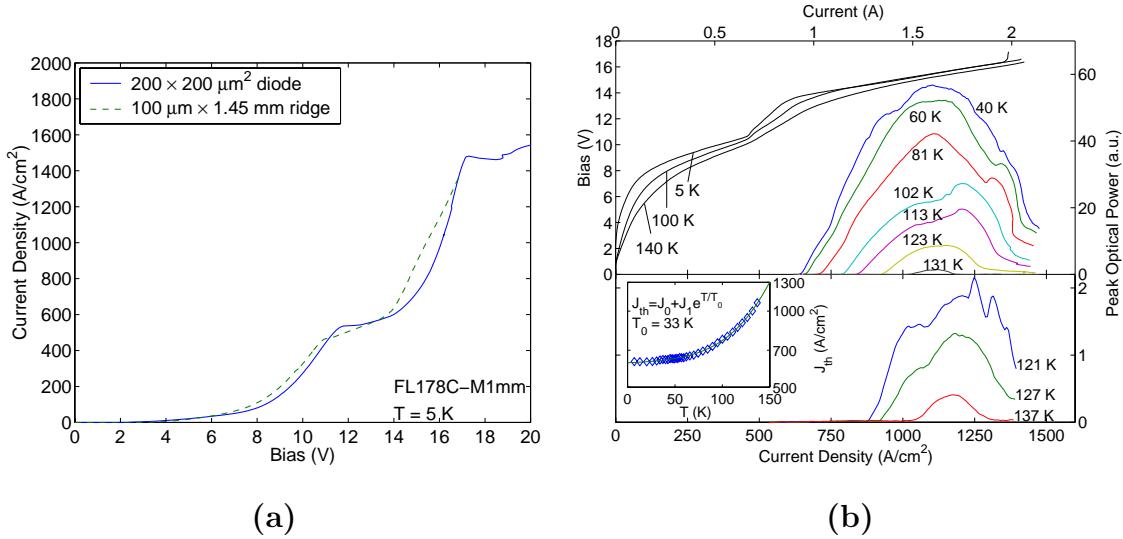


Figure 7-18: (a) Pulsed  $I$ - $V$  characteristic for a  $200 \times 200 \mu\text{m}^2$  FL178C-M1mm diode and  $100\text{-}\mu\text{m-wide, } 1.45\text{-mm-long ridge structure that lases.}$  (b) Typical  $L$ - $I$ s and  $V$ - $I$  from FL178C-M1mm devices. The  $L$ - $I$  in upper panel is from  $60\text{-}\mu\text{m-wide, } 2.48\text{-mm-long ridge.}$  The  $V$ - $I$  was taken using a smaller device ( $100\text{-}\mu\text{m-wide, } 1.45\text{-mm-long},$  so only the current density axis is valid. The  $L$ - $I$ s in the lower panel are from a  $150\text{-}\mu\text{m-wide, } 2.74\text{-mm-long FL178C-M1mm ridge.}$  Inset gives  $J_{th}$  versus temperature. Data was measured using 200-ns pulses repeated at 1 kHz with the photodetector.

This device was grown on a semi-insulating wafer with a 300-nm-thick  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  etch stop layer below a 800-nm-thick  $n^+$  contact layer doped at  $3 \times 10^{18} \text{ cm}^{-3}$ , to allow for processing using either a metal-metal or SI-surface-plasmon waveguide. Metal-metal processing was performed using the In-Au wafer bonding method, whereupon the substrate was removed and the contact layer was thinned down to approximately 200 nm. Ti/Au (200/4000 Å) contacts were evaporated and used as etch masks for dry etching in  $\text{BCl}_3:\text{N}_2$  (15:5 sccm). Waveguide losses are calculated to be  $\alpha_w = 14.2 \text{ cm}^{-1}$  with a confinement of  $\Gamma = 0.98$  (at 3.8 THz) (using  $\tau = 0.5 \text{ ps}$  for the active region,  $\tau = 0.1 \text{ ps}$  for the heavily doped regions, and  $\tau = 0.05 \text{ ps}$  for gold).

The  $I$ - $V$  for a small non-lasing  $200 \times 200 \mu\text{m}^2$  diode is shown in Fig. 7-18(a) along with the  $I$ - $V$  from a  $100\text{-}\mu\text{m-wide, } 1.45\text{-mm-long ridge structure that lases.}$  The diode  $I$ - $V$  clearly shows an NDR at 17 V which corresponds to the design bias, and a parasitic NDR at  $\sim 11 \text{ V}$ . The high value for the design bias is partially caused by a larger than usual ( $\sim 1 \text{ V}$ ) parasitic voltage drop, which can be seen as an offset in

the  $I$ - $V$  where no current flows. Usually, the voltage drop associated with the Ti/Au Schottky contact is only  $\sim 0.5$  V. This larger turn-on voltage is probably caused by over-etching the top  $n^+$  contact layer so that the metallization was deposited directly on the low-doped active region instead. However, given that the parasitic NDR is calculated to occur at only  $178 \times 46$  mV  $\approx 8.2$  V, a one volt parasitic drop is insufficient to completely account for the observed voltage of 11 V. This indicates that there is likely additional parasitic resistance, perhaps associated with the lower contact. The  $I$ - $V$  for the laser ridge is very similar to the diode  $I$ - $V$ , except that for biases greater than  $\sim 14$  V, it displays a higher differential conductance, which is due to the effect of stimulated emission above threshold on the transport.

Lasing was observed in all tested metal-metal waveguide devices, which are summarized in Table 7.5. Since devices that were facet coated tended to short out, uncoated devices were tested. The first device tested, which was 100- $\mu\text{m}$  wide and 1.45-mm long, lased up to 127 K in pulsed mode. However, despite this excellent pulsed temperature performance, these devices only lased for the first 10–30  $\mu\text{s}$  of an applied pulse. Just as for FL152Fmm (Sec. 7.3), the thermal resistance at the metal-metal bonding interface impedes heat removal and consequently limits the duty cycle. The highest peak power observed was low: roughly 200–300  $\mu\text{W}$  as measured by the Scientech thermopile detector.

Various  $L$ - $I$  curves taken at different temperatures are shown in Fig. 7-18(b). The highest operating temperature device was 150- $\mu\text{m}$  wide and 2.74-mm long, and lased up to 137 K, which is currently the highest reported operating temperature for a terahertz quantum cascade laser. Also shown are  $L$ - $I$  characteristics from a 60- $\mu\text{m}$ -wide, 2.48-mm-long ridge that lased up to 132 K. Lasing begins almost immediately beyond the parasitic channel, which indicates that  $J_{th}$  is still limited by the parasitic current density, rather than by intrinsic gain or loss. The sharp kink in the  $V$ - $I$  associated with the parasitic current channel at  $\sim 11$  V indicates the development of high-field domains. Once this occurs, the bias in the structure becomes hard to control, and propagation of these domains can break modules from their design biases. This is likely the reason for the occasional abrupt changes in the slope efficiency as

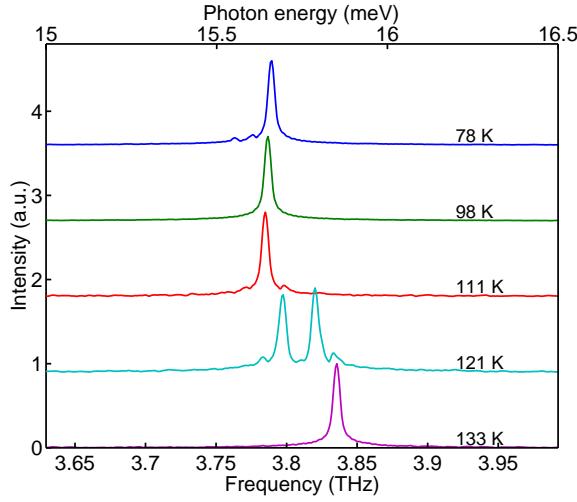


Figure 7-19: Spectra at various heat sink temperatures from a 100- $\mu\text{m}$ -wide, 2.74-mm-long FL178C-M1mm uncoated ridge. Spectra were taken using 200-ns pulses repeated at 5 kHz with the photodetector, all with approximately the same current densities of 1170–1200 A/cm<sup>2</sup>.

the bias is increased. Furthermore, when the photodetector output trace was viewed on an oscilloscope, the signal would flicker and display different pulse shapes, which indicated that the device was being biased to different conditions during each applied pulse. This effect may also explain the relatively low output powers measured from these metal-metal waveguide devices.

Spectra from a 100- $\mu\text{m}$ -long, 2.74-mm-wide FL178C-M1mm uncoated ridge are shown in Fig. 7-19 for heat sink temperatures up to 133 K. The device lased at approximately 3.8–4.0 THz ( $\lambda \approx 78 \mu\text{m}$ ), which corresponds to an energy of  $\sim 16$  meV. This is slightly higher than calculated and observed for FL175C. Since x-ray diffraction suggests roughly the same amount of undergrowth for the two devices, it is unlikely that growth fluctuations are the reason for this discrepancy. The most likely explanation is that the modification of the injector causes injection to occur at a higher bias, so that the radiative transition is Stark shifted and lasing takes place at a higher energy. The development of high-field domains in FL178C-M1mm, which did not occur in the original FL175C device, may cause modules to “jump” in voltage to a higher bias point than might otherwise occur with a smooth bias transition. The spontaneous emission linewidth was measured to be 6 meV ( $\sim 1.5$  THz). The line-

Table 7.4: Characteristics of FL178C-M1mm devices that successfully lased. Listed devices were uncoated.

Device	Etching	$J_{th}(\text{A}/\text{cm}^2)$ (5 K)	$T_{max}$ (K)	$\text{PW}_{max}$ ( $\mu\text{s}$ )
100 $\mu\text{m} \times 1.45$ mm	dry	632	127	10
200 $\mu\text{m} \times 2.18$ mm	dry	639	115	
150 $\mu\text{m} \times 2.74$ mm	dry	625	137	
100 $\mu\text{m} \times 2.74$ mm	dry		136	
80 $\mu\text{m} \times 2.74$ mm	dry	625	131	
400- $\mu\text{m}$ diameter	dry	675	105	
60 $\mu\text{m} \times 2.48$ mm	dry	630	132	30
80 $\mu\text{m} \times 2.48$ mm	dry	645	125	
80 $\mu\text{m} \times 2.48$ mm	dry	959	114	

shape was far different from a Lorentzian, and this broad linewidth is likely due to non-uniform alignment of different modules, as well as slightly different module alignments during every applied pulse. The longitudinal mode spacing is approximately  $0.45 \text{ cm}^{-1}$ , which corresponds to an effective index of  $4.05 \pm 0.2$ .

These devices were also notable in that excellent performance was observed in ridge structures as narrow as 60- $\mu\text{m}$  wide. As shown in Table 7.4, three side-by-side ridge structures of the same length and on the same die displayed only a slight decrease in  $T_{max}$  as narrower ridges were used. Since these devices were dry etched and have nearly vertical sidewalls, they do not suffer from the current spreading problems that reduced peak gain in wet-etched FL175C and FL152Fmm devices. Also, the metal-metal waveguides may be more compatible with narrower ridge structures than SI-surface-plasmon waveguides, since there is less mode leakage and no risk of coupling with lossy surface plasmons under the lateral contacts [138]. Furthermore, there is no evidence of any detrimental effects from the sidewall roughness (seen in Fig. 5-14), which would be expected increase in importance with narrower guides. This is to be expected, since the radiation wavelength inside the guide is approximately  $20 \mu\text{m}$ , and the typical length scale of the roughness is less than a few microns.

FL178C-M1 was also processed into SI-surface-plasmon waveguides. Ti/Au contacts were deposited to form non-alloyed ohmic contacts and ridges were wet-etched

Table 7.5: Characteristics of FL178C-M1 (SI-surface-plasmon) devices that successfully lased. Listed devices were rear facet HR coated.

Device size	Etching	$J_{th}$ (A/cm <sup>2</sup> ) (5 K)	$T_{max}$ (K)	Max duty cycle
100 $\mu$ m $\times$ 1.47 mm	wet	750	63	> 90%
150 $\mu$ m $\times$ 1.50 mm	wet	702	83	cw (for 3 s)
150 $\mu$ m $\times$ 2.95 mm	wet	739	68	
150 $\mu$ m $\times$ 2.97 mm	wet	670	92	

using H<sub>3</sub>PO<sub>4</sub>:H<sub>2</sub>O<sub>2</sub>:H<sub>2</sub>O (1:1:25). NiGeAu lateral contacts were deposited 50  $\mu$ m away from the ridges. Facets were coated using Ti/Au ( $\sim$ 150/1500 Å) evaporated over Al<sub>2</sub>O<sub>3</sub>. The waveguide loss was calculated to be  $\alpha_w = 8.2 \text{ cm}^{-1}$  with a confinement of  $\Gamma = 0.324$ . These devices did not attain the same high temperatures as the metal-metal devices. The maximum temperature obtained in pulsed mode was 92 K for a 150- $\mu$ m-wide, 2.97-mm-long, rear HR coated structure. The difference is mostly due to the extra loss and low confinement of the SI-surface-plasmon structure. The effect of current spreading and module bias non-uniformity may also reduce the available gain. For a 150- $\mu$ m-wide, 1.5-mm-long coated ridge, continuous wave operation was briefly observed for approximately 3 s before heat buildup in the cryostat stopped operation. A peak pulsed power of 2.5 mW was observed in pulsed mode for this device at 5 K. A summary of tested devices is given in Table 7.5. There is some uncertainty of device areas for the wet etched devices, i.e. a 100- $\mu$ m-wide device is approximately 81- $\mu$ m-wide at the top of the ridge and 101- $\mu$ m-wide at the bottom. Therefore, the listed threshold densities may be somewhat of an underestimate.

The modification of the injector region was successful in reducing the parasitic current to  $J_{para} \sim 500 \text{ A/cm}^2$ , especially compared to the large current density at design bias  $J_0 \approx 1500 \text{ A/cm}^2$ . This allowed threshold current densities as low as 625 A/cm<sup>2</sup> at 5 K. However, since the minimum  $J_{th}$  is still limited by the parasitic channel, substantial further reductions in the threshold are possible. Also, the development of the high-field domains, the low observed power, the broad inhomogeneous spontaneous emission linewidth, and the abrupt changes in the  $L$ - $I$  characteristic before NDR all suggest that even when lasing, some or many modules may not be at

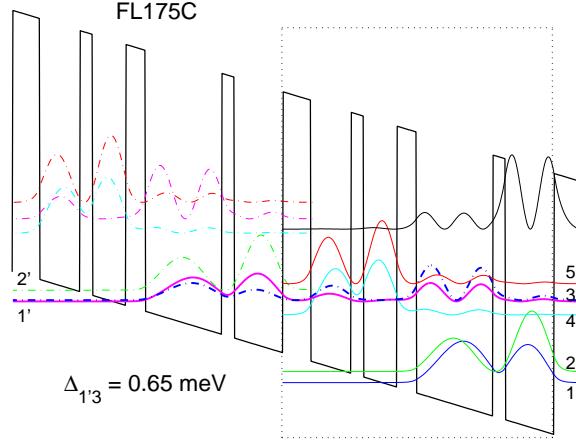


Figure 7-20: Conduction band profile for FL175C (80% barrier offset) calculated at the parasitic 1'-3 anticrossing point.

the proper bias, or may break off too early. These modules contribute no gain. If the parasitic channel is reduced and high-field domain development can be prevented, even lower thresholds and higher operating temperatures can be obtained, not to mention a much higher slope efficiency and output power. The excellent temperature performance of this device is partially due to the use of a low-loss, high-confinement metal-metal waveguide, and partially due to improved injection efficiency and reduced parasitics associated with the thicker barrier. As mentioned above, when lasing the radiative transition in this device may be more diagonal than FL175C, and this may also help to improve the injection selectivity and upper state lifetime, as described in Sec. 7.9.

## 7.7 Parasitic current channels

The parasitic current channel that appears as a shoulder in the  $V$ - $I$  for FL175C (Fig. 7-5), FL152Fmm (Fig. 7-10), FL148I-M1mm (Fig. 7-14), and FL178C-M1mm (Fig. 7-18) places a lower limit on the threshold current densities of these lasers. This in turn limits the maximum operating temperature and duty cycle. Below this bias point, most current flows through the parasitic channel rather than the upper radiative state and does not contribute to gain. This is indicated by the two slope

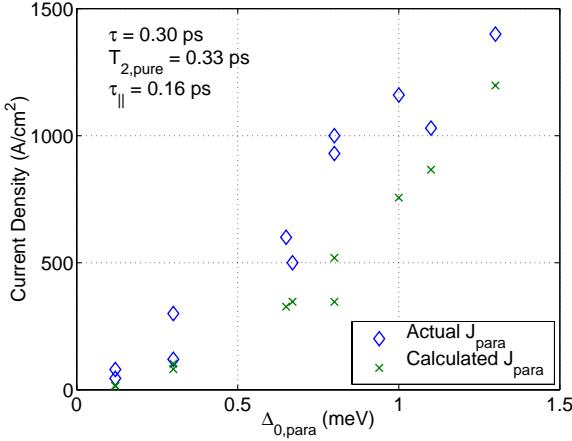


Figure 7-21: Actual and predicted parasitic current densities using Eq. (7.3), with  $\tau_{\parallel}^{-1} = \tau^{-1} + T_2^{*-1}$  as per Eq. (2.61).

behavior present in the subthreshold  $L$ - $I$  characteristics for these devices (Fig. 7-3). It is likely that this channel corresponds to injection of current from the  $n = 1'$  (and sometimes  $n = 2'$ ) injector state into the excited state in the wide well in the next module, which is known as  $n = 3$  in FL175C. Such a bias condition is shown for FL175C in Fig. 7-20, and corresponds to the anticrossing of states  $1'$  and  $3$  with the gap  $\Delta_{1'3} = 0.65$  meV. Since the state  $n = 3$  always has an extremely short lifetime due to resonant LO-phonon scattering ( $\tau_3 \approx 0.3$  ps for FL175C) it can be an efficient transport mechanism even with a small tunneling interaction. In fact, this parasitic channel is present in most devices where resonant-phonon depopulation is used, although it varies greatly in magnitude. For devices like FL125 and FFL140 the channel is relatively small, whereas in other devices such as FL179E, FL152F, FL162G, and FL137H the channel is much larger. For these last four devices, the parasitic channel is so strong relative to the design injection that high-field domains develop when the devices are biased past the parasitic point. This can be seen in the sharp features in the  $I$ - $V$  characteristics in Figs. 7-10, 6-35(b). The peak and parasitic current densities as well as the associated anticrossing gaps are listed in Table 7.6 for various devices.

The tight-binding density matrix formalism for resonant tunneling developed in Sec. 2.5.2 can be used to analyze these parasitic current channels. For tunneling

Table 7.6: Current densities  $J_0$  and  $J_{para}$  at injection point and parasitic point respectively, and calculated injection and parasitic anticrossing gaps  $\Delta_0$  and  $\Delta_{para}$  for various devices. Devices with no data for  $J_0$  had no clearly defined NDR at the design point. The sheet density per module  $N_{mod}$  is given, although only for FL125 was the doping measured. All data calculated with 80% conduction band offset.

Device	$N_{mod}$ (cm $^{-2}$ )	$J_0$ (A/cm $^2$ )	$J_{para}$ (A/cm $^2$ )	$\Delta_0$ (meV)	$\Delta_{para}$ (meV)
FL125	$3.4 \times 10^{10}$	200-300	45	1.0	0.12
FL175B	$4.0 \times 10^{10}$	330	80	0.75	0.12
FL175C	$2.8 \times 10^{10}$	1000	600	1.8	0.65
FL175D	$3.2 \times 10^{10}$		120	0.8	0.3
FL179E	$2.9 \times 10^{10}$		1400	2.5	1.3
FL152F	$3.0 \times 10^{10}$	1500	1000	2.45	0.8
FL162G	$2.9 \times 10^{10}$		1160	2.25	1.0
FL137H	$2.8 \times 10^{10}$		1030	3.0	1.1
FFL140	$4.0 \times 10^{10}$	1250	300	1.7	0.3
FL178C-M1	$2.8 \times 10^{10}$	1500	500	1.9	0.67
O169	$2.0 \times 10^{10}$		930	1.9	0.8

between two localized states, Eq. (2.64) gives the maximum current at resonance (zero detuning) as

$$J_{max} = \frac{eN}{2} \frac{(\frac{\Delta_0}{\hbar})^2 \tau_{||}}{1 + (\frac{\Delta_0}{\hbar})^2 \tau \tau_{||}}, \quad (7.3)$$

where  $\tau_{||}$  is the dephasing time,  $\tau$  is the relaxation time of the target state,  $\Delta_0$  is the anticrossing gap, and  $N$  is the total population of the two levels of interest. The condition  $(\Delta_0/\hbar)^2 \tau \tau_{||} \gg 1$  is known as the coherent limit and the condition  $(\Delta_0/\hbar)^2 \tau \tau_{||} \ll 1$  is known as the incoherent limit. This expression was used to predict the various parasitic current densities using a lifetime of  $\tau = 0.3$  ps and a pure dephasing time of  $T_2^* = 0.33$  ps (see Fig. 7-21). This value of  $\tau$  was chosen to correspond with the calculated time for FL175C, but might vary from 0.15–0.4 ps depending on how close the scattering is to the LO-phonon resonance. The value for dephasing is even more uncertain, but as described in Sec. 2.5.2, it is chosen to correspond with a homogeneous radiative linewidth broadening of  $\sim 1$  THz ( $\sim 4$  meV). The population  $N$  is taken to be  $N_{mod}$ , since at this low bias point and low temperature, most of the carriers are in state  $n = 1$ . The calculated values for  $J_{para}$  display the same trend as the measured

values, but underestimate the values slightly. This is not surprising due to the highly uncertain value for the dephasing rates. For all of these devices the parasitic tunneling is in the “incoherent” regime and transport is limited by dephasing and tunneling through the barriers rather than the final state lifetime  $\tau$ . Because of this, increasing  $\tau$  actually will raise  $J_{para}$ , due to the concomitant reduction in dephasing. A better fit to the experimental data might also be obtained if a value for the conduction band offset smaller than 80% was used, since the recalculated anticrossing gaps would be larger.

For laser designs, it is desirable to reduce  $J_{para}$  with respect to  $J_0$ .  $J_0$  can be estimated using the same equation (7.3), although the upper radiative state can be expected to have a much longer lifetime  $\tau_0 \approx 1 - 5$  ps compared with  $\tau_{para} \approx 0.3$  ps. Also, at design bias, the electron population will be more evenly distributed among the subbands than at the parasitic bias, so the term  $N = N_0$  may be smaller than  $N_{para}$ . Using the expression (7.3), we can write the ratio of the two currents as

$$\frac{J_0}{J_{para}} = \frac{N_0}{N_{para}} \left( \frac{\Delta_0}{\Delta_{para}} \right)^2 \left( \frac{\tau_{||,0}}{\tau_{||,para}} \right) \frac{1 + (\frac{\Delta_{para}}{\hbar})^2 \tau_{para} \tau_{||,para}}{1 + (\frac{\Delta_0}{\hbar})^2 \tau_0 \tau_{||,0}}, \quad (7.4)$$

where any differences in the dephasing times for the two cases have been explicitly included via  $\tau_{||,0}$  and  $\tau_{||,para}$ . There are three cases that can arise.

1. Both injection and parasitic tunneling are incoherent:

$$\frac{J_0}{J_{para}} = \frac{N_0}{N_{para}} \left( \frac{\Delta_0}{\Delta_{para}} \right)^2 \left( \frac{\tau_{||,0}}{\tau_{||,para}} \right). \quad (7.5)$$

2. Injection is coherent and parasitic tunneling is incoherent:

$$\frac{J_0}{J_{para}} = \frac{N_0}{N_{para}} \left( \frac{\Delta_0}{\Delta_{para}} \right)^2 \left( \frac{\tau_{||,0}}{\tau_{||,para}} \right) \frac{1}{1 + (\frac{\Delta_0}{\hbar})^2 \tau_0 \tau_{||,0}} = \frac{N_0}{N_{para}} \frac{\hbar^2}{\Delta_{para}^2 \tau_{||,para} \tau_0}. \quad (7.6)$$

3. Both injection and parasitic tunneling are coherent:

$$\frac{J_0}{J_{para}} = \frac{N_0}{N_{para}} \frac{\tau_{para}}{\tau_0}. \quad (7.7)$$

To maximize  $J_0/J_{para}$ , both transport channels should operate in the incoherent limit, where the difference in anticrossing gaps has the most effect. Once the injection enters the coherent limit,  $J_0$  saturates and depends on  $\tau_0$  rather than  $\Delta_0$ , so the advantage of the anticrossing gap ratio is wasted. During design it is often straightforward to increase or decrease the two anticrossing gaps by adjusting barrier thicknesses while keeping  $\Delta_0/\Delta_{para}$  constant. The requirement for both channels to be incoherent puts an upper limit on how large  $\Delta_0$  can be made while still controlling the parasitic channel. The precise limit depends on the upper radiative state lifetime and the dephasing time, both of which are highly uncertain. Condition (3), where both channels are coherent, is highly undesirable, as it would result in a ratio  $J_0/J_{para} < 1$ . Luckily, for typical values of anticrossing gaps and lifetimes, this condition is difficult to realize.

The nature of injection tunneling is somewhat difficult to classify, the radiative upper state lifetime as well as the dephasing time is somewhat uncertain. Again assuming  $T_2^* = 0.33$  ps, to be in the coherent limit requires anticrossing gaps  $\Delta_0 > 0.5 - 1.15$  meV for upper state lifetimes ranging from  $\tau = 5$  ps to  $\tau = 1$  ps. This suggests that many of the current structures may be marginally in the coherent limit for injection (see Table 7.6).

In summary, the tight-binding resonant tunneling model gives a rough agreement with experimental data for the parasitic channel strength. To minimize the effect of the parasitic current channels one should design structures for the maximum ratio of injection to parasitic anticrossing gaps. This is an intuitive conclusion, as it simply means that one should reduce coupling to parasitic states from the injector states by increasing spatial separation, barrier thicknesses, etc.. However, for a given ratio of anticrossing gaps, the injection should not be made too strong, or tunneling will enter the coherent regime, which throws away the advantage provided by a large anticrossing gap ratio. However, there are many other design criteria that limit the implementation of this strategy. For example, traditional QCL design philosophy favors coherent injection into the upper radiative state so that tunneling does not provide a bottleneck to transport. Also, the collector barrier cannot be made overly thick to block the parasitic channel, since this may slow depopulation at design

bias. Our current design methodology is to make the injection marginally coherent ( $\Delta_0 \sim 1.5$  meV), while minimizing  $\Delta_{para}$ , primarily by optimizing injector alignment. Adding wells to the injector region can help prevent parasitic coupling below design bias (see FL148I-M1, Sec. 7.4), but this potentially creates a wide mini-band injector, which is a potential source of optical cross-absorption.

The development of high-field domains was seen in a number of devices, including FL179E, FL152F, FL162G, FL137H, and FL178C-M1. This phenomenon results from a region of negative differential resistance beyond the parasitic bias. Ideally, as a device is biased into an NDR region, the electric field would remain uniform and the subbands would move out of alignment in all modules simultaneously, which would cause the current to decrease. Instead, charge will build up at bottleneck points and will cause one or many modules to jump to a higher bias point so that constant current can be maintained. This can occur even if  $J_0 > J_{para}$ , if the bias point in between the parasitic and design bias displays a local current minimum. Following the onset of high-field domains, a uniform bias condition may be recovered at a later point, but there is no guarantee that all the modules will be at the correct bias. In order to avoid high-field domain development, there must be a smooth transition from parasitic injection to design injection. In other words, injection should not be perfectly selective. The best strategy to prevent high-field domain development is simply to minimize  $J_{para}$  with respect to  $J_0$ . That will allow injection to be made as selective as possible while allowing a gradual shift between bias conditions, and a large bias current range where the injection efficiency is high. If  $J_{para}$  cannot be reduced sufficiently, it may be necessary to move the two bias voltages closer together to allow a smoother transition.

## 7.8 Intra-miniband oscillator strengths

The resonant phonon QC-lasers discussed in this chapter were based on a radiative transition between two states that were composed of anticrossed single-quantum-well ground states. No excited states—where the wavefunction has a node inside a well—

were used. To make an analogy, one could describe such an intersubband transition as an intra-miniband transition, if one considers the coupled double quantum well to be a two-well “superlattice.” This is in contrast to the chirped superlattice QCLs, where the radiative transition is an inter-miniband transition. The oscillator strength of the first chirped superlattice structure was  $f \sim 2.7$ , which is large compared to that of FL175C, where  $f = 0.96$ . Obviously, larger oscillator strengths could be valuable in improving the gain of resonant phonon devices. This was the motivation for devices such as FFL140, FL162G, and FL137H, where additional wells were added to the active region to increase the oscillator strength.

To evaluate the relative merits of various intra-miniband transitions, the oscillator strengths between various subbands in  $n$ -well “superlattices” have been evaluated.<sup>1</sup> To allow an accurate comparison, the well and barrier thicknesses have been adjusted to keep the transition energy approximately constant ( $14.0 \pm 0.2$  meV), and the final state at the same energy ( $\sim 26.8$  meV above the well bottom). These oscillator strengths have been listed in Table 7.7. Take for example, the FL-series of structures where the radiative transition is between the two lower states ( $2 \rightarrow 1$ ). There is clearly an advantage to using more wells, up to about 4 wells. This benefit was seen in FL162G and FL137H, where 3-well active regions were used. Increasing the number of wells beyond four yields little additional oscillator strength, and may even hurt overall gain, since the module length will increase slightly. Additional oscillator strength can be gained by using the  $3 \rightarrow 2$  transition, as in FFL140. However, the disadvantage to this strategy is that below the lower radiative state an additional subband  $n = 1$  must be depopulated along with  $n = 2$ , which proved difficult to achieve in FFL140.

The results of Table 7.7 can be simply understood by considering the oscillator strength and sum rule discussed in Sec. 2.3.1. Recall from Eqs. (2.25) and (2.26) that

---

<sup>1</sup>Care should be taken in comparing these calculated oscillator strengths with actual devices, since the lower radiative state in resonant-phonon designs is also coupled with the state in the wide phonon well, which enhances  $f$ .

Table 7.7: Oscillator strengths for various transitions for  $n$ -well superlattices. Quantities in parentheses are well/barrier thicknesses for structures in monolayers (1 ML=2.825 Å). 15% aluminum barriers were used. Structures were adjusted to keep the relevant transition energy approximately constant at 14.0 meV, and the lower radiative state 26.8 meV from the well bottom.

	$ f_{2 \rightarrow 1} $	$ f_{3 \rightarrow 2} $	$ f_{4 \rightarrow 3} $
2-well	0.65 (30/7)		
3-well	0.84 (24.6/5.3)	1.43 (33.5/3.8)	
4-well	0.90 (20.3/4.0)	1.70 (28.3/3.7)	2.43 (36.6/1.6)
5-well	0.91 (17.0/3.2)	1.78 (23.9/3.0)	2.64 (30.3/1.8)
6-well	0.93 (14.5/2.7)	1.78 (25.9/1.6)	2.68 (25.8/1.5)

the sum rule states that

$$\sum_{f \neq i} f_{i \rightarrow f} = 1, \quad (7.8)$$

where

$$f_{i \rightarrow f} = \frac{2m^*(E_f - E_i)|z_{i \rightarrow f}|^2}{\hbar^2}. \quad (7.9)$$

However, this rule does not actually require that  $|f_{i \rightarrow f}| < 1$  for  $i > 1$ . Since downward transitions contribute negative oscillator strength, the presence of other subbands below a state requires correspondingly larger oscillator strengths for upward transitions to compensate so that (7.8) holds. For example, consider an infinite quantum well, where  $f_{1 \rightarrow 2} = 0.96$ . Since  $f_{2 \rightarrow 1} = -0.96$ , the sum rule for  $i = 2$  would then read

$$\sum_{n \geq 3} f_{2 \rightarrow n} = 1.96. \quad (7.10)$$

For this reason, oscillator strengths between higher subbands in the miniband tend to be larger.

Considering our  $n$ -well “superlattices,” for constant transition energies, as more wells are added, the wells and barriers tend to become narrower and we approach the limit of a single quantum well. Note that  $|f_{2 \rightarrow 1}|$  in Table 7.7 is approaching the value 0.96. The oscillator strength between transitions of opposite parity in an infinite

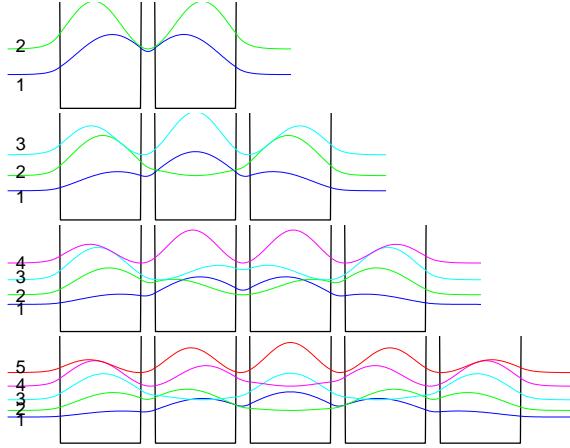


Figure 7-22: Magnitude squared wavefunctions for multi-well 30/5 ML superlattices with 15% Al barriers.

quantum well is [37]

$$f_{i \rightarrow f} = \frac{64}{\pi^2} \frac{i^2 f^2}{(i^2 - f^2)^3}; \quad (7.11)$$

$f_{i \rightarrow f} = 0$  for states of like parity. Note that this expression has no width dependence. Thus the saturation of the oscillator strength with increasing number of wells is thus easily understood.

However, the choice of an intra-miniband radiative transition should also be governed by the wavefunction probability distributions for the various states. The magnitude squared wavefunctions for various  $n$ -well superlattices with identical layer thicknesses are shown in Fig. 7-22. One disadvantage of using a  $2 \rightarrow 1$  transition in a 3-well active region is that state  $n = 2$  has very little amplitude in the middle well. This large amplitude in the outer wells will potentially increase coupling with lower states and decrease its lifetime. This effect can be seen in FL137H, which has a relatively short upper state lifetime of  $\tau_{6 \rightarrow (3,2,1)} \approx 3.5$  ps. This effect can be suppressed somewhat by making the transitions slightly diagonal. Nevertheless, because state  $n = 2$  has larger middle well amplitudes, a four-well active region may be a better choice than a three-well active region to obtain large oscillator strength and comparatively longer upper state lifetimes.

## 7.9 High temperature operation

As of this writing, the maximum operating temperatures for terahertz QCLs are 137 K for a resonant phonon design (Sec. 7.6), 103 K for a bound-to-continuum design [45], and 74 K for a chirped superlattice design [175]. As the operating temperature of a quantum cascade laser rises, losses increase and the available gain is reduced, and higher current densities are required to reach threshold. The upper limit is placed on  $J_{th}$  by the current density at the NDR point, beyond which the injection levels become misaligned, and carriers are no longer injected into the upper state. Often, this behavior in lasers is fit by a phenomenological model  $J_{th} = J_0 + J_1 \exp(T/T_0)$ . This function appears to provide a reasonably good fit, albeit only for the relatively low temperatures achieved. However, this is not a physical model, and the characteristic temperature  $T_0$  merely characterizes a combination of processes that act to decrease the gain and increase losses as temperature rises.

Using the generalized model for a three-level QCL, where the radiative transition takes place between  $n = 3$  and 2, we can write the following expression for the threshold density  $J_{th}$ :

$$J_{th} \propto \frac{g_{th}\Delta\nu}{f_{32}\Delta N_{th}} \propto \frac{g_{th}\Delta\nu}{f_{32}\tau_3(1 - \frac{\tau_2}{\tau_{32}})} \quad (7.12)$$

where  $\Delta N_{th}$  is the threshold population inversion, the threshold gain is  $g_{th} = (\alpha_w + \alpha_m)/\Gamma$ ,  $\Delta\nu$  is the transition linewidth, and we have assumed unity injection efficiency. The upper state lifetime is given by  $\tau_3^{-1} = \tau_{32}^{-1} + \tau_{3,para}^{-1}$ , where  $\tau_{3,para}^{-1}$  is a parasitic scattering time. This model is described in more detail Appendix C. Note that this generalized model is slightly different than for resonant phonon lasers where the lower radiative state is the upper state of a doublet, and hence two thermally activated scattering rates should be accounted for (one for each state in the doublet).

### Active region

In a terahertz QCL, the increase in  $J_{th}$  versus temperature is primarily due to two factors. First, population inversion can be reduced due to thermal backfilling of carriers from the injector states into the lower radiative state  $n = 2$ . This effect can be

thought of as an effective increase in the depopulation time  $\tau_2$ . This is expected to be less of a problem for our resonant phonon designs than for chirped superlattice and bound-to-continuum terahertz QCLs, because of the large ( $\sim 36$  meV) energy separation between the injector states and the lower radiative state. The second effect, which is common to all terahertz QCLs, is the reduction of the upper state scattering time  $\tau_{32}$  with increasing temperature due to thermally activated LO-phonon scattering. As described in Sec. 2.4, LO-phonon scattering for low electron temperatures is suppressed between the radiative states, but becomes the dominant lifetime limiting mechanism as electrons gain in-plane momentum. Using Eq. (2.49), we can write  $\tau_{32}$  as

$$\tau_{32}^{-1} = W_{32} = W_{32,0} + W_{32}^{(hot)} \exp\left(\frac{E_{32} - E_{LO}}{k_B T_e}\right), \quad (7.13)$$

where  $W_{32,0}$  is the  $3 \rightarrow 2$  rate due to electron-electron scattering at low temperatures, and  $W_{32}^{(hot)}$  is the raw scattering rate of an electron with sufficient in-plane energy to emit an LO-phonon for intersubband relaxation. Note that the temperature in (7.13) is the electron gas temperature  $T_e$ . As discussed in Sec. 2.4, at low lattice temperatures  $T_e$  will remain substantially elevated in an operating device. This is a result of the fact that LO-phonon emission is the only efficient cooling mechanism, so the electron gas will heat up until a sufficient number of optical phonons are emitted to balance the input power. Monte Carlo simulations of FL175C indicate that with a lattice temperature of 10 K,  $T_e$  may be 70–100 K hotter [84]. However, as the lattice temperature increases  $T_e$  will eventually track  $T$ , albeit with some offset ( $\sim 50$  K). Thus at elevated lattice temperatures ( $T > 70$  K), it is legitimate to approximate  $T_e \approx T + T_{off}$  in Eq. (7.13).

We can now explicitly describe the temperature dependence  $J_{th}$ :

$$J_{th}(T) \propto \frac{g_{th} \Delta \nu}{f_{32}} \left[ \frac{\tau_{3,para}^{-1} + W_{32,0} + W_{32}^{(hot)} \exp\left(\frac{E_{32} - E_{LO}}{k_B T_e}\right)}{1 - \tau_2 \left( W_{32,0} + W_{32}^{(hot)} \exp\left(\frac{E_{32} - E_{LO}}{k_B T_e}\right) \right)} \right] \quad (7.14)$$

For the moment, we will neglect the temperature dependence of  $g_{th}$  and  $\Delta \nu$ . At high temperatures, the thermally activated scattering dominates the lifetime and we

obtain

$$J_{th}(T) \propto \frac{g_{th}\Delta\nu}{f_{32}} \left[ \frac{W_{32}^{(hot)} \exp\left(\frac{E_{32}-E_{LO}}{k_B T_e}\right)}{1 - \tau_2 W_{32}^{(hot)} \exp\left(\frac{E_{32}-E_{LO}}{k_B T_e}\right)} \right] \quad (7.15)$$

While lasers with smaller photon energies  $E_{32}$  will naturally be more thermally robust, the best way to extend terahertz QCL operation for a given frequency is to use the fastest possible depopulation  $\tau_2$ , and minimize the raw scattering rate  $W_{32}^{(hot)}$ . While minimizing  $\tau_2$  is always desirable, in general reducing  $W_{32}^{(hot)}$  is accomplished by making the radiative transition more diagonal, and is accompanied by a reduction of the oscillator strength  $f_{32}$ . However, if one assumes that for a given design the reduction in  $f_{32}$  is proportional to a reduction in the rates  $\tau_{3,para}^{-1}$ ,  $W_{32,0}$ , and  $W_{32}^{(hot)}$ , examination of Eq. (7.12) reveals a net improvement in gain, since the term  $f_{32}\tau_3$  will remain approximately constant, while a slower time  $\tau_{32}$  is obtained. This should result in a lower  $J_{th}$  at low as well as high temperatures. This is illustrated in Fig. 7-23 for several parameter sets that follow this scaling. In fact, such an approximation is only broadly true, and should be used with caution, since the details of a given active region design are important. Using a more diagonal radiative transition should also improve the injection efficiency, since the lower state will be spatially separated from the injector states. The benefit of a more diagonal radiative transition is born out by the improved performance of bound-to-continuum designs over chirped superlattice designs. However, using a more diagonal transition has the potential to increase the linewidth  $\Delta\nu$  and reduce the gain. This can occur homogeneously due to the increase in interface roughness scattering associated with interwell transitions, or inhomogeneously due to the stronger dependence of the photon energy on applied bias. Hence there may be a lower limit to how diagonal the radiative transition should be made. This is especially true if the radiative states are delocalized over many wells, where dephasing scattering might even degrade their coherence, and reduce the oscillator strength.

In summary, to maximize operating temperatures for terahertz QCLs, several strategies should be employed for active region design: First and most importantly, the depopulation should be made as fast and as temperature robust as possible. If

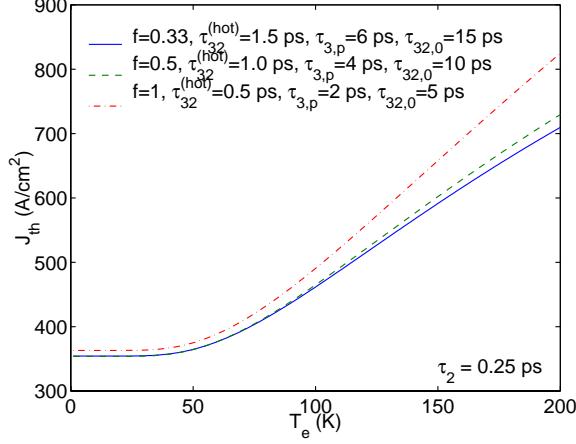


Figure 7-23: Calculated threshold current density using Eq. (7.14) for various parameters. The radiative energy was  $E_{32} = 14$  meV,  $\tau_2 = 0.25$  ps,  $g_{th} = 30$  cm $^{-1}$ , and  $\Delta\nu = 1$  THz.

$\tau_2 < \tau_{32}^{(hot)}$ , then some degree of population inversion will always be maintained, even at very high temperatures. Second, the radiative transition should be made somewhat diagonal to maximize  $\tau_{32}^{(hot)}$  and obtain a longer upper state lifetime. The reduction in the oscillator strength  $f_{32}$  should be balanced by a corresponding increase in  $\tau_3$ . A diagonal transition has the added benefit of improving injection efficiency. Third, the injector design and doping level should be optimized to maximize the selectivity of injection into the upper state and to enable the largest possible current density at design bias.

## Waveguide

The increase in waveguide loss  $\alpha_w$  with temperature also contributes to the increase in  $J_{th}$ . The absorption due to phonon effects will increase with temperature as the *Reststrahlen* resonance broadens, and may provide several times more loss, as shown in Fig. 4-5. Aside from this, the increase in loss comes from free carrier absorption and the reduction of the Drude scattering times for metals and semiconductors.

Metal cladding layers are potentially sensitive to temperature, i.e. between 80 and 300 K, the dc electrical conductivity of gold drops by a factor of  $\sim 5$ :  $\tau_{dc} \approx 0.12$  ps at 80 K, and  $\tau_{dc} \approx 0.027$  ps at 300 K. However, optical measurements of the Drude

Table 7.8: Drude relaxation times and waveguide losses for SI-surface-plasmon waveguide and metal-metal waveguide calculated for FL175C using the Drude times in this chapter, which roughly corresponds to 77 K, and at 300 K.

	$\tau_{Au}$	$\tau_{AR}$	$\tau_{n+}$	$\alpha_{mm}/\Gamma$	$\alpha_{SI}/\Gamma$
This chapter ( $\sim 77$ K)	0.05 ps	0.5 ps	0.1 ps	$16.1 \text{ cm}^{-1}$	$25.9 \text{ cm}^{-1}$
300 K	0.024 ps	0.23 ps	0.08 ps	$32.1 \text{ cm}^{-1}$	$42.8 \text{ cm}^{-1}$

scattering time in the mid-infrared showed a much smaller Drude relaxation time  $\tau$  at 77 K ( $\sim 0.04$  ps) [131], which reduces to  $\tau = 0.025$  ps at room temperature [130]. The low temperature conductivity of metal depends sensitively on the impurity level, which is relatively unknown for our samples. For most waveguide simulations in this thesis,  $\tau = 0.05$  ps has been used for Au to roughly correspond with the optical Drude model fitting at 77 K. Actual loss from metal may increase by as much as a factor of five, but recalculating with  $\tau = 0.025$  ps will only bring a factor of two change over waveguide loss calculated in this chapter.

As for GaAs, since the mobility in the heavily doped contact layers is dominated by impurity scattering, there is hardly any temperature dependence of their loss. Free carrier absorption in the moderately doped active region will increase however, although it is difficult to characterize because of the quasi-two-dimensional carrier confinement (see Sec. 4.3.3). However, the extent of two-dimensional nature of the active region electron gas is certain to diminish with increasing temperature. The mobility of bulk GaAs for  $n$ -type doping in the  $10^{15} - 10^{16} \text{ cm}^{-3}$  range drops by approximately a factor of four, with  $\tau \approx 1$  ps at 77 K and  $\tau \approx 0.23$  ps at 300 K [128]. However, in this chapter  $\tau = 0.5$  ps has been used for the active region Drude relaxation time.

To obtain a rough estimate of the temperature dependence of waveguide loss, the losses for typical SI-surface-plasmon and metal-metal guides (FL175C) were calculated, using 300 K Drude times and the Drude times used in this chapter, which roughly correspond to 77 K or worse. The relaxation times and calculated losses are shown in Table 7.8. The confinement factors are the same for both temperatures:

$\Gamma = 0.29$  for the SI waveguide and  $\Gamma = 0.98$  for the metal waveguide. The calculated losses approximately double. The SI-surface-plasmon waveguide observes a smaller relative increase in loss, since most of its loss is due to the heavily doped contact layer, which is temperature insensitive. Despite this, the metal-metal waveguide is still calculated to perform better at room temperature. Also, due to its smaller confinement factor, it is most vulnerable to the increase in phonon loss in GaAs. The largest uncertainty in loss is in the active region. Also, these figures will vary for lasers at different frequencies, since the active region free carrier loss may make up a larger or smaller fraction of the total loss. However, while the loss may as much as double at high temperatures, it remains to be seen how severely this will impact lasing performance.

## Heat removal

Obtaining high duty cycle and cw operation of terahertz lasers at elevated temperatures is an exercise in thermal management. High temperature operation is made more difficult in quantum cascade devices due to the reduced thermal conductivity of the active region. Superlattices and multiple quantum well regions typically display lower thermal conductivity  $\kappa$  perpendicular to the epitaxial layers than bulk alloys of the same average composition [176]. This is due both to increased interface scattering of phonons and a flattening of the dispersion relations which reduces phonon group velocity. The actual value of the active region thermal conductivity  $\kappa_{AR}$  for our structure is unknown. Recent measurements of mid-infrared GaAs/Al<sub>0.33</sub>Ga<sub>0.67</sub>As QCL facet temperatures using a microprobe photoluminescence technique have estimated values of  $\kappa_{AR} \approx 0.05$  W/(cm K) at approximately 80 K [177]. Transient photoluminescence was used at lower temperatures to obtain  $\kappa_{AR} \approx 1.3$  W/(cm K) at 30 K, which decreased to approximately 0.2 W/(cm K) at 100 K [178]. These values were approximately 10 times less than those for comparable bulk alloys. Using a spatially resolved interferometric technique, the thermal conductivity for a GaAs-based QCL at room temperature was measured to be approximately 0.015 W/(cm K) perpendicular to the layers, and 0.25 W/(cm K) in the plane of the layers [179].

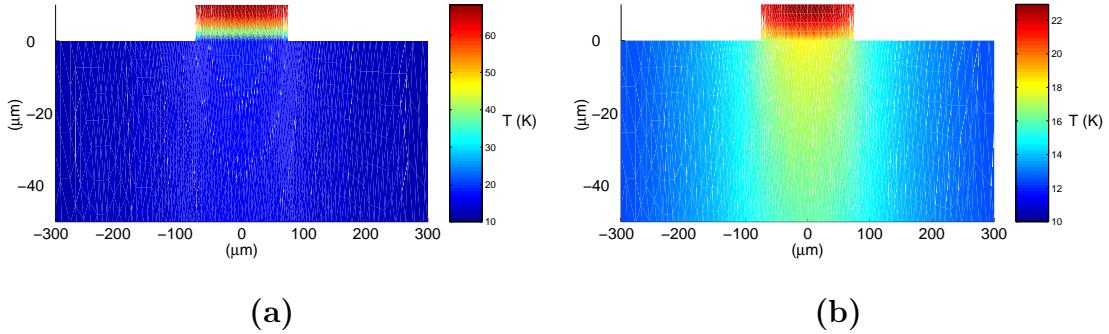


Figure 7-24: Thermal simulations of 150- $\mu\text{m}$ -wide 10- $\mu\text{m}$ -high ridge structure using an active region thermal conductivity of (a)  $\kappa_{AR} = 0.1 \text{ W}/(\text{cm K})$  and (b)  $\kappa_{AR} = 1.0 \text{ W}/(\text{cm K})$ . A heat sink temperature of 10 K is used, the substrate thickness is 600  $\mu\text{m}$  with  $\kappa_{sub} = 10 \text{ W}/(\text{cm K})$ , and the power density is  $10^7 \text{ W cm}^{-3}$  in the active region.

For terahertz QCLs, the reduction in  $\kappa_{AR}$  may be somewhat less severe than these reported results, since the average aluminum content is lower. For example, 15.6% of FL175C is composed of  $\text{Al}_{0.15}\text{Ga}_{0.85}\text{As}$  barriers, so that the equivalent bulk alloy is only 2.3% aluminum, compared to 17% for a mid-IR QCL [178].

The dc temperature profile is modeled for FL175C for a 150- $\mu\text{m}$ -wide ridge in Fig. 7-24 using both  $\kappa_{AR} = 0.1 \text{ W}/(\text{cm K})$  and  $\kappa_{AR} = 1 \text{ W}/(\text{cm K})$ . A heat sink temperature of 10 K and  $\kappa_{sub} = 10 \text{ W}/(\text{cm K})$  is used for the substrate, which is the value for SI-GaAs at 50 K. For the case with lower  $\kappa_{AR}$ , nearly all the temperature drop takes place inside the active region, and the maximum temperature is close to 70 K. This would explain the lack of cw lasing for FL175C. For the case with larger  $\kappa_{AR}$ , the maximum active region temperature is only 22 K, which is only 12 K above the substrate temperature. However, for such a scenario, the substrate resistance is comparable to that of the active region, and thinning the substrate may benefit heat sinking.

In fact, during our experiments, the thermal resistance of the copper mount and indium foil elevated the temperature significantly above the LHe cryogen temperature. For example, for the 1.19-mm-long, 150- $\mu\text{m}$ -wide FL175C device shown in Fig. 7-5, at a 50% duty cycle, the measured temperature from the sensor was 15 K. Since the sensor was mounted below the actual chip carrier, the actual heat sink

temperature was probably even higher. Thus it is not unreasonable to assume a heat sink temperature of 20–30 K during cw testing of FL175C, in which case only an additional 20–30 K of heating is enough to prevent lasing, which is roughly consistent with FL175C performance.

In addition to simply reducing the available gain, the increase in temperature in a ridge structure produces thermoelastic stress that can degrade performance and even lead to catastrophic device failure. This strain is most severe around the edges of a ridge. Indeed, the FL175C device mentioned above failed under dc bias due to a stress related failure in the coating. Similar devices were observed to completely “explode” due to stress buildup, leaving only a ridge shaped crater. Recent experiments by Spagnolo *et al.* on strain buildup in mid-infrared QCLs suggests device damage may occur for power densities greater than  $5 \text{ kW/cm}^2$  [180]. At 5 K, the threshold power density for FL175C is at best  $800 \text{ A/cm}^2 \times 12 \text{ V} = 9.6 \text{ kW/cm}^2$ .

Clearly the large dc power dissipated by our first generation of devices is an impediment to cw operation. Therefore, reduction of  $J_{th}$  is a priority, which is most easily done by substantially reducing the parasitic current channel. Even the modest reduction of  $J_{th}$  to  $700 \text{ A/cm}^2$  in FL178C-M1 was enough to allow cw lasing, if only for a few seconds. The threshold voltage is not so easily reduced, since devices with an explicit LO-phonon relaxation channel will always require a relatively large design bias. However, FL175C for example has a design bias of  $\sim 13.5 \text{ V}$ , which is equivalent to  $77 \text{ mV/module}$ . Since the calculated design bias is  $\sim 64 \text{ mV/module}$ , and emission of an LO-phonon and a 14 meV photon only accounts for  $50 \text{ mV/module}$ , reduction of parasitic voltage drops and trimming of the design bias may be possible.

Advanced fabrication techniques to facilitate lateral heat removal from the laser ridge such as epitaxial regrowth, ion implantation isolation [48], or thick metallic plating [139] may be useful. Lateral heat removal is especially promising for QCL structures, as the in-plane thermal conductivity is substantially larger than  $\kappa$  perpendicular to the layers [179]. Due to the need for lateral contacts, implementation of these techniques is more difficult for SI-surface plasmon waveguides. However, metal-metal waveguides are contacted on the back of the substrate, which allows

more straightforward implementation of these techniques, as well as the use of epi-layer down mounting. Furthermore, above 100 K, the thermal conductivity of Au is better than that of GaAs, so that metal waveguides could potentially enhance heat removal. Copper has a better thermal conductivity than GaAs at all temperatures, so if Cu-Cu bonding is used [181] even better heat sinking is possible. Also, the use of metal-metal waveguides more easily allows the growth of thinner active regions, which will help reduce power dissipation. In general, obtaining cw operation at temperatures comparable with  $T_{max}$  for pulsed operation should be easier in the terahertz than for mid-IR GaAs-based QCLs. Terahertz QCLs have much lower threshold current densities: as low as 95 A/cm<sup>2</sup> [182] compared to 1.6 kA/cm<sup>2</sup> for a mid-IR GaAs QCL at 77 K [48]. Also, the average aluminum content of terahertz QCLs is much lower, which will likely result in higher thermal conductivity of the active region.

## 7.10 Conclusions and Discussion

In this thesis, I have described the study and development of terahertz quantum cascade lasers, successfully culminating in the fabrication of several lasing devices. The research conducted over the course of this thesis was wide ranging, and primarily consisted of device theory and design, fabrication of structures, and experimental testing and characterization. A wide variety of quantum cascade structures were designed, modeled, and simulated in order to obtain laser structures as well as to study intersubband transport, terahertz emission, and various terahertz waveguides. Over the course of this effort, many simulations and design tools were written, two of the most significant being a one-dimensional complex waveguide mode solver and a complex (interface and confined) phonon mode intersubband scattering simulation. Fabrication procedures were developed and implemented for 10- $\mu$ m-deep wet and dry etch laser structures, including the development of new methods for wafer bonding and the fabrication of metal-metal waveguides. Finally, many different terahertz emission structures were tested using transport, Fourier transform emission spectroscopy, and magnetotunneling spectroscopy methods performed at cryogenic temperatures.

Some of the major accomplishments and experimental observations that contributed to our understanding include the following.

- The first narrow linewidth intersubband emission from a non-intrawell quantum cascade structure was observed at both 5 K and 77 K (L30). Although it did not produce gain, this active region was the precursor of the resonant phonon designs, and the temperature robustness of the emission foreshadowed the viability of terahertz QCLs above 77 K. The higher power and narrower linewidth of this device compared to a similar device with a diagonal transition suggested that vertical transitions with higher oscillator strengths were more suitable for successful lasers.
- Magnetotunneling spectroscopy was used to resolve an anticrossing gap of approximately 2 meV between two lower states (T65). This provided direct evidence as to the coherent nature of the tunneling interaction between the two states, and the applicability of the delocalized anticrossed picture for final states with small energy separations, or at least for those that have long lifetimes. Furthermore, this anticrossing picture was roughly confirmed by spectral measurements, where the transfer of oscillator strength to the  $3 \rightarrow 2$  transition was observed, as predicted by wavefunction parity arguments.
- The incoherent sequential nature of tunneling through the injection barrier into the excited state was experimentally confirmed via the series of measurements on the T-series devices. The increase in peak current density and excited state population as the barrier was thinned was a clear signature of this phenomenon. This demonstrated the inapplicability of the semiclassical delocalized transport model for injection, and pointed to a need for a description of injection that accounted for dephasing and localization. Such a description using a tight-binding density matrix description, first given by Kazarinov and Suris [31] and Sirtori *et al.* [104] was extensively discussed in Sec. 2.5.2 and was shown to give a qualitatively good description of parasitic current channels in Sec. 7.7.

- Extensive calculations of electron-LO-phonon intersubband scattering were performed taking account of the heterostructure and its effect on the complex phonon spectra. These simulations clarified the role of higher energy phonon modes associated with the barrier material, and rebutted suggestions that terahertz QCLs should be designed with 47 meV subband separation instead of 36 meV in order to take advantage of the extra modes for scattering. The effect of raising the depopulation subband separation was not found to be beneficial for structures with  $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$  barriers. The effect of higher energy modes is expected to be even less in the current group of designs where  $\text{Al}_{0.15}\text{Ga}_{0.85}\text{As}$  barriers are used, and hence the 36 meV design rule has been retained.
- Extensive fabrication and testing of metal-metal waveguide devices was performed, allowing refinement of fabrication techniques and characterization of device yield. A new recipe featuring In-Au reactive bonding was developed, and was shown to be superior in yield and mechanical strength, although the thermal conductance of the bonding layer needs to be improved. Prior to lasing, the low loss of the metal-metal waveguide was shown by the larger spontaneous emission power that was available compared to a plasmon waveguide. Also, a lateral Fabry-Pérot mode effect was observed in the spectra, which is consistent with low values of waveguide loss.
- Terahertz quantum cascade lasers have been demonstrated using resonant tunneling of the lower radiative state combined with direct LO-phonon scattering for selective depopulation. This depopulation method is fundamentally different in character than the miniband depopulation used for the prior chirped superlattice design. LO-phonon scattering provides an efficient and temperature insensitive depopulation mechanism, and an intrinsic protection from thermal backfilling effects due to the large ( $\sim 36$  meV) energy separation between the lower radiative state and the injector. Both properties are important in allowing higher temperature operation at longer wavelengths. Lasing at 3.4 THz ( $\lambda \approx 88 \mu\text{m}$ ) was observed in a device with a semi-insulating-surface-plasmon

waveguide in pulsed mode up to 87 K. The peak pulsed power was 14 mW at 5 K. A maximum duty cycle of 60% was obtained. Additionally, the first terahertz QCL using a metal-metal waveguide for mode confinement was been demonstrated at 3.0 THz ( $\lambda \approx 100 \mu\text{m}$ ). Such a waveguide has a confinement factor close to unity and is expected to be advantageous over a SI-surface-plasmon guide due to the latter's low confinement factor. This was borne out by the fact that the metal-metal device lased in pulsed mode up to 77 K, whereas the same device fabricated with a SI-surface-plasmon guide failed to lase. The highest lasing temperature to date is 137 K in pulsed mode from a modification of the original design fabricated with a metal-metal waveguide.

In less than two years since the first terahertz QCL was demonstrated, major advances have been observed, but there is every reason to believe terahertz QCLs will still see substantial improvement. At this time, the maximum operating temperature is 137 K in pulsed mode from a resonant-phonon design (Sec. 7.6), and is 68 K in cw mode from a bound-to-continuum design [45]. It is not yet clear what the maximum operating temperature for terahertz QCLs is, and if room temperature operation can be obtained. The reduction of the upper state lifetime by thermally activated LO-phonon scattering is the primarily limitation, which causes a rapid increase of  $J_{th}$  above 100 K. Even in the best devices, the slope efficiencies are far below the theoretical limit, which suggests there is much room for improvement. Even for FL178C-M1, which operated up to 137 K, it appears that many of the modules were not biased properly, and thus a great deal of potential gain was wasted. Further benefits will result from optimizing the selectivity of injection.

Since the upper state lifetime drops strongly with temperature, maintaining extremely fast depopulation is of the utmost importance. In the opinion of this author, LO-phonon depopulation is the best means, which gives resonant phonon designs an advantage over other current designs. However, at some high temperatures the resonant tunneling interaction may be impeded by dephasing scattering. Also, because the lower state is part of an anticrossed doublet, each state has twice the relaxation time of a true intrawell LO-phonon relaxation event. For these reasons, if an excited

state in a wide well is used as the lower radiative state, the LO-phonon scattering will be intrawell and extremely fast, as well as insensitive to dephasing. However, the upper state lifetime will be short in such a structure, and reaching threshold at any temperature may be difficult.

Additional challenges are present for high temperature cw operation. Aside from minimizing the dissipated power, the major challenge is to remove heat from the active region. Techniques such as epi-down mounting and lateral heat removal from the ridge will play a role. Also, the larger thermal conductivity of metal at high temperatures may grant an advantage to metal-metal waveguides in terms of heat sinking. Indeed, if the thermal conductivity of the bonding layer can be improved, metal-metal waveguides may prove to be critical for higher temperature operation, not only because of their low loss, but because their high confinement factor lessens the need for long, thick, and wide ridges which dissipate more power, as well as the need for failure-prone facet coatings.

In this thesis, I have described the theory and design, fabrication, and characterization of terahertz quantum cascade lasers. Extensive modeling of transport and optical properties was performed for both the multiple-quantum-well intersubband gain medium, and the various waveguides. Fabrication methods were detailed, including those for metallic wafer bonding processes used to fabricate low-loss metal-metal waveguides. A large variety of quantum cascade structures were characterized for their potential as lasers as well as to gain fundamental insight into design. Finally, I have presented the results of a new class of terahertz quantum cascade lasers based on resonant phonon depopulation, and discussed their prospects for improvement.

# Appendix A

## Notes on dephasing and homogeneous broadening

The role of population relaxation (lifetime broadening) and dephasing (collisional broadening) on homogeneous optical transitions and other coherent systems warrants review, since the literature involving intersubband transitions is sometimes unclear due to the large number of conventions and notations used. To illustrate the role these processes play, I will briefly review the classical oscillator model for atomic transitions. For a more detailed treatment, see Chapters 2 and 3 in Siegman [183].

Consider an ensemble of classical electron oscillators with resonance frequency  $\omega_0$ . The equation of motion for the dipole moment  $\mu = -ex(t)$  of a single oscillator in the presence of an electric field is given by

$$\frac{d^2}{dt^2}\mu_x(t) + \gamma\frac{d}{dt}\mu_x(t) + \omega_0^2\mu_x(t) = (e^2/m)E_x(t), \quad (\text{A.1})$$

where  $E_x(t)$  is the driving field,  $m$  is the mass,  $e$  is the elementary charge, and  $\gamma$  is the energy decay rate of the oscillator. In a quantum description,  $\gamma$  would be equivalent to sum of the energy (population) relaxation rates for the upper and lower transition levels. In the absence of a driving field, each oscillator  $i$  will have the solution

$$\mu_{x,i} = |\mu_{x0,i}|e^{-(\gamma/2)t}e^{-i\omega_0 t}e^{i\phi_i}, \quad (\text{A.2})$$

where  $\mu_{x0,i}$  is the polarization at time  $t = 0$ , and  $\phi_i$  is the phase of the  $i$ th oscillator. The macroscopic polarization produced by the ensemble is given by summing over the ensemble:

$$p_x(t) = \sum_i \mu_{x,i}(t). \quad (\text{A.3})$$

Assume that the ensemble of  $N_0$  oscillators is in phase at time  $t = 0$ , and all add coherently to produce a large macroscopic polarization. If collisions or other phase-randomizing processes occur with a characteristic time  $T_2^*$  (known as the pure dephasing time), the polarization will decay, since the polarization produced by those oscillators that have dephased will average out to zero ( $\langle \mu_{dephased} \rangle = 0$ ). So only the number  $N(t)$  that haven't dephased will contribute to the polarization. Since  $N(t)$  will decay according to

$$N(t) = N_0 e^{-t/T_2^*}, \quad (\text{A.4})$$

the polarization likewise decays as

$$p_x(t) = N(t)\mu_x(t) = p_{x0}e^{-t/T_2^*}e^{-(\gamma/2)t}e^{-i\omega_0 t} = p_{x0}e^{-(\gamma/2+1/T_2^*)t}e^{-i\omega_0 t}. \quad (\text{A.5})$$

The effect of dephasing is to replace the single oscillator decay rate with a decay rate for the coherent polarization, i.e.  $\gamma/2 \rightarrow (\gamma/2 + 1/T_2^*)$ .

As a result of this substitution, the equation of motion for the polarization is written

$$\frac{d^2}{dt^2}p_x(t) + \left(\gamma + \frac{2}{T_2^*}\right) \frac{d}{dt}p_x(t) + \omega_0^2 p_x(t) = (e^2/m)E_x(t). \quad (\text{A.6})$$

Assuming a harmonic time dependence  $e^{-i\omega t}$  for  $p_x(t)$  and  $E_x(t)$ , we obtain the complex susceptibility for the oscillator:

$$\chi(\omega) = \frac{P(\omega)}{\epsilon_0 E(\omega)} = \frac{Ne^2}{\epsilon_0 m} \frac{1}{(\omega_0^2 - \omega^2) - i\omega\Delta\omega_0}, \quad (\text{A.7})$$

where  $\Delta\omega_0 = \gamma + 2/T_2^*$ . By making the resonance approximation  $\omega^2 - \omega_0^2 \approx 2\omega_0(\omega -$

$\omega_0$ ), we can obtain the imaginary part of the susceptibility

$$\chi''(\omega) = -\frac{Ne^2}{4\epsilon_0 m} \frac{\Delta\omega_0}{(\omega - \omega_0)^2 + (\Delta\omega_0/2)^2} = -\frac{Ne^2}{4\epsilon_0 m} \frac{(\Delta\nu_0/2\pi)}{(\nu - \nu_0)^2 + (\Delta\nu_0/2)^2}, \quad (\text{A.8})$$

which has a Lorentzian lineshape with a full width half maximum linewidth  $\Delta\omega_0$  ( $\Delta\nu_0$ ). We note that the contribution of  $\gamma$  to the linewidth is half that of the pure dephasing rate. This is a result of the fact that  $T_2^{*-1}$  is an amplitude decay rate for the polarization  $p_x(t)$ , and  $\gamma$  is an energy decay rate that reduces the intensity  $|p_x(t)|^2$ . Essentially, lifetime broadening contributes half as much to the homogeneous broadening as does pure dephasing. In a quantum picture,  $\gamma$  is replaced by the relaxation rates of the upper and lower states of the transition  $\gamma_u$  and  $\gamma_l$ . Thus the linewidth can be written

$$\Delta\omega_0 = \gamma_u + \gamma_l + \frac{2}{T_2^*} = \frac{2}{T_2}, \quad (\text{A.9})$$

where the right side of the equation represents the combination of all terms to form a total dephasing time  $T_2$  which determines the linewidth. The notation for these terms varies considerably in the literature, and it is not always clear whether a dephasing time or pure dephasing time is being referenced. Care should be taken when interpreting equations and interpretations of linewidths.

A quantum mechanical derivation of the complex susceptibility can be performed using density matrix formalism; such a derivation is found in Yariv [77]. However, in these cases the dephasing and population relaxation is described phenomenologically, rather than being derived from first principles. For a full quantum mechanical derivation of dephasing, see Appendix 6A of Mukamel [105]. Dephasing with respect to resonant tunneling is considered quantum mechanically in Sec. 2.5.2.



# Appendix B

## Processing recipes

### B.1 Dry etching

Dry etching was performed in the Plasmaquest using the GAASETCH.RCP recipe, listed below. Running etch cleans and conditioning the chamber before performing the etch is essential, since the Plasmaquest is frequently contaminated.

**BCl<sub>3</sub>** 15 sccm

**N<sub>2</sub>** 5 sccm

**pressure** 5 mTorr

**ECR** 600 W

**RF** 15 W

**Temperature** 80 C



# Appendix C

## QCL rate equations

In this section I will derive the rate equations for a generic quantum cascade laser taking account of non-unity injection efficiency. A schematic diagram of the system is shown in Fig. C-1. Electrons are injected into the upper radiative state  $|3\rangle$  at some rate  $P\eta$ , and into the lower radiative state  $|2\rangle$  with rate  $P(1 - \eta)$ . The various non-radiative relaxation rates are given by  $\tau_2$  and  $\tau_3$ , where the latter includes a parasitic channel  $\tau_{3,para}$  as well as the direct  $|3\rangle \rightarrow |2\rangle$  channel  $\tau_{32}$ . The total spontaneous emission time into all modes is given by  $\tau_{sp}$ , and the spontaneous emission time into the laser mode of interest is  $\tau'_{sp}$ . The total stimulated emission rate is given by  $\tau_{st}^{-1} = m\tau'_{sp}$ , as given in Eq. 2.22. The photon population density of that mode is given by  $m$  and the electron population densities of each level are  $n_3$  and  $n_2$ .

The notation used in this derivation is summarized below.

$n_3, n_2$  3D population densities of level 3, 2.

$m$  Photon population density of mode.

$\beta$  Number of modes.

$\tau_{sp}$  Spontaneous emission time (into all modes).

$\tau'_{sp}$  Spontaneous emission time into single mode of interest. ( $\frac{1}{\tau_{sp}} \approx \frac{\beta}{\tau'_{sp}}$ ).

$\tau_p$  Photon lifetime ( $\tau_p^{-1} = v_g\alpha$ )

$N_{mod}$  Number of modules.

$V$  Volume of one module.

$$P = \frac{J}{et} \quad \tau_3^{-1} = \tau_{32}^{-1} + \tau_{3,para}^{-1}$$

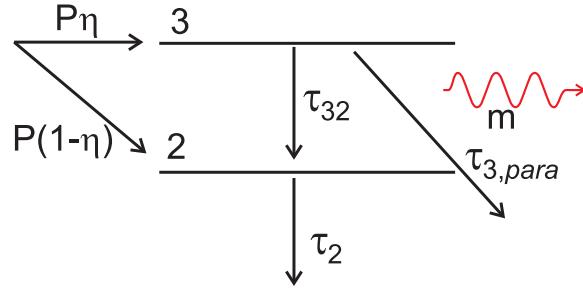


Figure C-1: Schematic for coupled carrier and photon population model for quantum cascade laser with non-unity injection efficiency  $\eta$ .

$V_p$  Volume of photon cavity.

$\Gamma$  active region confinement factor ( $\Gamma = \frac{V}{V_p} N_{mod}$ ).

$A$  Area of device.

$t$  Thickness of one module ( $At = V$ ).

$J$  Current Density ( $J = \frac{I}{A}$ )

$P$  Pumping rate ( $P = \frac{J}{et}$ )

Rate equations can be written for the populations of each level, as well as that of the optical “lasing” mode. These can be first written as number rate equations.

$$V \frac{dn_3}{dt} = \frac{\eta I}{e} - \frac{(n_3 - n_2)V^2 m}{\tau'_{sp}} - \frac{n_3 V}{\tau_{sp}} - \frac{n_3 V}{\tau_3} \quad (C.1)$$

$$V \frac{dn_2}{dt} = \frac{(1-\eta)I}{e} + \frac{(n_3 - n_2)V^2 m}{\tau'_{sp}} + \frac{n_3 V}{\tau_{sp}} + \frac{n_3 V}{\tau_{32}} - \frac{n_2 V}{\tau_2} \quad (C.2)$$

$$V_p \frac{dm}{dt} = \frac{(n_3 - n_2)V^2 m N_{mod}}{\tau'_{sp}} + \frac{n_3 V N_{mod}}{\tau'_{sp}} + \frac{n_3 V N_{mod}}{\tau'_{sp}} - \frac{m V_p}{\tau_p} \quad (C.3)$$

From here, it is convenient to rewrite these as density rate equations.

$$\frac{dn_3}{dt} = \eta P - \frac{(n_3 - n_2)V m}{\tau'_{sp}} - \frac{n_3}{\tau_{sp}} - \frac{n_3}{\tau_3} \quad (C.4)$$

$$\frac{dn_2}{dt} = (1-\eta)P + \frac{(n_3 - n_2)V m}{\tau'_{sp}} + \frac{n_3}{\tau_{sp}} + \frac{n_3}{\tau_{32}} - \frac{n_2}{\tau_2} \quad (C.5)$$

$$\frac{dm}{dt} = \frac{(n_3 - n_2)Vm\Gamma}{\tau'_{sp}} + \frac{n_3\Gamma}{\tau'_{sp}} - \frac{m}{\tau_p} \quad (\text{C.6})$$

Next, we will use these rate equations to analyze QCL behavior in the steady state. We obtain:

$$\eta P = \frac{(n_3 - n_2)Vm}{\tau'_{sp}} + \frac{n_3}{\tau_3} \quad (\text{C.7})$$

$$(1 - \eta)P = -\frac{(n_3 - n_2)Vm}{\tau'_{sp}} - \frac{n_3}{\tau_{32}} + \frac{n_2}{\tau_2} \quad (\text{C.8})$$

$$\frac{m}{\tau_p} = \frac{(n_3 - n_2)Vm\Gamma}{\tau'_{sp}} + \frac{n_3\Gamma}{\tau'_{sp}}, \quad (\text{C.9})$$

where the spontaneous emission term was dropped in Eqs. (C.7) and (C.8), since  $\tau_{32} \ll \tau_{sp}$  in the terahertz.

## C.1 Above threshold

Above lasing threshold, there are a large number of photons in the cavity and  $Vm \gg 1$ . Therefore the spontaneous emission term in (C.9) becomes negligible and we have

$$\frac{(n_3 - n_2)V\Gamma}{\tau'_{sp}} = \frac{1}{\tau_p}. \quad (\text{C.10})$$

Since the photon lifetime  $\tau_p = (v_g\alpha)^{-1}$  where  $\alpha$  is the total loss and  $v_g$  is the group velocity, this is simply a restatement of the threshold gain condition of Eq. (4.1). Above threshold, the population inversion is pinned at this level.

The next step is to use Eqs. (C.7) and (C.8) to solve for the population inversion  $(n_3 - n_2)$ . After some algebra we obtain

$$n_3 = \frac{P \left( \frac{\eta}{\tau_2} + \frac{Vm}{\tau'_{sp}} \right)}{\frac{Vm}{\tau'_{sp}\tau_3} + \frac{Vm}{\tau'_{sp}\tau_2} + \frac{1}{\tau_3\tau_2} - \frac{Vm}{\tau'_{sp}\tau_{32}}}. \quad (\text{C.11})$$

Using this, the population inversion can be written as

$$n_3 - n_2 = P \left[ \frac{-\frac{1}{\tau_3} + \frac{\eta}{\tau_3} + \frac{\eta}{\tau_2} - \frac{\eta}{\tau_{32}}}{\frac{V_m}{\tau'_{sp}} \left( \frac{1}{\tau_3} + \frac{1}{\tau_2} - \frac{1}{\tau_{32}} \right) + \frac{1}{\tau_{32}}} \right]. \quad (\text{C.12})$$

If this expression is inserted back into Eq. (C.10) we can obtain an expression for the photon density

$$m = \left( \frac{\tau'_{sp}}{V} \right) \frac{\tau_p \frac{PV\Gamma}{\tau'_{sp}} \left( -\frac{1}{\tau_3} + \frac{\eta}{\tau_3} + \frac{\eta}{\tau_2} - \frac{\eta}{\tau_{32}} \right) - \frac{1}{\tau_3 \tau_2}}{\frac{1}{\tau_3} + \frac{1}{\tau_2} - \frac{1}{\tau_{32}}} \quad (\text{C.13})$$

$$= \frac{\tau_p \Gamma \left( \eta \tau_3 \left( 1 - \frac{\tau_2}{\tau_{32}} \right) - (1 - \eta) \tau_2 \right)}{\tau_2 + \tau_3 \left( 1 - \frac{\tau_2}{\tau_{32}} \right)} (P - P_{th}). \quad (\text{C.14})$$

In this expression the threshold pumping  $P_{th}$  has been defined as

$$P_{th} = \frac{\tau'_{sp}}{\tau_p V \Gamma} \frac{1}{\eta \tau_3 \left( 1 - \frac{\tau_2}{\tau_{32}} \right) - (1 - \eta) \tau_2}. \quad (\text{C.15})$$

$P_{th}$  is the pumping necessary to meet the threshold condition (C.10). We can relate the photon density  $m$  to the emitted light power  $L$  via the expression:

$$L = \frac{\hbar \omega V_p m}{\tau_m}, \quad (\text{C.16})$$

where  $\tau_m$  is the mirror photon escape time, defined by  $\tau_m^{-1} = \alpha_m v_g$ . Given that the pumping rate is defined by

$$P = \frac{J}{et} = \frac{I}{Aet}, \quad (\text{C.17})$$

we can write the slope efficiency as

$$\frac{dL}{dI} = \left( \frac{\hbar \omega V_p}{\tau_m Aet} \right) \frac{dm}{dP} \quad (\text{C.18})$$

$$= \frac{\hbar \omega}{e} N_{mod} \frac{\alpha_m}{\alpha_m + \alpha_w} \left[ \frac{\eta \tau_3 \left( 1 - \frac{\tau_2}{\tau_{32}} \right) - (1 - \eta) \tau_2}{\tau_2 + \tau_3 \left( 1 - \frac{\tau_2}{\tau_{32}} \right)} \right]. \quad (\text{C.19})$$

The quantity in brackets is a dimensionless quantity that takes the role of an internal efficiency  $\eta_i$ . It reflects that slope efficiency is reduced by possible non-unity injection

$\eta$  and the fact that the lower radiative state may not be empty (i.e.  $\tau_2 \neq 0$ ). Note that the expression given for this term in Köhler *et al.* [69] is incorrect—it ought to have an additional factor of  $\eta$  in front.

Note that the expression for slope efficiency includes the number of modules  $N_{mod}$ . This term allows for the achievement of differential quantum efficiencies greater than unity, which is an essential feature of cascade lasers. In such a device, a single injected electron can emit more than one photon. Examination of Eq. (C.19) makes clear that a large slope efficiency requires large  $\eta$  and small  $\tau_2/\tau_{32}$ . By making the radiative transition more diagonal, both of these parameters are optimized. In addition, a diagonal scheme is favorable for high temperature operation (see Sec. 7.9).

## C.2 Subthreshold

Below lasing threshold, there are very few photons in the mode so that  $Vm \ll 1$  and stimulated emission has no effect on transport and electron populations. Eq. (C.9) gives the photon density in the mode as simply

$$m = \frac{n_3 \Gamma \tau_p}{\tau'_{sp}}. \quad (\text{C.20})$$

The total spontaneous emission power will consist of emission into all the modes however, so we should use the total spontaneous emission time  $\tau_{sp}$  to obtain

$$L = \frac{\hbar \omega V_p \tau_p}{\tau_m} \frac{\Gamma n_3}{\tau_{sp}} \quad (\text{C.21})$$

$$= \eta \hbar \omega N_{mod} \frac{\alpha_m}{\alpha_m + \alpha_w} \frac{\tau_3}{\tau_{sp}} \frac{I}{e}. \quad (\text{C.22})$$

In other words, the subthreshold spontaneous emission is proportional to the upper state population  $n_3$ , and thereby to the upper state lifetime  $\tau_3$ . This makes the spontaneous emission power level a useful probe of level populations. Caution must be taken when examining  $L-V$  or  $L-I$  characteristics for to extract upper state populations or lifetimes however. Changes in bias can often result in changes in the injection

efficiency  $\eta$  as injection states may move into and out of alignment with both the upper radiative state as well as parasitic states. Hence, a sublinear  $L$ - $I$  characteristic could due to either a reduction in  $\tau_3$  or  $\eta$ .

One can likewise compute the subthreshold population inversion from Eqs. (C.7) and (C.8) in the limit  $Vm \ll 1$ . This gives the expression

$$n_3 - n_2 = P \left[ \eta \tau_3 \left( 1 - \frac{\tau_2}{\tau_{32}} \right) - (1 - \eta) \tau_2 \right] \quad (\text{C.23})$$

$$= \frac{J}{et} \left[ \eta \tau_3 \left( 1 - \frac{\tau_2}{\tau_{32}} \right) - (1 - \eta) \tau_2 \right]. \quad (\text{C.24})$$

This simply reflects the intuitive fact that it is necessary to have  $\tau_2 < \tau_{32}$  to obtain a population inversion. A long upper state lifetime  $\tau_3$  is also desirable, as is efficient injection ( $\eta = 1$ ).

# Bibliography

- [1] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho. Quantum cascade laser. *Science*, 264:553, 1994.
- [2] J. C. Wiltse. History of millimeter and submillimeter waves. *IEEE Trans. Microwave Theory Tech.*, MTT-32:1119–1127, 1984.
- [3] F. C. De Lucia. Spectroscopy in the terahertz spectral regime. In D. Mittleman, editor, *Sensing with terahertz radiation*, pages 39–115. Springer, Berlin, 2003.
- [4] D. Leisawitz *et al.* . Scientific motivation and technology requirements for the SPIRIT and SPECS far-infrared/submillimeter space interferometers. *Proc. SPIE*, 4013:36–46, March. 29-31, 2000.
- [5] T. G. Phillips and J. Keene. Submillimeter astronomy. *Proc. IEEE*, 80:1662–1677, 1992.
- [6] Herschel. <http://sci.esa.int/home/first/>.
- [7] P. H. Siegel. Terahertz technology. *IEEE Trans. Microwave Theory Tech.*, 50:910, 2002.
- [8] Aura Atmospheric Chemistry. <http://eos-aura.gsfc.nasa.gov/>.
- [9] N. C. Luhmann and W. A. Peebles. Instrumentation for magnetically confined fusion plasma diagnostics. *Rev. Sci. Instrum.*, 55:279–331, 1984.
- [10] B. B. Hu and M. C. Nuss. Imaging with terahertz waves. *Opt. Lett.*, 20:1717–1718, 1995.
- [11] Z. Jiang and X. C. Zhang. Terahertz imaging via electrooptic effect. *IEEE Trans. Microwave Theory Tech.*, 47:2644–2650, 1999.
- [12] K. S. Siebert, H. Quast, R. Leonhardt, T. Löfller, M. Thomson, T. Bauer, and H. G. Roskos. Continuous-wave all-optoelectronic terahertz imaging. *Appl. Phys. Lett.*, 80:3003–3005, 2002.
- [13] A. Nahata, J. T. Yardley, and T. F. Heinz. Two-dimensional imaging of continuous-wave terahertz radiation using electro-optic detection. *Appl. Phys. Lett.*, 81:963–965, 2002.

- [14] P. H. Siegel, R. J. Dengler, E. Mueller, C. Readhead, and S. Fraser. Scanned-pixel terahertz heterodyne imaging. Poster at Meeting of the British Royal Society on Terahertz Gap, June 5 2003.
- [15] D. M. Mittleman, J. Cunningham, M. C. Nuss, and M. Geva. Noncontact semiconductor wafer characterization with the terahertz hall effect. *Appl. Phys. Lett.*, 71:16–18, 1997.
- [16] E. K. Duerr. *Distributed photomixers*. PhD thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2002.
- [17] A. Maestrini, G. Chattopadhyay, E. Schlecht, and I. Mehdi. 1400–1900 GHz membrane based Schottky diode triplers. In *Proceedings of the Thirteenth International Space Terahertz Technology Symposium*, Cambridge, MA, Mar. 26–28 2002. Paper 5.1.
- [18] A. Staprans, E. McCune, and J. Ruetz. High-power linear-beam tubes. *Proc. IEEE*, 61:299, 1973.
- [19] W. Shi, Y. J. Ding, N. Fernelius, and K. Vofopyanov. Efficient, tunable, and coherent 0.18–5.27-THz source based on GaSe crystal. *Optics Lett.*, 27:1454–1456, 2002.
- [20] D. Mittleman. Terahertz imaging. In D. Mittleman, editor, *Sensing with terahertz radiation*, pages 117–153. Springer, Berlin, 2003.
- [21] Z. Jiang and X. C. Zhang. Free-space electro-optic techniques. In D. Mittleman, editor, *Sensing with terahertz radiation*, pages 155–192. Springer, Berlin, 2003.
- [22] M. Inguscio, G. Moruzzi, K. M. Evenson, and D. A. Jennings. A review of frequency measurements of optically pumped lasers from 0.1 to 8 THz. *J. Appl. Phys.*, 60:R161, 1986.
- [23] E. R. Mueller, Jr. W. E. Robotham, R. P. Meisner, R. A. Hart, J. Kennedy, and L. A. Newman. 2.5 THz laser local oscillator for the EOS CHEM 1 satellite. In *Proceedings of the Ninth International Space Terahertz Technology Symposium*, pages 563–572, Pasadena, CA, Mar. 17–19 1998.
- [24] E. Gornik and A. A. Andronov, editors. *Opt. Quantum Electron.*, volume 23. Chapman and Hall, London, 1991. Special issue on far infrared semiconductor lasers.
- [25] E. Bründermann, D. R. Chamberlin, and E. E. Haller. High duty cycle and continuous emission from germanium. *Appl. Phys. Lett.*, 76:2991, 2000.
- [26] J. N. Hovenier, M. C. Diez, T. O. Klassen, W. T. Wenckenbach, A. V. Muraviov, S. G. Pavlov, and V. N. Shastin. The  $p$ -Ge terahertz laser—properties under pulsed- and mode-locked operation. *IEEE Trans. Microwave Theory Tech.*, 48:670, 2000.

- [27] I. V. Altukhov, E. G. Chirkova, M. S. Kagan, K. A. Korolev, V. P. Sinis, and F. A. Smirnov. Hot-hole far-IR emission from uniaxially compressed germanium. *Sov. Phys. JETP*, 74:404, 1992.
- [28] Y. P. Gousev, I. V. Altukhov, K. A. Korolev, V. P. Sinis, M. S. Kagan, E. E. Haller, M. A. Odnoblyudov, I. N. Yassievich, and K.-A. Chao. Widely tunable continuous-wave THz laser. *Appl. Phys. Lett.*, 75:757, 1999.
- [29] M. S. Kagan, I.V. Altukhov, V. P. Sinis, S. G. Thomas, K. L. Wang, K. A. Chao, and I. N. Yassievich. Terahertz emission of SiGe/Si quantum wells. *Thin Solid Films*, 380:237, 2000.
- [30] A. Blom, M. A. Odnoblyudov, H. H. Cheng, I. N. Yassievich, and K. A. Chao. Mechanism of terahertz lasing in SiGe/Si quantum wells. *Appl. Phys. Lett.*, 79:713, 2001.
- [31] R. F. Kazarinov and R. A. Suris. Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice. *Soviet Physics - Semiconductors*, 5:707–709, 1971.
- [32] L. Esaki and R. Tsu. Superlattice and negative differential conductivity in semiconductors. *IBM J. Res. Dev.*, 14:61, 1970.
- [33] L. C. West and S. J. Egash. First observation of an extremely large-dipole infrared transition within the conduction band of a GaAs quantum well. *Appl. Phys. Lett.*, 46:1156–1158, 1985.
- [34] F. Capasso, K. Mohammed, and A. Y. Cho. Sequential resonant tunneling through a multiquantum-well superlattice. *Appl. Phys. Lett.*, 48:478–480, 1986.
- [35] M. Helm, E. Colas, P. England, F. Derosa, and S. J. Allen. Observation of grating-induced intersubband emission from GaAs/AlGaAs superlattice. *Appl. Phys. Lett.*, 53:1714–1716, 1988.
- [36] M. Helm, E. Colas, P. England, F. Derosa, and S. J. Allen. Intersubband emission from semiconductor superlattices excited by sequential resonant tunneling. *Phys. Rev. Lett.*, 63:74–77, 1989.
- [37] J. H. Smet. *Intrawell and interwell intersubband transitions in single and multiple quantum well heterostructures*. PhD dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1995.
- [38] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, and A. Y. Cho. Quantum cascade lasers. In H. C. Liu and F. Capasso, editors, *Intersubband Transitions in Quantum Wells: Physics and Device Applications*, number 2, chapter VIII. Academic Press, New York, 2000.

- [39] 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. Quantum cascade lasers: ultrahigh-speed operation, optical wireless communication, narrow linewidth, and far-infrared emission. *J. Quantum Electron.*, 38:511, 2002.
- [40] J. Faist, D. Hofstetter, M. Beck, T. Aellen, M. Rochat, and S. Blaser. Bound-to-continuum and two-phonon resonance quantum-cascade lasers for high duty cycle, high-temperature operation. *J. Quantum Electron.*, 38:533–546, 2002.
- [41] J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, S. N. G. Chu, and A. Y. Cho. Short wavelength ( $\lambda \sim 3.4 \mu\text{m}$ ) quantum cascade laser based on strained compensated InGaAs/AlInAs. *Appl. Phys. Lett.*, 72:680–682, 1998.
- [42] R. Colombelli, F. Capasso, C. Gmachl, A. L. Hutchinson, D. L. Sivco, A. Tredicucci, M. C. Wanke, A. M. Sergent, and A. Y. Cho. Far-infrared surface-plasmon quantum-cascade lasers at  $21.5 \mu\text{m}$  and  $24 \mu\text{m}$  wavelengths. *Appl. Phys. Lett.*, 78:2620–2622, 2001.
- [43] N. Ulbrich, G. Scarpa, A. Sigl, J. Roßkopf, G. Böhm, G. Abstreiter, and M.-C. Amann. High-temperature ( $T > 470 \text{ K}$ ) pulse operation of  $5.5 \mu\text{m}$  quantum cascade lasers with high-reflection coating. *Electron. Lett.*, 37:1341–1342, 2001.
- [44] M. Beck, D. Hofstetter, T. Aellen, J. Faist, U. Oesterle, M. Illegems, E. Gini, and H. Melchior. Continuous wave operation of a mid-infrared semiconductor laser at room temperature. *Science*, 295:301, 2002.
- [45] J. Faist. Personal communication.
- [46] C. Gmachl, A. Straub, R. Colombelli, F. Capasso, D. L. Sivco, A. M. Sergent, and A. Y. Cho. Single-mode, tunable distributed-feedback and multiple-wavelength quantum cascade lasers. *J. Quantum Electron.*, 38:569–581, 2002.
- [47] A. A. Kosterev and F. K. Tittel. Chemical sensors based on quantum cascade lasers. *J. Quantum Electron.*, 38:582–591, 2002.
- [48] C. Sirtori, H. Page, C. Becker, and V. Ortiz. GaAs-AlGaAs quantum cascade lasers: physics, technology and prospects. *IEEE J. Quantum Electron.*, 38:547, 2002.
- [49] G. Dehlinger, L. Diehl, U. Gennser, H. Sigg, J. Faist, K. Ensslin, D. Grütmacher, and E. Müller. Intersubband electroluminescence from silicon-based quantum cascade structures. *Science*, 290:2277–2280, 2000.
- [50] L. Diehl, S. Mentese, E. Müller, D. Grütmacher, H. Sigg, U. Gennser, I. Sagnes, Y. Campidelli, O. Kermarrec, and D. Bensahel. Electroluminescence from strain-compensated  $\text{Si}_{0.2}\text{Ge}_{0.8}/\text{Si}$  quantum cascade structures based on a bound-to-continuum transition. *Appl. Phys. Lett.*, 81:4700, 2002.

- [51] C. Gmachl, Hock M. Ng, and A. Y. Cho. Intersubband absorption in degenerately doped GaN/Al<sub>x</sub>Ga<sub>1-x</sub>N coupled double quantum wells. *Appl. Phys. Lett.*, 79:1590–1592, 2001.
- [52] J. D. Heber, C. Gmachl, H. M. Ng, and A. Y. Cho. Comparative study of ultrafast intersubband electron scattering times at  $\sim 1.55 \mu\text{m}$  wavelength in GaN/AlGaN heterostructures. *Appl. Phys. Lett.*, 81:1237, 2002.
- [53] N. Iizuka, K. Kaneko, N. Suzuki, T. Asano, S. Noda, and O. Wada. Ultrafast intersubband relaxation ( $\leq 150 \text{ fs}$ ) in AlGaN/GaN multiple quantum wells. *Appl. Phys. Lett.*, 77:648–650, 2000.
- [54] P. F. Yuh and K. L. Wang. Novel infrared band aligned superlattice laser. *Appl. Phys. Lett.*, 51:1404, 1987.
- [55] S. I. Borenstain and J. Katz. Evaluation of the feasibility of a far-infrared laser based on intersubband transitions in GaAs quantum wells. *Appl. Phys. Lett.*, 55:654, 1989.
- [56] Q. Hu and S. Feng. Feasibility of far-infrared lasers using multiple semiconductor quantum wells. *Appl. Phys. Lett.*, 59:2923, 1991.
- [57] A. N. Korotkov, D. V. Averin, and K. K. Likharev. Tasers: possible dc pumped terahertz lasers using interwell transitions in semiconductor heterostructures. *Appl. Phys. Lett.*, 65:1865–1867, 1994.
- [58] J. H. Smet, C. G. Fonstad, and Q. Hu. Intrawell and interwell intersubband transitions in multiple quantum wells for far-infrared sources. *J. Appl. Phys.*, 79:9305–9319, 1996.
- [59] B. Xu. *Development of intersubband terahertz lasers using multiple quantum well structures*. PhD dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1998.
- [60] Bin Xu, Qing Hu, and Michael R. Melloch. Electrically pumped tunable terahertz emitter based on intersubband transition. *Appl. Phys. Lett.*, 71:440, 1997.
- [61] M. Rochat, J. Faist, M. Beck, U. Oesterle, and M. Illegems. Far-infrared ( $\lambda = 88 \mu\text{m}$ ) electroluminescence in a quantum casacde structure. *Appl. Phys. Lett.*, 73:3724, 1998.
- [62] J. Ulrich, R. Zobl, N. Finger, K. Unterrainer, G. Strasser, and E. Gornik. Terahertz-electroluminescence in a quantum cascade structure. *Physica B*, 272:216, 1999.
- [63] B. S. Williams, B. Xu, Q. Hu, and M. R. Melloch. Narrow-linewidth terahertz intersubband emission from three-level systems. *Appl. Phys. Lett.*, 75:2927, 1999.

- [64] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi. Terahertz semiconductor-heterostructure laser. *Nature*, 417:156, 2002.
- [65] M. Rochat, L. Ajili, H. Willenberg, J. Faist, H. Beere, G. Davies, E. Linfield, and D. Ritchie. Low-threshold terahertz quantum-cascade lasers. *Appl. Phys. Lett.*, 81:1381, 2002.
- [66] B. S. Williams, H. Callebaut, S. Kumar, Q. Hu, and J. L. Reno. 3.4-THz quantum cascade laser based on longitudinal-optical-phonon scattering for de-population. *Appl. Phys. Lett.*, 82:1015, 2003.
- [67] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno. 3.4 THz quantum cascade laser operating above liquid nitrogen temperature. *Electron. Lett.*, 39:915, 2003.
- [68] L. Ajili, G. Scalari, D. Hofstetter, M. Beck, J. Faist, H. Beere, G. Davies, E. Linfield, and D. Ritchie. Continuous-wave operation of far-infrared quantum cascade lasers. *Electron. Lett.*, 38:1675–1676, 2002.
- [69] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, S. S. Dhillon, and C. Sirtori. High-performance continuous-wave operation of superlattice terahertz quantum-cascade lasers. *Appl. Phys. Lett.*, 82:1518, 2003.
- [70] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno. Terahertz quantum cascade laser at  $\lambda \approx 100 \mu\text{m}$  using metal waveguide for mode confinement. *to be published in Appl. Phys. Lett.*, 2003.
- [71] A. Tredicucci, F. Capasso, C. Gmachl, D. Sivco, A. L. Hutchinson, and A. Y. Cho. High performance interminiband quantum cascade lasers with graded superlattices. *Appl. Phys. Lett.*, 73:2101, 1998.
- [72] G. Scalari, L. Ajili, J. Faist, H. Beere, E. Linfield, D. Ritchie, and G. Davies. Far-infrared ( $\lambda \cong 87 \mu\text{m}$ ) bound-to-continuum quantum- cascade lasers operating up to 90 K. *Appl. Phys. Lett.*, 82:3165, 2003.
- [73] J. P. Loehr and M. O. Manares. Theoretical modeling of the intersubband transitions in III-V semiconductor multiple quantum wells. In M. O. Manares, editor, *Semiconductor quantum wells and superlattices for long-wavelength infrared detectors*, Artech House materials science library, chapter 2, pages 159–188. Artech, Boston, 1993.
- [74] D. F. Nelson, R. C. Miller, and D. A. Kleinman. Band nonparabolicity effects in semiconductor quantum wells. *Phys. Rev. B*, 35:7770, 1987.
- [75] R. P. Leavitt. Empirical two-band model for quantum wells and superlattices in an electric field. *Phys. Rev. B*, 44:11270, 1991.

- [76] M. G. Burt. On the validity and range of applicability of the particle in a box model. *Appl. Phys. Lett.*, 65:717–719, 1994.
- [77] A. Yariv. *Quantum Electronics*. John Wiley & Sons, New York, third edition, 1989.
- [78] C. Sirtori, F. Capasso, and J. Faist. Nonparabolicity and a sum rule associated with bound-to-bound and bound-to-continuum intersubband transitions in quantum wells. *Phys. Rev. B*, 50:8663–8674, 1994.
- [79] B. S. Williams. *GaAs/AlGaAs mid-infrared quantum cascade laser*. MS dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1998.
- [80] S. D. Brorson, H. Yokoyama, and E. P. Ippen. Spontaneous emission rate alteration in optical waveguide structures. *IEEE J. Quantum Electron.*, 26:1492, 1990.
- [81] R. Ferreira and G. Bastard. Evaluation of some scattering times for electrons in unbiased and biased single- and multiple-quantum-well structures. *Phys. Rev. B*, 40:1074–1086, 1989.
- [82] M. Hartig, S. Haacke, B. Deveaud, and L. Rota. Femtosecond luminescence measurements of the intersubband scattering rate in  $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$  quantum wells under selective excitation. *Phys. Rev. B*, 54:14269–14272, 1996.
- [83] J. Shah. Ultrafast luminescence studies of carrier relaxation and tunneling in semiconductor nanostructures. In J. Shah, editor, *Hot carriers in semiconductor nanostructures*, Academic Press Handbook Series, chapter IV.1, pages 169–188. Academic Press, San Diego, 1992.
- [84] H. Callebaut, S. Kumar, B. S. Williams, and Q. Hu. Analysis of transport properties of terahertz quantum cascade lasers. *Appl. Phys. Lett.*, 83:207–209, 2003.
- [85] J. L. Educato, J. P. Leburton, P. Boucaud, P. Vagos, and F. H. Julien. Influence of interface phonons on intersubband scattering in asymmetric coupled quantum wells. *Phys. Rev. B*, 47:12949, 1993.
- [86] H. Rücker, E. Molinari, and P. Lugli. Microscopic calculation of the electron-phonon interaction in quantum wells. *Phys. Rev. B*, 45:6747, 1992.
- [87] S. Borenstain and J. Katz. Intersubband auger recombination and population inversion in quantum-well subbands. *Phys. Rev. B*, 39:10852–10857, 1989.
- [88] J. H. Heyman, K. Unterrainer, K. Craig, B. Galdrikian, M. S. Sherwin, K. Campman, P. F. Hopkins, and A. C. Gossard. Temperature and intensity dependence of intersubband relaxation rates from photovoltage and absorption. *Phys. Rev. Lett.*, 74:2682–2685, 1985.

- [89] J. H. Heyman, K. Unterrainer, K. Craig, J. Williams, M. S. Sherwin, K. Campman, P. F. Hopkins, A. C. Gossard, B. N. Murdin, and C. J. G. M. Langerak. Far-infrared pump-probe measurement of the intersubband lifetime in an Al-GaAs/GaAs coupled-quantum well. *Appl. Phys. Lett.*, 68:3019, 1986.
- [90] K. Luo, H. Zheng, Z. Lu, J. Xu, Z. Xu, T. Zhang, C. Li, X. Yang, and J. Tian. Subband separation energy dependence of intersubband relaxation time in wide quantum wells. *Appl. Phys. Lett.*, 70:1155–1157, 1997.
- [91] M. Hartig, S. Haacke, P. E. Selbmann, B. Deveaud, R. A. Taylor, and L. Rota. Efficient intersubband scattering via carrier-carrier interaction in quantum wells. *Phys. Rev. Lett.*, 80:1940–1943, 1998.
- [92] M. Hartig, J. D. Ganiére, P. E. Selbmann, B. Deveaud, and L. Rota. Density dependence of carrier-carrier-induced intersubband scattering in GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As quantum wells. *Phys. Rev. B*, 60:1500–1503, 1999.
- [93] M. Dür, S. M. Goodnick, P. Lugli, and B. Deveaud. Effect of intercarrier scattering on intersubband transitions in GaAs/AlGaAs quantum well systems. *Physica B*, 272:230–233, 1999.
- [94] H. Callebaut. Personal communication.
- [95] P. Harrison. The nature of the electron distribution functions in quantum cascade lasers. *Appl. Phys. Lett.*, 75:2800, 1999.
- [96] A. Tomita, J. Shah, J. E. Cunningham, S. M. Goodnick, P. Lugli, and S. L. Chuang. Erratum: Femtosecond hole relaxation in *n*-type modulation-doped quantum wells. *Phys. Rev. B*, 52:5445, 1995.
- [97] P. Harison. *Quantum Wells, Wires and Dots*. John Wiley and Sons, Chichester, 2000.
- [98] S.-C. Lee and I. Galbraith. The intrasubband and intersubband relaxation of nonequilibrium electron populations in wide semiconductor quantum wells. *Physica E*, 7:229–232, 2000.
- [99] M. R. S. Tavares, G.-Q Hai, and S. Das Sarma. Carrier relaxation due to electron-electron interaction in coupled double quantum well structures. *Phys. Rev. B*, 64:0453250, 2001.
- [100] K. Kempa, P. Bakshi, J. Engelbrecht, and Y. Zhou. Intersubband electron transitions due to electron-electron interactions in quantum-well structures. *Phys. Rev. B*, 61:11083–11087, 2000.
- [101] P. Hyldgaard and J. W. Wilkins. Electron-electron scattering in far-infrared quantum cascade lasers. *Phys. Rev. B*, 53:6889, 1996.

- [102] R. C. Iotti and F. Rossi. Nature of charge transport in quantum-cascade lasers. *Phys. Rev. Lett.*, 87:146603, 2001.
- [103] F. Eickemeyer, K. Reimann, M. Woerner, T. Elsaesser, S. Barbieri, C. Sirtori, G. Strasser, T. Müller, R. Bratschitsch, and K. Unterrainer. Ultrafast coherent electron transport in semiconductor quantum cascade structures. *Phys. Rev. Lett.*, 89:047402, 2002.
- [104] C. Sirtori, F. Capasso, J. Faist, A. L. Hutchinson, D. L. Sivco, and A. Y. Cho. Resonant tunneling in quantum cascade lasers. *IEEE J. Quantum Electron.*, 34:1722, 1998.
- [105] S. Mukamel. *Principles of nonlinear optical spectroscopy*. Oxford University Press, New York, 1995.
- [106] M. V. Kisin, V. B. Gorfinkel, M. A. Stroscio, G. Belenky, and S. Luryi. Influence of complex phonon spectra on intersubband optical gain. *J. Appl. Phys.*, 82:2031, 1997.
- [107] J. K. Jain and S. Das Sarma. Role of discrete slab phonons in carrier relaxation in semiconductor quantum wells. *Phys. Rev. Lett.*, 62:2305, 1989.
- [108] M. Dutta and M. A. Stroscio. Comment on “Energy level schemes for far-infrared quantum well lasers” [Appl. Phys. Lett. 73, 300 (1998)]. *Appl. Phys. Lett.*, 74:2555, 1999.
- [109] V. M. Menon, W. D. Goodhue, A. S. Karakashian, and L. R. Ram-Mohan. Phonon mediated lifetimes in intersubband terahertz lasers. *J. Appl. Phys.*, 88:5262, 2000.
- [110] N. Mori and T. Ando. Electron-optical-phonon interaction in single and double heterostructures. *Phys. Rev. B*, 40:6175, 1989.
- [111] J. J. Licari and R. Evrard. Electron-phonon interaction in a dielectric slab: effect of the electronic polarizability. *Phys. Rev. B*, 15:2254, 1977.
- [112] K. W. Kim and M. A. Stroscio. Electron-optical-phonon interaction in binary-ternary heterostructures. *J. Appl. Phys.*, 68:6289, 1990.
- [113] H. B. Teng, J. P. Sun, G. I. Haddad, M. A. Stroscio, S. G. Yu, and K. W. Kim. Phonon assisted intersubband transitions in step quantum well structures. *J. Appl. Phys.*, 84:2155, 1998.
- [114] B. K. Ridley. *Electrons and Phonons in Semiconductor Multilayers*. Cambridge University Press, Cambridge, 1997.
- [115] K. Huang and B. Zhu. Dielectric continuum model and Fröhlich interaction in superlattices. *Phys. Rev. B*, 38:13377, 1988.

- [116] K. T. Tsen, K. R. Wald, T. Ruf, P. Y. Yu, and H. Morkoc. Electron-optical-phonon interactions in ultrathin GaAs/AlAs multiple quantum wells. *Phys. Rev. Lett.*, 67:2557–2560, 1991.
- [117] J. Zhang, B. Zhu, and K. Huang. Comparison of cooling rates for hot carriers in GaAs/AlAs quantum wells based on macroscopic and microscopic phonon models. *Phys. Rev. B*, 59:13184–13195, 1999.
- [118] S. G. Yu, K. W. Kim, M. A. Stroscio, G. J. Iafrate, J.-P. Sun, and G. I. Haddad. Transfer matrix method for interface optical-phonon modes in multiple-interface heterostructure systems. *J. Appl. Phys.*, 82:3363, 1997.
- [119] M. V. Kisin, M. A. Stroscio, G. Belenky, V. B. Gorfinkel, and S. Luryi. Effects of interface phonon scattering in three-interface heterostructures. *J. Appl. Phys.*, 83:4816, 1998.
- [120] H. C. Liu, I. W. Cheung, A. J. SpringThorpe, C. Dharma-wardana, Z. R. Wasilewski, D. J. Lockwood, and G. C. Aers. Intersubband Raman laser. *Appl. Phys. Lett.*, 78:3580, 2001.
- [121] C. Becker, C. Sirtori, H. Page, A. Robertson, V. Ortiz, and X. Marcadet. Influence of confined phonon modes on the thermal behavior of AlAs/GaAs quantum cascade structures. *Phys. Rev. B*, 65:085305, 2002.
- [122] H. Kogelnik. Theory of optical waveguides. In T. Tamir, editor, *Guided-Wave Optoelectronics*, number 26 in Springer Series in Electronics and Photonics, chapter 2, pages 7–87. Springer-Verlag, Berlin, second edition, 1990.
- [123] N. W. Ashcroft and N. D. Mermin. *Physics of Optoelectronic Devices*, chapter 1, pages 1–27. Harcourt Brace College Publishing, Fort Worth, first edition, 1976.
- [124] B. Jensen. The quantum extension of the Drude-Zener theory. In E. D. Palik, editor, *Handbook of optical constants of solids*, number 26 in Academic Press Handbook Series, chapter 9, pages 169–188. Academic Press, 1998.
- [125] T.-I. Jeon and D. Grischkowsky. Characterization of optically dense, doped semiconductors by reflection THz time domain spectroscopy. *Appl. Phys. Lett.*, 72:3032, 1998.
- [126] P. G. Huggard, J. A. Cluff, G. P. Moore, C. J. Shaw, S. R. Andrews, S. R. Keiding, E. H. Linfield, and D. A. Ritchie. Drude conductivity of highly doped GaAs at terahertz frequencies. *J. Appl. Phys.*, 87:2382, 2000.
- [127] M. Lundstrom. *Fundamentals of carrier transport*, section 4.3–4.4, pages 164–172. Cambridge University Press, Cambridge, second edition, 2000.
- [128] V. W. L. Chin, T. Osotchan, M. R. Vaughan, T. L. Tansley, G. J. Griffiths, and Z. Kachwalla. Hall and drift mobilities in molecular beam epitaxial grown GaAs. *J. Elec. Mat.*, 22:1317, 1993.

- [129] D. C. Look, C. E. Stutz, J. R. Sizelove, and K. R. Evans. On Hall scattering factors for holes in GaAs. *J. Appl. Phys.*, 80:1913, 1996.
- [130] M. A. Ordal, L. L. Long, R. J. Bell, S. E. Bell, R. R. Bell, Jr. R. W. Alexander, and C. A. Ward. Optical properties of the metals Al, Co, Cu, Au, Fe, Pf, Ni, Pd, Pt, Ag, Ti, and W in the infrared and far infrared. *Appl. Optics*, 22:1099, 1983.
- [131] Optical properties of pure metals and binary alloys. In K.-H Hellwege and O. Madelung, editors, *Landolt-Börnstein*, volume III/15b of *New Series*, chapter 4, pages 210–222. Springer-Verlag, Berlin, 1985.
- [132] I. Vurgaftman and J. R. Meyer. TE- and TM-polarized roughness-assisted free-carrier absorption at midinfrared and terahertz wavelengths. *Phys. Rev. B*, 60:14294, 1999.
- [133] T. Ando, A. B. Fowler, and F. Stern. Electronic properties of two-dimensional systems. *Rev. Mod. Phys.*, 54:437, 1982.
- [134] W. J. Moore and R. T. Holm. Infrared dielectric constant of gallium arsenide. *J. Appl. Phys.*, 80:6939, 1996.
- [135] E. D. Palik. Gallium arsenide. In E. D. Palik, editor, *Handbook of optical constants of solids*, Academic Press Handbook Series, pages 429–443. Academic Press, Orlando, 1985.
- [136] M. Rochat, M. Beck, J. Faist, and E. Oesterle. Measurement of far-infrared waveguide loss using a multisection single-pass technique. *Appl. Phys. Lett.*, 78:1967, 2001.
- [137] K. Unterrainer, R. Colombelli, C. Gmachl, F. Capasso, H. Y. Hwang, A. M. Sergent, D. L. Sivco, and A. Y. Cho. Quantum cascade lasers with double metal-semiconductor waveguide resonators. *Appl. Phys. Lett.*, 80:3060, 2002.
- [138] S. M. Kohen, B. S. Williams, and Q. Hu. Two-dimensional modeling of terahertz waveguides. *to be submitted to J. Appl. Phys.*
- [139] J. S. Yu, S. Slivken, A. Evans, J. David, and M. Razeghi. Very high average power at room temperature from  $\lambda \approx 5.9\text{-}\mu\text{m}$  quantum-cascade lasers. *Appl. Phys. Lett.*, 82:3397–3399, 2003.
- [140] Bin Xu and Qing Hu. Grating coupling for intersubband emission. *Appl. Phys. Lett.*, 70:2511–2513, 1997.
- [141] W. T. Welford and R. Winston. *High collection nonimaging optics*. Academic Press, San Diego, first edition, 1989.
- [142] D. A. Harper, R. H. Hildebrand, R. Stiening, and R. Winston. Heat trap: an optimized far infrared field optics system. *Applied Optics*, 15:53, 1976.

- [143] E. V. Loewenstein, D. R. Smith, and R. L. Morgan. Optical constants of far infrared materials. 2: Crystalline solids. *Appl. Optics*, 12:398–405, 1973.
- [144] L. Mertz. *Transformations in Optics*. John Wiley and Sons, New York, 1965.
- [145] I. Hallakoun, T. Boterashvili, G. Bunin, and Y. Shapira. Electronchemical etching impact on GaAs process, mask design and device performance. In *2000 GaAsMANTECH Conference*, pages 25–27, Washington DC, May. 1–4 2000.
- [146] Y. Zhao, Y. Tkachenko, and D. Bartle. Supression of electrochemical etching effects in GaAs PHEMTs. In *21st Annual GaAs IC Symposium Technical Digest 1999*, pages 163–166, Monterey, CA, Oct. 17–20 1999.
- [147] M. P. Patkar, T. P. Chin, J. M. Woodall, M. S. Lundstrom, and M. R. Melloch. Very low resistance nonalloyed ohmic contacts using low-temperature molecular beam epitaxy of GaAs. *Appl. Phys. Lett.*, 66:1412–1414, 1995.
- [148] L. Parameswaran. Silicon pressure sensor using wafer bonding technology. MS thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1993.
- [149] H. J. Ueng, V. R. Kolagunta, D. B. Janes, K. J. Webb, D. T McInturff, and M. R. Melloch. Annealing stability and device application of nonalloyed ohmic contacts using a low temperature grown GaAs cap on thin  $n^+$  GaAs layers. *Appl. Phys. Lett.*, 71:2496–2498, 1997.
- [150] L. C. Wang, X. Z. Wang, S. S. Lau, T. Sands, W. K. Chen, and T. F. Kuech. Stable and shallow PdIn ohmic contacts to  $n$ -GaAs. *Appl. Phys. Lett.*, 56:2129, 1990.
- [151] E. D. Marshall, B. Zhang, L. C. Wang, P. F. Jiao, W. X. Chen, T. Sawada, S. S. Lau, K. L. Kavanagh, and T. F. Kuech. Nonalloyed ohmic contacts to  $n$ -GaAs by solid phase epitaxy of Ge. *J. Appl. Phys.*, 62:942, 1987.
- [152] E. Yablonovitch, T. Sands, D. M. Hwang, I. Schnitzer, T. J. Gmitter, S. K. Shastry, D. S. Hill, and J. C. C. Fan. Van der Waals bonding of GaAs on Pd leads to a permanent solid-phase-topotaxial, metallurgical bond. *Appl. Phys. Lett.*, 59:3159, 1991.
- [153] H. Okamoto and T. B. Massalski. In T. B. Massalski, editor, *Binary Alloy phase diagrams*, pages 381–383. ASM International, Metals Park, Ohio, 1990.
- [154] C. C. Lee, C. Y. Wang, and G. Matijasevic. Au-In bonding below the eutectic temperature. *IEEE Trans. Comp., Hybrids, Manuf. Technol.*, 16:311, 1993.
- [155] T. B. Wang, Z. Z. Shen, R. Q. Ye, X. M. Xie, F. Stubhan, and J. Freytag. Die bonding with Au/In isothermal solidification technique. *J. Electron. Mat.*, 29:443, 2000.

- [156] R. Zhao, W. S. Lau, T. C. Chong, and M. F. Li. A comparison of the selective etching characteristics of conventional and low-temperature-grown GaAs over AlAs by various etching solutions. *Jpn. J. Appl. Phys.*, 35:22–25, 1996.
- [157] K. Fobelets, R. Vounckx, and G. Borghs. A GaAs pressure sensor based on resonant tunneling diodes. *J. Micromech. Microeng.*, 4:123–128, 1994.
- [158] M. Tong, D. G. Ballegeer, A. Ketterson, E. J. Roan, K. Y. Cheng, and I. Adesida. A comparative study of wet and dry selective etching processes for GaAs/AlGaAs/InGaAs pseudomorphic MODFETs. *J. Electron. Mat.*, 21:9–15, 1992.
- [159] C. Carter-Coman, R. Bicknell-Tassius, R. G. Benz, A. S. Brown, and N. M. Jokerst. Analysis of GaAs substrate removal etching with citric acid:H<sub>2</sub>O<sub>2</sub> and NH<sub>4</sub>OH:H<sub>2</sub>O<sub>2</sub> for application to compliant substrates. *J. Electrochem. Soc.*, 144:L29–L31, 1997.
- [160] E. Yablonovitch, T. Gmitter, J. P. Harbison, and R. Bhat. Extreme selectivity in the lift-off of epitaxial GaAs films. *Appl. Phys. Lett.*, 51:2222–2224, 1987.
- [161] S. Adachi. GaAs, AlAs, and Al<sub>x</sub>Ga<sub>1-x</sub>As: Material parameters for use in research and device applications. *J. Appl. Phys.*, 58:R1–R28, 1985.
- [162] E. T. Yu, J. O McCaldin, and T. C. McGill. Band offsets in semiconductor heterojunctions. In H. Ehrenreich and D. Turnbull, editors, *Solid State Physics: Advances in Research and Applications*, number 46, pages 33–41. Academic Press, Boston, 1992.
- [163] S. Blaser, M. Rochat, M. Beck, , D. Hofstetter, and J. Faist. Terahertz intersubband emission in strong magnetic fields. *Appl. Phys. Lett.*, 81:67–69, 2002.
- [164] C. Sirtori. Personal communication.
- [165] J. Ulrich, R. Zobl, K. Unterrainer, G. Strasser, and E. Gornik. Magnetic-field-enhanced quantum-cascade emission. *Appl. Phys. Lett.*, 76:19, 2000.
- [166] S. Blaser, M. Rochat, M. Beck, and J. Faist. Far-infrared emission and Stark-cyclotron resonances in a quantum-cascade structure based on photon-assisted tunneling transition. *Phys. Rev. B*, 61:8369, 2000.
- [167] J. H. Smet, C. G. Fonstad, and Q. Hu. Magnetotunneling spectroscopy in wide InGaAs/InAlAs double quantum wells. *Appl. Phys. Lett.*, 63:2225, 1993.
- [168] B. S. Williams, H. Callebaut, Q. Hu, and J. L. Reno. Magnetotunneling spectroscopy of resonant anticrossing in terahertz intersubband emitters. *Appl. Phys. Lett.*, 79:4444, 2001.
- [169] B. W. Hakki and T. L. Paoli. Gain spectra in gaas double-heterostructure injection lasers. *J. Appl. Phys.*, 46:1299–1306, 1976.

- [170] M. A. Stroscio, M. Kisin, G. Belenk, and S. Luryi. Phonon enhanced inverse population in asymmetric double quantum wells. *Appl. Phys. Lett.*, 75:3258, 1999.
- [171] H. Callebaut. *GaAs/AlGaAs far-infrared quantum cascade laser*. M.S. thesis, Massachusetts Institute of Technology, (unpublished), (2001).
- [172] S. M. Kohen, 2003. personal communication.
- [173] J. Faist, A. Tredicucci, Federico Capasso, C. Sirtori, D. L. Sivco, J. N. Biallargeon, A. L. Hutchinson, and A. Y. Cho. High-power continuous-wave quantum cascade lasers. *IEEE J. Quantum Electron.*, 34:336, 1998.
- [174] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno. Terahertz quantum-cascade laser operating up to 137 k. *submitted to Appl. Phys. Lett.*, 2003.
- [175] R. Köhler. Personal communication.
- [176] W. S. Capinski, H. J. Maris, T. Ruf, M. Cardona, K. Ploog, and D. S. Katzer. Thermal-conductivity measurements of GaAs/AlAs superlattices using a picosecond optical pump-and-probe technique. *Phys. Rev. B*, 59:8105–8113, 1999.
- [177] V. Spagnolo, M. Troccoli, G. Scamarcio, C. Becker, G. Glastre, and C. Sirtori. Thermal resistance and temperature characteristics of GaAs/Al<sub>0.33</sub>Ga<sub>0.67</sub>As quantum-cascade lasers. *Appl. Phys. Lett.*, 78:1177–1179, 2001.
- [178] A. J. Borak, C. C. Phillips, and C. Sirtori. Temperature transients and thermal properties of GaAs/AlGaAs quantum-cascade lasers. *Appl. Phys. Lett.*, 82:4020–4022, 2003.
- [179] C. Pflügl, M. Litzenberger, W. Schrenk, D. Pogany, E. Gornik, and G. Strasser. Interferometric study of thermal dynamics in GaAs-based quantum-cascade lasers. *Appl. Phys. Lett.*, 82:1664–1666, 2003.
- [180] V. Spagnolo, G. Scamarcio, D. Marano, H. Page, and C. Sirtori. Thermoelastic stress in GaAs/AlGaAs quantum cascade lasers. *Appl. Phys. Lett.*, 82:4639–4641, 2003.
- [181] K. N. Chen, A. Fan, C. S. Tan, R. Reif, and C. Y. Wen. Microstructure evolution and abnormal grain growth during copper wafer bonding. *Appl. Phys. Lett.*, 81:3774–3776, 2002.
- [182] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, and D. A. Ritchie. Low-threshold quantum-cascade lasers at 3.5 THz ( $\lambda = 85 \mu\text{m}$ ). *Opt. Lett.*, 28:810–812, 2003.
- [183] A. E. Siegman. *Lasers*. University Science Books, Mill Valley, California, first edition, 1986.