

7

GaAs/AlGaAs Mid-Infrared Quantum Cascade Laser

by

Benjamin S. Williams

Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degree of

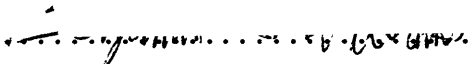
Master of Science in Electrical Engineering and Computer Science

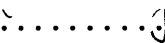
at the


MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1998

© Massachusetts Institute of Technology 1998. All rights reserved.

Author 
Department of Electrical Engineering and Computer Science
May 8, 1998

Certified by 
Qing Hu
Associate Professor of Electrical Engineering and Computer Science
Thesis Supervisor

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

JUL 28 1998

LIBRARIES

GaAs/AlGaAs Mid-Infrared Quantum Cascade Laser

by

Benjamin S. Williams

Submitted to the Department of Electrical Engineering and Computer Science
on May 8, 1998, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

Abstract

This thesis explores the development of mid-infrared ($\lambda \sim 8 - 12 \mu\text{m}$) quantum cascade lasers developed using GaAs/Al_xGa_{1-x}As heterostructures. Such a laser could fill the need for compact, powerful, tunable mid-infrared coherent sources operating at room temperature. A mid-infrared laser operating in the 8 – 12 μm transmission window could be used for spectroscopy and remote sensing of chemicals and pollutants. Although such lasers have been recently developed using the InGaAs/InAlAs material system, GaAs based technology is widespread, making such a quantum cascade laser desirable from the production point of view.

The design of a unipolar quantum cascade emission structure was undertaken using a set of numerical simulation programs to self-consistently solve Schrödinger's and Poisson's equations. The structure was grown by MBE and fabricated into testable structures. Cryogenic current-voltage measurements verify electron transport through the multiple quantum well structure by resonant tunneling. The intersubband electroluminescence spectrum was resolved at 10 K, revealing a homogeneously broadened emission peak at $\hbar\omega_0 = 120 \text{ meV}$ ($\lambda \sim 10.3 \mu\text{m}$), with a FWHM as small as 11 meV. These measurements confirm the design of the structure, and indicate its suitability for use as a gain medium in a QCL. A waveguide structure based on metallic and plasma confinement was designed and evaluated for use with the quantum cascade structure. Although this new structure awaits testing, estimates indicate that lasing should be possible.

Thesis Supervisor: Qing Hu

Title: Associate Professor of Electrical Engineering and Computer Science

Acknowledgments

I owe a great deal to the many people that have made it possible to complete this research. Foremost, I want to thank Professor Qing Hu for the opportunity to undertake this research as well as providing the financial support that made it possible. I am tremendously grateful for his confidence in my ability to handle this project, and for all the help he gave me as I began, and continues to give. I also want to thank Brian Riely for working with me on this project. As well as assisting with the MQW and waveguide design, Brian assisted me with and performed many experiments. He also was responsible for the incorporation of nonparabolicity into SEQUAL, construction of the bias supply, development of a waveguide solver, and development of the new mask design. His work on this project has proven invaluable.

I want to thank Bin Xu for training me on his THz laser project during my first year, as well as providing a great deal of guidance and assistance since then. I had many valuable conversations with both him and Ilya Lyubomirsky regarding this project, as well as their own far-infrared laser projects. They have been primarily responsible for my training since I have been here, and I am extremely grateful. I am also thankful to have had the opportunity to work in such a friendly and intellectually stimulating atmosphere. All of the members of the group members, including Gert de Lange, Erik Duerr, Konstantinos Konistis, Yanko Sheiretov, Noah Zamdmer, and Farhan Rana, have provided me with assistance and support in various forms. I especially want to thank Erik for giving such a demonstration on how to write a thesis, as well as providing a humorous outlet for all the frustrations that occur when performing research.

I also want to thank Laura DeStefano for giving me so much emotional support during the course of this research. No doubt she tired of hearing about dewars, cleanish rooms, and the laser that wouldn't, but she kept a brave face through it all. I will forgive her constant attempts to lure me to Miami. I also want to thank my parents for all their love and support. They were always ready to visit or merely to bolster my spirits during difficult times. I also want to thank my brother, if only for always being ready to argue about anything but physics, whether it be constitutional law or the meaning of the previous episode of

“Seinfeld.” This is excepting of course the rare occasion that he would reveal one of his physics “theories.” I also want to thank the rest of my friends and family for all of their love and support during this thesis.

This work was supported by the U. S. Army Research Laboratory Cooperative Agreement DAAL01-96-2-0001.

Contents

1	Introduction	15
1.1	Motivation	15
1.2	Quantum cascade lasers	16
1.3	Project statement	19
1.4	Approach	20
1.5	Thesis overview	21
2	Background - quantum cascade laser theory	23
2.1	Introduction	23
2.2	Intersubband radiative transitions	24
2.3	Fundamental laser theory	29
2.3.1	Optical gain	29
2.3.2	Optical loss	30
2.3.3	Threshold conditions	30
2.3.4	Gain saturation	31
2.4	Cavity loss - slab waveguide	33
2.4.1	Fundamental equations	34
2.4.2	Multilayer waveguide	35
2.4.3	Implementation	37
2.4.4	Calculation results	38
2.5	Drude model	39
2.6	Non-radiative relaxation (LO-phonon scattering)	42

3	Design	47
3.1	Introduction	47
3.2	SEQUAL	48
3.2.1	Theoretical formulation and propagating states	48
3.2.2	Bound states	51
3.2.3	Self-consistency	52
3.2.4	Nonparabolicity	53
3.3	Multiple quantum well design	56
3.3.1	Overview	56
3.3.2	Active region	58
3.3.3	Digitally graded gap	65
3.3.4	Injection and exit barriers	66
3.3.5	Self-consistency	68
3.3.6	General Discussion	69
3.4	Waveguide design	71
3.4.1	A40 design	72
3.4.2	Plasma waveguide	74
3.4.3	Metal-metal waveguide	80
4	Experiments	83
4.1	Setup	83
4.2	Fabrication	88
4.3	Measurements (A40)	91
4.3.1	Emission measurements	91
4.3.2	Heating	97
5	Discussion	101
5.1	A40 Results	101
5.1.1	Linewidth broadening	102
5.2	Laser design PL50	106
5.3	Conclusion	107

List of Figures

1-1	<i>Conduction band profile and squared magnitude wavefunctions for original quantum cascade laser developed at Bell Labs by Faist et al. . Figure taken from Ref. [6]</i>	17
2-1	<i>Geometry of $N + 2$ layer slab waveguide.</i>	34
2-2	<i>Plasma frequency for bulk GaAs as a function of free carrier density n using $\epsilon(\lambda = 10 \mu\text{m}) = 10.71\epsilon_0$. Note that our QCL is designed to operate at approximately 30 THz ($\lambda \sim 10 \mu\text{m}$).</i>	40
2-3	<i>Complex refractive index $N = \sqrt{\epsilon/\epsilon_0} = n + ik$ versus wavelength for three free carrier densities. The calculation was performed using a momentum relaxation time of $\tau = 0.1$ ps and a permittivity $\epsilon(\lambda = 10 \mu\text{m}) = 10.71\epsilon_0$</i>	41
2-4	<i>Free carrier loss in GaAs calculated with Drude model using a momentum relaxation time of $\tau = 0.1$ ps and a permittivity of $\epsilon(\lambda = 10 \mu\text{m}) = 10.71\epsilon_0$</i>	42
3-1	<i>Conduction band profile for mid-infrared (10 μm) quantum cascade laser.</i>	57
3-2	<i>Double well active region. The layer thicknesses are 42.4, 67.8, 11.3, 53.7, and 25.4 \AA (15/24/4/19/9 ML). The intrawell lasing transition occurs between E_3 and E_2. The envelope functions are obtained from (3.21) and the squared magnitude $\psi_i(z) ^2$ plotted at their respective energies E_i.</i>	58
3-3	<i>Subband dispersion in active region. The wavy line indicates the calculated 124 meV optical transition between E_3 and E_2. LO-phonon scattering is the dominant scattering mechanism; both intersubband and intrasubband transitions are indicated by the solid arrows. E_{32} is set at the LO phonon resonance (36 meV) to provide fast depopulation of E_2.</i>	60
3-4	<i>Evolution of states E_1 and E_2 under bias. Shown are three bias points: (a) zero bias, (b) anticrossing point $V = 0.138$ V, (c) operating point $V = 0.265$ V. The E_3 level is shown for reference, although it was not calculated using nonparabolicity.</i>	63
3-5	<i>Energy difference E_{21} as a function of bias in two structures. The actual design structure has a center barrier 4 ML thick and a large Δ_0, while a similar structure with a 7 ML center barrier has a small Δ_0. The operating voltage where E_{21} is indicated by V_{op}.</i>	63

3-6	<i>Digitally graded gap. The layer thicknesses are 25.4, 36.7, 25.4, 31.1, 25.4, 25.4, 25.4, 22.6, 25.4, 25.4, 28.3 and 42.4 Å (9/13/9/11/9/9/9/8/9/9/10/15 ML). The six-transmission state squared-magnitude wavefunctions that make up the miniband are shown at their respective energies. Also shown is the first of the continuum states above the minigap. The lowest state G is aligned with E_3 in the next module.</i>	64
3-7	<i>Transmission through six-well digitally graded gap under design bias shown in Fig. 3-6. The energy bands of high and low transmission are labeled the “miniband” and “minigap” respectively. The energies of the three states from the previous active region are indicated by arrows.</i>	65
3-8	<i>Injection barrier anticrossing. (a) Energy difference $E_G - E_3$. (b) Interface module at anticrossing voltage $V_{mod} = 0.234$ V.</i>	68
3-9	<i>Self-consistent band structure (dashed) calculated with the electron population of 1.5×10^{11} cm⁻² split 50/50 between the injection level G and the excited state at E_3. The flat-band structure is shown for comparison.</i>	69
3-10	<i>Calculated bound TM_0 mode for A40 design with a loss of $\alpha_c = 149$ cm⁻¹ and a confinement of $\Gamma = 0.58$. The axis orientation is shown in (a) and the simulated structure is shown in (b). The high doping density in the buffer layer adjacent to the metal contact is meant to simulate the effect of annealing. The real part of the refractive index profile is displayed in (c). The horizontal line gives the real part of the mode effective index $n_{eff} = \beta/k_0$. The field components are shown in (d), (e) and (f), with the real and imaginary parts given by the solid and dotted lines respectively.</i>	73
3-11	<i>(a) Proposed plasma-metal waveguide structure with non-alloyed ohmic contact. (b) Variation of cavity loss α_c and confinement factor Γ with thickness of the core active region. For this plot, the doping of the upper buffer layer and lower cladding layer are both taken as $N_D = 2 \times 10^{18}$ cm⁻³.</i>	76
3-12	<i>Variation of the cavity loss α_c and confinement factor Γ of proposed plasma-metal waveguide for different doping levels of the (a) lower cladding layer and (b) upper buffer layer. For each plot, the doping of the other layer was taken to be $N_D = 2 \times 10^{18}$ cm⁻³ and the thickness of the core region is 2.5 μm.</i>	76
3-13	<i>Bound TM_0 mode for proposed PL50 design, with cavity loss $\alpha_c = 64$ cm⁻¹ and confinement $\Gamma = 0.85$. The axis orientation is shown in (a) and the simulated structure with non-alloyed ohmic contact is shown in (b). The real part of the refractive index profile is displayed in (c). The horizontal line gives the real part of the mode effective index $n_{eff} = \beta/k_0$. The field components are shown in (d), (e) and (f), with the real and imaginary parts given by the solid and dotted lines respectively.</i>	78
3-14	<i>Drude model complex refractive index $N = n + ik$ for gold. The temperature dependence is illustrated with typical momentum relaxation times τ, taken from DC conductivity data[51].</i>	79
4-1	<i>MBE growth profile for spontaneous emitter design A40.</i>	84

4-2	<i>Copper mount used to hold electroformed cone and chip carrier with device. The mount was screwed to the cold plate, providing good heat sinking for the device. All metal-metal surfaces were plated with indium to ensure maximum thermal conductivity. A temperature sensor (not shown) was attached to the mount just below the chip carrier.</i>	85
4-3	<i>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 less than $\theta_{max} = \theta$.</i>	86
4-4	<i>Experimental setup for emission measurements.</i>	87
4-5	<i>Fabrication sequence used for A40 device.</i>	90
4-6	<i>Device B: (a) Pulsed current and optical power versus device bias. (b) Pulsed ohmic power $P = IV$ compared with measured optical power. . . .</i>	92
4-7	<i>Emission spectra from Device B. (a) Three characteristic spectra at different biases measured with 25 μs pulse width, 1% duty cycle, at 14 K. (b) Emission spectrum and its Lorentzian fit with narrowest observed linewidth collected using 15 μs pulse width at $V_D = 9.9$ V. Inset shows the corresponding interferogram.</i>	94
4-8	<i>Optical power vs. bias for device C. Also plotted is the heat sink temperature as a gauge of average power dissipation during the measurement. The device was biased with 25 μs voltage pulses with a 1% duty cycle.</i>	95
4-9	<i>Narrowest emission spectra from Device C at $V_D = 11.3$ V, 25 μs pulse width, 1% duty cycle. Lorentzian and Gaussian fits to this data are shown for comparison.</i>	96
4-10	<i>Emission spectra from Device C. (a) Four characteristic spectra at different biases measured with 25 μs pulse width, 1% duty cycle, at 9-10 K. (b) Lorentzian fits to the four spectra. The two insets show the various peak positions and linewidths measured for all similarly measured device C data. . . .</i>	96
4-11	<i>High bias emission spectra from Device B. The emission spectrum for $V_D = 12.5$V is shown with a measured 300 K blackbody spectrum for comparison. Note that the peak in the blackbody spectrum is not the true peak, which is located at 75 meV, but rather due to the detector response near the roll-off at 80 meV. The narrow intersubband emission spectrum is shown for comparison.</i>	98
4-12	<i>Time-resolved high bias emission spectra from Device B. A gated boxcar averager was used to collect the spectra during 4 sequential 5 μs periods of the 25 μs pulse. A narrow intersubband emission spectrum (scaled differently) is shown for linewidth comparison. The inset shows the change in the current and emitted optical power over the course of the pulse.</i>	99
5-1	<i>MBE growth diagram for proposed PL50 laser structure.</i>	108
5-2	<i>Processing sequence for PL50 laser design.</i>	109

This page intentionally left blank.

List of Tables

- 3.1 *Active region properties* 59
- 3.2 *Exit barrier WKB tunneling times* 68
- 3.3 *Effect of nonparabolicity* 70

- 4.1 *Emission devices tested* 91

Chapter 1

Introduction

1.1 Motivation

This research project involved the development of mid-infrared quantum cascade emitters and lasers operating in the technologically important 8-12 μm spectral range. The mid-infrared, extending from 2 – 20 μm , encompasses two atmospheric transmission windows, from 3 – 5 μm and from 8 – 12 μm . Outside of these “windows”, common atmospheric gases such as nitrogen, carbon dioxide and water vapor strongly absorb mid-infrared radiation. High resolution gas spectroscopy is one of the major applications of mid-infrared lasers, with many systems for hazardous gas control and processing based around such methods. Lasers operating inside these transmission windows are useful for remote chemical sensing, pollution monitoring, and point-to-point communications. Most molecules, including many environmental pollutants such as CO and SO₂ exhibit characteristic absorption and emission lines in this range. Provided it can be tuned to their resonances, an infrared laser can be used to make high sensitivity measurements of these absorption and emission lines. A laser spectroscopy system operated in one of the atmospheric transmission windows could be used for remote chemical detection, useful in environmental monitoring or chemical weapons detection. Another application is in semiconductor manufacturing where the system can be use in end-point detection for dry etch processes. When etching through a material layer grown upon a substrate of a different material, the machine can detect that the upper layer has been removed by monitoring the chamber gas for the

identifying absorption lines of the substrate.

A tunable, compact, room temperature laser operating in this spectral range has been long sought after. Currently available lasers in the mid-infrared are all lacking in one or more important aspects. Gas lasers, such as CO₂ lasers, offer high output powers, but offer only limited tunability, as the output frequencies are restricted to transitions between rotational-vibrational energy levels. Gas lasers are also somewhat unwieldy, being themselves large and requiring high voltage supplies. Mid-infrared semiconductor interband laser diodes have been developed from II-VI or IV-VI lead salt materials[1]. While these lead-salt lasers have been quite successful for high resolution spectroscopy, they are still limited to cryogenic operation and provide relatively low power. High power operation is fundamentally limited by high Auger band to band recombination rates, which dramatically increase at higher temperatures.

1.2 Quantum cascade lasers

The recently developed quantum cascade laser (QCL) offers an attractive new alternative for mid-infrared coherent sources. The QCL is a unipolar quantum well laser, where the lasing transition is between conduction subbands in a series of quantum wells. The one-dimensional wells are created in the conduction band edge by depositing thin epitaxial layers of material of differing bandgaps. As a result, electron energies are quantized in the growth direction, while two-dimensional free electron motion is retained in each “subband”. This concept of using a periodic modulation of the conduction band edge to form a semiconductor “superlattice” was first suggested by Esaki and Tsu[2] in 1970. Soon afterward, Kazarinov and Suris proposed using a superlattice to amplify radiation from photon-assisted tunneling between subbands[3]. Helm *et al.* successfully observed spontaneous emission from subbands in a superlattice populated through resonant tunneling in 1989[4]. Despite this, it was realized that space charge buildup in a symmetric superlattice would lead to the development of high field domains and instability when high currents were applied[5].

The breakthrough quantum cascade laser was first developed at Bell Labs by Faist *et al.*

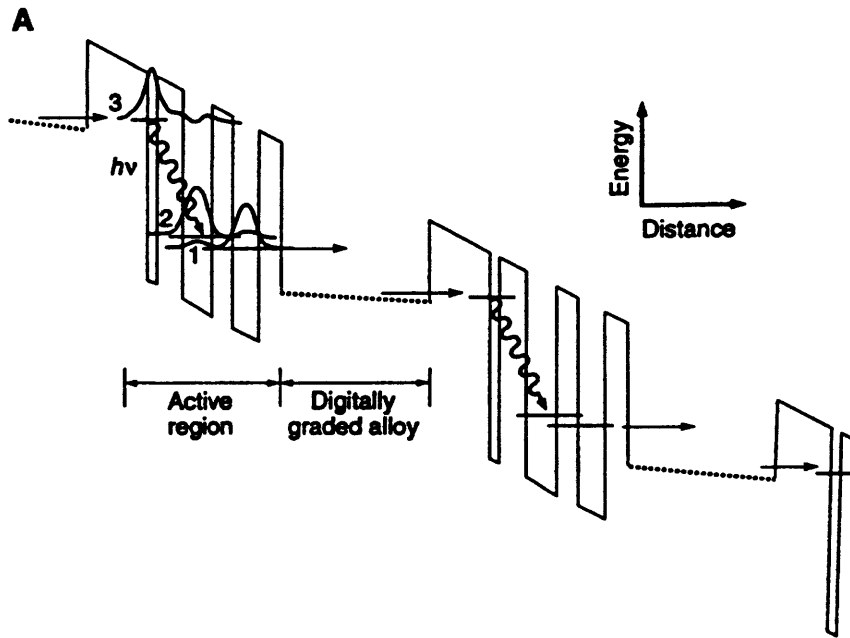


Figure 1-1: *Conduction band profile and squared magnitude wavefunctions for original quantum cascade laser developed at Bell Labs by Faist et al. . Figure taken from Ref. [6]*

in 1994[6]. This design featured a triple well active region, with each well containing one subband. The profile of the conduction band edge with envelope wavefunctions in the active region is shown in Fig. 1-1. The lasing transition takes place between level 3 and 2. Since level 2 and level 1 are separated by the longitudinal-optical (LO) phonon resonance energy, the second level is quickly depopulated through fast LO-phonon scattering. LO-scattering from level 3 to 2 is slowed by the large in-plane momentum transfer, and thus a population inversion between levels is achieved. Electrons are injected into level 3 and extracted from level 1 through the use of a digitally graded region designed to create a highly transmissive path for electrons to travel to the next module. Each digitally graded region also is doped to provide carriers to prevent the buildup of charge from the contacts and to maintain global charge neutrality.

One great advantage of the QCL over interband lasers is the ability to engineer the subband separation, and thus the lasing wavelength, merely by choosing the thickness of the wells and barriers. This is in contrast to a conventional semiconductor laser, where the lasing transition is between the conduction and valence bands, and the frequency of the

emitted radiation is principally determined by the bandgap of the material. This flexibility allows QCLs to cover unavailable spectral ranges using technologically mature III-V materials. This is especially convenient for infrared applications where appropriate small bandgap materials become difficult to find and work with. The emission frequency can be theoretically made quite small, allowing the development of far-infrared emitters in the range 30 – 300 μm [7, 8]. Also, the QCL is a unipolar device (no electron-hole recombination is involved), and hence a single electron has the opportunity to emit many photons before exiting the device.

Faist's QCL was first developed to operate at 4.2 μm in pulsed mode with peak power of 8 mW at 10 K with a threshold current density of 14 kA/cm^2 [6]. Since then, the quantum cascade lasers have improved dramatically in performance, with many improvements made in the quantum well design, waveguiding, and heat sinking. Bell Labs has demonstrated lasers at wavelengths from 3 – 12 μm , at room temperature pulsed operation[9, 10] and continuous wave cryogenic operation. Continuous wave (CW) operation has been obtained with power as high as 0.7 W per facet at 20 K, with CW operation reported up to 110 K[11]. The longer wavelength lasers tend to be less efficient, due to reduced stimulated emission efficiency and higher free-carrier losses. Most of the QCL's operate in a Fabry-Perot cavity formed by cleaving the ends of the wafer and have multimode operation. However, Gmachl *et al.* has obtained single mode operation with gain guided distributed feedback structures operating at 8.5 μm with a CW output power of 10 mW at 120 K. Recently, Faist *et al.* have used oscillator strength tuning to produce a laser tunable from 6.2 – 6.6 μm (100 cm^{-1}) [12]. Tunability has also been obtained in distributed feedback structures over 20 cm^{-1} by varying the operating temperature to change the dielectric properties of the waveguide. In fact, such a laser was recently used by Namjou *et al.* to make the first high sensitivity gas spectroscopy measurements using a QCL[13]. Quantum cascade lasers have proven themselves to be powerful, compact, efficient sources of mid-infrared radiation, useful for many spectroscopic applications.

At this time, only one group outside of Bell Labs has successfully fabricated an electrically pumped quantum cascade laser. Slivken *et al.* used gas source molecular beam epitaxy of InGaAs/InAlAs to fabricate a QCL operating at 8.5 μm [14]. This device displayed peak

pulsed output power of 35 mW per facet at 80 K, and lasing up to 270 K. On a slightly different note, Gauthier-Lafaye *et al.* recently developed an optically pumped quantum well laser in GaAs/AlGaAs quantum wells operating at 15.5 μm [15]. Pulsed operation was demonstrated with a peak output power of 0.4 W at 77 K, and operation up to 110 K. Long wavelength lasers are easier to achieve in optically pumped configurations, since heavily doped contact layers that dominate the loss at long wavelengths are not necessary. This is currently the longest wavelength laser based on quantum wells.

A new type of interband quantum cascade laser based on type-II InAs/GaInSb/AlSb quantum wells was recently demonstrated[16, 17]. In this design, the lasing transition is between a conduction subband in a InAs well and an adjacent valence subband in a GaInSb well. These adjacent wells overlap in energy space, as is characteristic of type-II wells, and allow similar frequency tunability and cascade structures as in traditional unipolar QCLs. The main advantage is that the dominant non-radiative relaxation mechanism, fast LO-phonon scattering, can be circumvented. They recently displayed lasing at 3.9 μm operating in pulsed mode at temperatures up to 170 K with peak power of 0.5 W per facet observed at 80 K ($J_{th} \sim 0.29 \text{ kA/cm}^2$). This material system is somewhat less developed than the more standard GaAs/AlGaAs system.

1.3 Project statement

The goal of this research project is to develop a QCL using GaAs/AlGaAs heterostructures which operates in the atmospheric transmission window of 8-12 μm . At the time of this writing, all operational QCLs have been formed using the InGaAs/AlInAs material system. There are several reasons for this. For the lattice matched InGaAs/InAlAs system, the conduction band discontinuity is $E_c = 520 \text{ meV}$, substantially larger than the discontinuity for $\text{Al}_x\text{Ga}_{1-x}\text{As}$ with feasible Al fractions ($E_c = 320 \text{ meV}$ for $x = 0.4$). These higher barriers allow larger subband separations and thus lasers at higher frequencies. Since the radiation efficiency of an intersubband transition decreases and free carrier losses increase at longer wavelengths, it is advantageous to attempt lasing at higher frequencies. Also, $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ has a smaller effective mass ($m^* = 0.043m_0$), and thus has a larger oscilla-

tor strength for an intersubband transition than one in GaAs ($m^* = 0.067m_0$). Perhaps most importantly, the InGaAs/InAlAs system is grown on InP wafers, which have a lower refractive index than either material. Thus, the substrate provides a natural waveguide cladding. GaAs/AlGaAs heterostructures are always grown on GaAs substrates, which provide no intrinsic confinement.

However, there are several reasons for developing a GaAs/AlGaAs quantum cascade laser. Most importantly, GaAs/AlGaAs molecular beam epitaxy is a mature and widespread technology, and there are a large number of epitaxial machines available for production. It would be advantageous if companies could produce these new lasers using existing facilities. Another advantage is that $\text{Al}_x\text{Ga}_{1-x}\text{As}$ is lattice matched to GaAs for all values of x , so one can vary the well depth and barrier height within some range. This is an extra degree of freedom when designing multiple quantum well structures. Also, GaAs is a binary alloy and hence has higher thermal conductivity than ternary alloys, and hence promises improved high temperature operation.

At the time of this writing, two other groups have made steps towards a GaAs/AlGaAs based quantum cascade laser. Strasser *et al.* [18] and Li *et al.* [19] have both observed electroluminescence in quantum cascade structures. Strasser adopts a structure patterned after the Bell Labs lasers, with a digitally graded region providing carrier injection. Li on the other hand, takes advantage of the lattice matching of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ to GaAs, and substitutes a single wide, shallow well composed of $\text{Al}_{0.15}\text{Ga}_{0.85}\text{As}$ to provide injection into the excited state. The final non-trivial step for these groups in construction of a quantum cascade laser is the design of an appropriate waveguide structure.

1.4 Approach

The goal of this project was to undertake the initial steps in designing a quantum cascade laser, and to fabricate a working laser if possible. A necessary first step in the construction of any laser is to observe spontaneous emission from the gain medium. Following the lead of Bell Labs, a quantum cascade structure was designed to provide a gain medium. The epitaxial growth was performed by Michael Melloch at Purdue University. The MBE layers

were then fabricated into testable structures at the Microsystems Technology Laboratory at MIT. Following construction of an appropriate experimental setup, the current-voltage characteristics and electroluminescence spectrum of the device were measured. These measurements were used to verify the design of the structure and to predict whether enough gain is present to achieve lasing. Following this step, an optical cavity was designed to provide feedback and any modifications necessary were made to the structure. Following growth and fabrication, the device must be tested to see if the structure lases. At present, all the steps have been completed except the final laser testing.

1.5 Thesis overview

Chapter 2 describes some of the background theory necessary for laser design, such as spontaneous emission and stimulated emission rates, basic laser threshold equations, nonradiative relaxation rates, free carrier absorption, and simulation of a one-dimensional waveguide. Chapter 3 describes the design process of the multiple quantum well structure and waveguide structure. Chapter 4 details the measurements made of the structure designed for spontaneous emission. Chapter 5 discusses the results, and concludes with a proposal for a complete laser structure.

Chapter 2

Background - quantum cascade laser theory

2.1 Introduction

The development of a laser requires a medium that will amplify electromagnetic radiation given some spontaneous emission input. When this gain medium is placed inside a confining cavity that provides optical feedback, lasing action will occur. Optical transitions in a QCL take place between subbands in a multiple quantum well (MQW) structure designed to establish a population inversion between the subbands. In parallel to atomic transitions, spontaneous and stimulated transition rates can be calculated to provide an estimate of the gain. This is discussed in Section 2.2 and 2.3. The calculation of cavity loss due to free carrier absorption in the waveguide walls is described in Section 2.4. A brief review of the Drude model for free carrier absorption and optical confinement is provided in Section 2.5. The emission efficiency of the structure is determined by the ratio of the radiative transition rate to the non-radiative transition rate. In a mid-infrared quantum cascade laser, the primary non-radiative relaxation mechanism is scattering due to longitudinal-optical (LO) phonons. The calculation of LO-phonon scattering rates is treated in Section 3.3.2.

2.2 Intersubband radiative transitions

The optical gain for a QCL is provided by stimulated emission from electron transitions between subbands in the MQW structure. 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[20] and Smet[21, 22].

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

$$W_{i \rightarrow f} = \frac{2\pi}{\hbar} \left| \langle f, n_{\mathbf{q}, \sigma} | H' | i, m_{\mathbf{q}, \sigma} \rangle \right|^2 \delta(E_f(\mathbf{k}_f) - E_i(\mathbf{k}_i) \pm \hbar\omega_{\mathbf{q}}), \quad (2.1)$$

where

$$H' = -\frac{e}{m^*} \mathbf{A} \cdot \mathbf{p} \quad (2.2)$$

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 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, ignoring the barrier material, as the electron spends the majority of its time there. We also neglect for the moment the effects of nonparabolicity of the band structure, as this will modify the oscillator strengths and selection rules. More complete expressions including the spatial dependence of the effective mass and nonparabolicity are discussed in Section 3.2.4.

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

$$\mathbf{A} = \sqrt{\frac{\hbar}{2\epsilon\omega_{\mathbf{q}}V}} \hat{\mathbf{e}}_{\mathbf{q}, \sigma} \left[a_{\mathbf{q}, \sigma} e^{i\mathbf{q}\cdot\mathbf{r}} + a_{\mathbf{q}, \sigma}^\dagger e^{-i\mathbf{q}\cdot\mathbf{r}} \right]. \quad (2.3)$$

In this expression, ϵ 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:

$$|\langle H'_{i \rightarrow f} \rangle|^2 / C = m_{\mathbf{q},\sigma} \delta_{m_{\mathbf{q},\sigma}-1, n_{\mathbf{q},\sigma}} \left| \langle f | e^{i\mathbf{q}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{p} | i \rangle \right|^2 \\ + (m_{\mathbf{q},\sigma} + 1) \delta_{m_{\mathbf{q},\sigma}+1, n_{\mathbf{q},\sigma}} \left| \langle f | e^{-i\mathbf{q}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{p} | i \rangle \right|^2, \quad (2.4)$$

$$C = \frac{e^2 \hbar}{2m^* \epsilon \omega_{\mathbf{q}} V}. \quad (2.5)$$

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 in the linear model of the Fermi golden rule, and will not appear unless higher order terms in the interaction Hamiltonian are 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 and the cell periodic Bloch function.

$$\langle \mathbf{r} | i \rangle = \frac{1}{\sqrt{S_{xy}}} e^{i\mathbf{k}_{i,\perp} \cdot \mathbf{r}_{\perp}} \psi_i(\mathbf{k}_{i,\perp}, z) U_0(\mathbf{r}) \quad (2.6)$$

$$\langle \mathbf{r} | f \rangle = \frac{1}{\sqrt{S_{xy}}} e^{i\mathbf{k}_{f,\perp} \cdot \mathbf{r}_{\perp}} \psi_f(\mathbf{k}_{f,\perp}, z) U_0(\mathbf{r}), \quad (2.7)$$

where S_{xy} is the in-plane area. The longitudinal envelope functions $\psi_i(\mathbf{k}_{i,\perp}, z)$ and $\psi_f(\mathbf{k}_{f,\perp}, z)$ are calculated from the 1D Schrödinger wave equation as described in Section 3.2.2.

Since we are considering mid-infrared radiation ($\lambda \approx 10 \mu\text{m}$), 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

$$\langle f | e^{\pm i\mathbf{q}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{p} | i \rangle \cong \langle f | \hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{p} | i \rangle, \quad (2.8)$$

$$= [\hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \mathbf{k}_{i,\perp}] \frac{\hbar}{S_{xy}} \int d\mathbf{r}_{\perp} e^{i(\mathbf{k}_{i,\perp} - \mathbf{k}_{f,\perp}) \cdot \mathbf{r}_{\perp}} \int dz \psi_f^*(\mathbf{k}_{f,\perp}, z) \psi_i(\mathbf{k}_{i,\perp}, z) \\ - \hat{\mathbf{z}} \frac{i\hbar}{S_{xy}} \int d\mathbf{r}_{\perp} e^{i(\mathbf{k}_{i,\perp} - \mathbf{k}_{f,\perp}) \cdot \mathbf{r}_{\perp}} \int dz \psi_f^*(\mathbf{k}_{f,\perp}, z) \frac{\partial \psi_i(\mathbf{k}_{i,\perp}, z)}{\partial z}, \quad (2.9)$$

$$= [\hat{\mathbf{e}}_{\mathbf{q},\sigma} \cdot \hat{\mathbf{z}}] \delta_{\mathbf{k}_{f,\perp}, \mathbf{k}_{i,\perp}} \langle \psi_f | p_z | \psi_i \rangle. \quad (2.10)$$

Note that in the previous expression, the cell periodic part $U_0(\mathbf{r})$ was dropped from the integration, as it is orthonormal over a unit cell. The first term of (2.9) vanishes due to the orthonormality of the envelope functions. This is the well known intersubband selection rule: only transitions with the \mathbf{E} field polarized along the growth axis z are permitted. Note that in the presence of band nonparabolicity, this selection rule is relaxed due to mixing with the valence band. This is discussed in Section 3.2.4. The delta function in (2.10) ensures conservation of in-plane momentum over a transition.

Equation (2.10) 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

$$\frac{i}{\hbar}[H_0, z] = \frac{p_z}{m^*}. \quad (2.11)$$

We then obtain

$$\langle f | p_z | i \rangle = \frac{im^*}{\hbar}(E_f - E_i) \langle \psi_f | z | \psi_i \rangle. \quad (2.12)$$

where the integral $\langle \psi_f | z | \psi_i \rangle$ is the dipole matrix element $z_{i \rightarrow f}$. Inserting this matrix element into (2.5) and (2.1) give the following spontaneous and stimulated transition rates:

$$W_{i \rightarrow f / mode}^{(sp)} = \frac{\pi e^2 \omega_{\mathbf{q}}}{\epsilon V} |\hat{\mathbf{e}}_{\mathbf{q}, \sigma} \cdot \hat{\mathbf{z}}|^2 |z_{i \rightarrow f}|^2 \delta(E_f - E_i - \hbar \omega_{\mathbf{q}}), \quad (2.13)$$

$$W_{i \rightarrow f / mode}^{(st)} = \frac{\pi e^2 \omega_{\mathbf{q}}}{\epsilon V} |\hat{\mathbf{e}}_{\mathbf{q}, \sigma} \cdot \hat{\mathbf{z}}|^2 |z_{i \rightarrow f}|^2 \delta(E_f - E_i - \hbar \omega_{\mathbf{q}}) m_{\mathbf{q}, \sigma}, \quad (2.14)$$

where the delta functions conserving in-plane momentum and photon number have been dropped to simplify notation. To obtain the total spontaneous emission rate, we must sum (2.13) 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

$$\rho(\mathbf{q}) d^3\mathbf{q} = \frac{d^3\mathbf{q}}{8\pi^2/V} = \frac{q^2 dq \sin \theta d\theta d\phi V}{8\pi^3}, \quad (2.15)$$

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

$$\begin{aligned}
W_{i \rightarrow f}^{(sp)} &= \frac{e^2}{8\pi^2\epsilon} |z_{i \rightarrow f}|^2 \int_0^\infty \int_0^\pi \int_0^{2\pi} \left(\frac{n\omega_{\mathbf{q}}}{c}\right)^3 \sin^3 \theta \delta(E_f - E_i + \hbar\omega_{\mathbf{q}}) d\hbar\omega d\theta d\phi \\
&= \frac{e^2 n \omega_0^3}{3\pi\epsilon_0 \hbar c^3} |z_{i \rightarrow f}|^2 = \frac{e^2 n \omega_0^2}{6\pi m^* \epsilon_0 c^3} f'_{i \rightarrow f},
\end{aligned} \tag{2.16}$$

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

$$f'_{i \rightarrow f} = \frac{m^*}{m_0} f_{i \rightarrow f} = \frac{2m^* (E_f - E_i) |z_{i \rightarrow f}|^2}{\hbar^2}. \tag{2.17}$$

This scaled oscillator strength $f'_{i \rightarrow f}$ is a dimensionless quantity defined as the ratio of the quantum radiative transition rate to three times the radiative rate of a classical dipole oscillating at the same frequency. Since it is proportional to the gain, it is convenient to use it to measure the strength of a transition. 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:

$$\sum_f f'_{i \rightarrow f} = 1. \tag{2.18}$$

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 nonparabolicity are discussed in Sirtori *et al.* [23].

The expression (2.16) 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[24, 21]. However, for our design, the wavelength of the radiation in the semiconductor is $\lambda_{GaAs} \sim 3 \mu\text{m}$, about the same as the thickness of the active region, so any reductions in $\tau_{32}^{(sp)}$ will be modest. We use the 3D mode density in our gain calculations to find an upper bound on the radiative lifetime.

The stimulated emission rate can be found by using the single mode transition rate

(2.14) to examine the interaction of the dipole with a monochromatic incident wave at frequency ν , mode ℓ . Since each energy level has a certain finite width due to its finite lifetime, it is useful to introduce a normalized lineshape function $g(\nu)$ for the transition. A homogeneously broadened transition will have the Lorentzian lineshape

$$g(\nu) = \frac{(\Delta\nu/2\pi)}{(\nu - \nu_0)^2 + (\Delta\nu/2)^2}, \quad (2.19)$$

where $\Delta\nu = (\pi T_2)^{-1}$ is the full-width half maximum linewidth of the transition centered about ν_0 . T_2 is the total phase breaking time given by

$$T_2^{-1} = \sum_i \tau_i^{-1}, \quad (2.20)$$

where τ_i are the characteristic times of all collisions and transitions that interrupt the coherent atom-field interaction.

The stimulated emission rate is given by

$$W_{i \rightarrow f}^{(st)} = \frac{\eta \lambda^2 I_\nu}{8\pi h \nu n^2 \tau_{i \rightarrow f}^{(sp)}} g(\nu), \quad (2.21)$$

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

$$I_{\nu,\ell} = \frac{cm_\ell h \nu_\ell}{nV}, \quad (2.22)$$

with m_ℓ photons in the mode of frequency ν_ℓ . The interaction of the incident field with the atom-like dipoles is described by the parameter

$$\eta = \frac{3|z_{i \rightarrow f}|^2}{|\mathbf{r}_{i \rightarrow f}|^2}. \quad (2.23)$$

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 Fundamental laser theory

2.3.1 Optical gain

We will now use the expressions for spontaneous and stimulated emission derived above to consider the optical gain in our laser medium. All lasers are based around a two-level system, with other levels present only to create the desired population conditions on the principal levels at energies E_1 and E_2 . In the presence of a field at frequency ν , 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

$$\frac{\text{Power}}{\text{Volume}} = [N_2 W_{21} - N_1 W_{12}] h\nu, \quad (2.24)$$

where $W_{12}^{(st)} = W_{21}^{(st)}$ is proportional to the field intensity I_ν as seen in (2.21). If $N_2 > N_1$, the induced power will be added to the inducing wave I_ν so that it will grow as it propagates according to

$$\frac{dI_\nu}{dz} = \gamma(\nu) I_\nu. \quad (2.25)$$

We can insert the stimulated transition rate (2.21) into (2.24) 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 ΔN is

$$\gamma(\nu) = \frac{\Delta N e^2 \pi \nu_0 |z_{i \rightarrow f}|^2}{\hbar c n \epsilon_0} g(\nu), \quad (2.26)$$

$$= \frac{\Delta N e^2 f'_{i \rightarrow f}}{4m^* c n \epsilon_0} g(\nu). \quad (2.27)$$

The population inversion density is commonly expressed in terms of $\Delta N = \Delta N^{(2D)} / a_{mod}$, where $\Delta N^{(2D)}$ is the two-dimensional population inversion density, and a_{mod} is the length of a module. Neglecting nonparabolicity, both the initial and final subbands have the same density of states, thus we can neglect differences in level degeneracy.

2.3.2 Optical loss

Radiation losses in a laser cavity can be separated into cavity losses α_c occurring along the propagation length, and facet losses α_f occurring from imperfect reflection at the ends of the laser. In a quantum cascade laser, the primary contribution to cavity loss is from free carrier absorption in the doped contact layers. This loss results from an electron absorbing making an upward intraband transition due to absorption of a photon, and absorption or emission of a phonon. In this study, we largely neglect the quantum aspects of the problem, and approximate the effects of free carriers using the Drude model, as discussed in Section 2.5. The overall cavity loss of any given waveguide is simulated using a one-dimensional slab waveguide model with complex permittivities determined by the Drude model, and is discussed below in Section 2.4.

We can neglect absorption due to phonon interactions, as our frequency of interest is much higher than the optical phonon resonance. Free carriers are present in the active region as well as the doped contacts, although we neglect their contribution to the loss. There are several reasons for this. First, the overall doping level in the wells ($1.5 \times 10^{11} \text{ cm}^{-2}$ over 500 \AA , corresponding to a 3D density of $3 \times 10^{16} \text{ cm}^{-3}$) is quite low compared to that in the bulk cladding. Second, the wells are modulation doped, with most of the carriers residing far from the fast scattering doped impurities. Third, the carriers are confined in two-dimensional subbands, perpendicular to the direction of the z polarized field. They are no longer free carriers in the z direction and cannot interact very strongly with a z -polarized electric field.

2.3.3 Threshold conditions

In order to obtain lasing, peak gain must exceed total losses. The threshold condition for lasing is achieved when the wave reproduces itself after one round trip. That is, for a cavity mode ℓ to sustain oscillation, the round trip phase delay must be a multiple of 2π , and the gain at that mode frequency ν_ℓ must satisfy

$$R_1 R_2 e^{2(\gamma(\nu_\ell) - \alpha_c)L} = 1, \quad (2.28)$$

where α_c is the cavity loss, L is the cavity length, and R_1 and R_2 are the facet reflectivities. If the facet loss is prorated per unit length as $\alpha_f = -\ln(R_1 R_2)/2L$ then the threshold small-signal gain can be written as

$$\gamma_{th} = \alpha_c + \alpha_f = \alpha. \quad (2.29)$$

One can use this value for the gain to calculate the threshold value for the population inversion ΔN_{th} .

For a homogeneously broadened transition, the laser will oscillate at only one mode ℓ , the mode that has the highest $\gamma(\nu_\ell) - \alpha$, with a frequency [20]

$$\nu \approx \nu_\ell - (\nu_\ell - \nu_0) \frac{\Delta\nu_{1/2}}{\Delta\nu}, \quad (2.30)$$

where $\Delta\nu$ is the FWHM gain linewidth, and $\Delta\nu_{1/2}$ is the FWHM of the resonator cavity:

$$\Delta\nu_{1/2} = \frac{1}{2\pi\tau_p} = \frac{c(1 - R_1 R_2)}{2\pi n L}. \quad (2.31)$$

In this expression, τ_p is the photon lifetime, i. e. the time it takes a photon to leak out of the cavity. If the cavity modes are spaced sufficiently close together, the gain at the oscillation frequency can be considered equal to the peak bulk gain $\gamma(\nu_\ell) = \gamma(\nu_0)$.

2.3.4 Gain saturation

When the lasing threshold is reached, the intensity of the field will rapidly increase, clamping the upper state population just below threshold. The basic dynamics of this laser behavior can be analyzed by using a simple rate equation approach common in the literature[25, 26]. While this approach lacks the phase information present in a density matrix approach, it is useful for observing qualitative laser behavior above and below the threshold. We write a pair of coupled rate equations for the population of the upper state N_2 , and the photon number m in the lasing mode. The depopulation of the lower level is assumed to be

sufficiently fast so that $N_1 \approx 0$. We obtain

$$\begin{aligned}\frac{dm}{dt} &= KN_2(m+1) - \frac{m}{\tau_p} \\ \frac{dN_2}{dt} &= R_p - KN_2m - \frac{N_2}{\tau_2},\end{aligned}\quad (2.32)$$

where R_p is the pump rate, τ_2 is the upper state lifetime due to nonradiative transitions, and τ_p is the photon lifetime. K is the stimulated transition rate per incident photon given by

$$K = W_{21}^{(st)}/m = \eta W_{21}^{(sp)}/p. \quad (2.33)$$

where $\eta = 3$ is defined in (2.23), and p is the number of cavity modes. The assumption has been made in writing (2.33), that spontaneous emission rate is essentially the same to each mode. While this is not strictly true, it leads to qualitatively correct results for the following calculations.

We can examine the steady state solutions to (2.32) to examine the behavior above and below the threshold[25].

$$m = \frac{N_2}{(\tau_p K)^{-1} - N_2} = \frac{N_2}{N_{2,th} - N_2}, \quad (2.34)$$

$$N_2 = \frac{R_p}{\tau_2^{-1} + Km} = R_p \tau_2 \times \frac{1}{1 + (\tau_2/\tau_{21})^{-1} \times (m/p)}. \quad (2.35)$$

Below the threshold, we see that the number of photons m in the mode remains small until N_2 attains a significant fraction of the threshold value. Also, for $\tau_2 \ll \tau_{21}^{(sp)}$, which is the case for our QCL as well as many other lasers, we see that N_2 essentially varies linearly with the pump power R_p .

We can rearrange the steady state solutions to shed light on the above threshold case:

$$N_2 = (\tau_p K)^{-1} \times \frac{m}{m+1} = N_{2,th} \times \frac{m}{m+1}, \quad (2.36)$$

$$m = \frac{R_p - N_2/\tau_2}{KN_2} = \frac{\tau_2 p}{\tau_{21}^{(sp)}} \left[\frac{N_{2,th}}{N_2} r - 1 \right], \quad (2.37)$$

where r is a normalized pumping rate $r \equiv R_p/R_{p,th}$. After the lasing threshold is reached,

the population inversion is “locked” just below its threshold value. The photon number m is on the order of p , which is much larger than its value below the threshold. Any additional pumping power goes into increasing the stimulated emission rate, resulting in a constant N_2 and an optical power that increases proportionally.

2.4 Cavity loss - slab waveguide

To achieve lasing, an appropriate waveguide must be designed to confine the radiation and provide optical feedback. The design of a waveguide structure should maximize the quantity $(\gamma - \alpha)$, where $\gamma = \Gamma\gamma_{bulk}$ is the modal gain, and the total α is given by the sum of the cavity loss α_c and the facet loss α_f . The facet loss can be minimized by making the cavity long. The cavity loss is principally determined by free carrier absorption in the contact and cladding layers.

To design a structure that maximizes confinement and minimizes cavity loss, a one-dimensional slab waveguide with $N + 2$ layers was solved. The general form of the problem is of a layered dielectric medium with infinite transverse extent carrying the optical wave. The effect of free carriers in each layer was taken into account by using complex permittivities given by the Drude model (discussed in Section 2.5). For our mid-infrared laser design, the wavelength in the semiconductor ($\sim 3 \mu m$) is somewhat smaller than the width of a typical bar structure ($\sim 10 \mu m$). Thus it is reasonable to use a slab waveguide for calculations. Lateral optical confinement is provided by oxide or air dielectric confinement from the mesa edge. With no free carriers, the lateral confinement should be minimally lossy and the large refractive index difference between GaAs and air (or oxide) should provide good confinement. Due to the dipole selection rule for intersubband transitions, all emitted radiation will be polarized with the E field along the growth direction, thus only TM modes are expected to be excited.

The waveguide geometry is shown in Fig. 2-1, with the propagation along \hat{y} , and infinite extent in the x direction. This coordinate system is different than that typically used in electromagnetic texts, where the direction of propagation is along \hat{z} . This particular set of coordinates is chosen to be consistent with the convention that the MBE growth direction

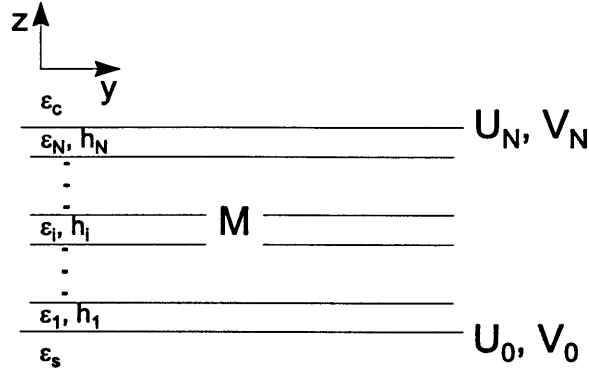


Figure 2-1: Geometry of $N + 2$ layer slab waveguide.

is along z . The lowest layer, known as the substrate, and the uppermost layer, known as the cover, sandwich N layers. The power loss for each guided mode is given by $2\text{Im}\{\beta\}$, where β is the complex propagation constant of the mode. The derivation that follows is quite general, but the formalism relating to the multi-layer treatment has been borrowed from Kogelnik[27].

2.4.1 Fundamental equations

The electric field in the waveguide has the following harmonic form:

$$\tilde{\mathbf{E}}(t) = \mathbf{E} \exp(-i\omega t) + \mathbf{E}^* \exp(i\omega t), \quad (2.38)$$

where $\mathbf{E} = \mathbf{E}(x, y, z)$ is the complex amplitude of the field. The boundary conditions require the x and y components of \mathbf{E} and \mathbf{H} to be continuous at each layer interface.

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

$$\mathbf{H}(x, y, z) = \hat{\mathbf{x}} H_{x,\nu}(z) \exp(i\beta_\nu y), \quad (2.39)$$

where ν is the mode index, and β_ν is the propagation constant of the mode. For simplicity, we will generally omit the mode index ν in future equations. This solution must satisfy the

wave equation for H_x

$$\frac{\partial^2 H_x}{\partial z^2} = (\beta^2 - \mu_0 \epsilon \omega^2) H_x. \quad (2.40)$$

There are several types of solutions to this waveguide equation, analogous to the solutions of Schrödinger's equation in quantum mechanics. We are most interested in the guided (bound) modes, where the energy is concentrated 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. Ideally this leakage can be made negligibly small.

2.4.2 Multilayer waveguide

To treat the case of multilayer stacks, it is convenient to use a characteristic matrix formalism for each layer[27]. The stack is composed of N layers sandwiched between the substrate and cover, with indices of $i = 1$ to N beginning from the substrate. Each layer has thickness h_i and permittivity ϵ_i (index n_i). We define two field variables U and V by

$$U = H_x \quad , \quad V = \omega \epsilon_0 E_y. \quad (2.41)$$

These field variables correspond to the quantities which are continuous at layer boundaries. The general solution in each layer that obeys the transverse wave equation (2.40) are given by

$$\begin{aligned} U &= A \exp(i\kappa z) + B \exp(-i\kappa z) \\ V &= -(\kappa/n^2)[A \exp(i\kappa z) - B \exp(-i\kappa z)], \end{aligned} \quad (2.42)$$

where κ is the transverse wavenumber given by

$$\kappa^2 = \mu_0 \epsilon \omega^2 - \beta^2. \quad (2.43)$$

The constants A and B can be replaced by the input values of U and V at each layer. This gives

$$\begin{aligned} A &= \frac{1}{2}(U_0 - V_0(n^2/\kappa)) \\ B &= \frac{1}{2}(U_0 + V_0(n^2/\kappa)). \end{aligned} \quad (2.44)$$

The evolution of the fields through the i th layer can be described by the matrix equation

$$\begin{pmatrix} U_{i-1} \\ V_{i-1} \end{pmatrix} = \mathbf{M}_i \begin{pmatrix} U_i \\ V_i \end{pmatrix} \quad (2.45)$$

where \mathbf{M}_i is the characteristic matrix of the layer. This equation relates the input quantities U_{i-1} and V_{i-1} at the start of the layer with the output quantities U_i and V_i at the interface with layer $i + 1$. The characteristic matrix for each layer is given by

$$\mathbf{M}_i = \begin{vmatrix} \cos(\kappa_i h_i) & i(n_i^2/\kappa_i) \sin(\kappa_i h_i) \\ i(\kappa_i/n_i^2) \sin(\kappa_i h_i) & \cos(\kappa_i h_i) \end{vmatrix}. \quad (2.46)$$

The individual matrices for each layer can be multiplied together to form a characteristic matrix \mathbf{M} for the entire stack. This relates the input variables U_0, V_0 at the substrate to the output variables U_N, V_N at the cover. This is written as

$$\begin{pmatrix} U_0 \\ V_0 \end{pmatrix} = \mathbf{M}_1 \cdot \mathbf{M}_2 \cdot \mathbf{M}_3 \dots \mathbf{M}_N \begin{pmatrix} U_N \\ V_N \end{pmatrix} = \mathbf{M} \begin{pmatrix} U_N \\ V_N \end{pmatrix} \quad (2.47)$$

Applying the appropriate boundary conditions to the characteristic matrix of the stack yields a dispersion relation which can be solved to find the field solution for a multilayer waveguide. For a guided mode, the field decays exponentially away from the core in the substrate and cover layers. This is expressed by writing the fields in each layer as

$$\begin{aligned} U &= A_i \exp(\gamma_i(z - z_i)) + B_i \exp(-\gamma_i(z - z_i)) \\ V &= (i\gamma_i/n_i^2)[A_i \exp(\gamma_i(z - z_i)) - B_i \exp(-\gamma_i(z - z_i))], \end{aligned} \quad (2.48)$$

where the z_i is the z coordinate at the start of each layer. This form of the fields uses the decay constant $\gamma^2 = -\kappa^2$. Using these variables, the decaying boundary conditions are

expressed as

$$U_0 = A_s, \quad V_0 = i\gamma_s/n_s^2; \quad U_N = B_c, \quad V_N = -i\gamma_c/n_c^2, \quad (2.49)$$

where the γ_s and γ_c are the decay constants in the substrate and the cover respectively. Inserting these variables into (2.47) yields the dispersion relation

$$i(M_{11}\gamma_s/n_s^2 + M_{22}\gamma_c/n_c^2) = M_{21} - (\gamma_s\gamma_c/n_s^2n_c^2)M_{12}. \quad (2.50)$$

The solutions of this transcendental equation correspond to the guided modes of the waveguide. For a given value of the propagation constant β , one can use the characteristic matrices to find the field pattern that corresponds to each mode. For lossless guides, β is real. For guides with loss, leakage, or gain, the propagation constant is complex given by

$$\beta = \beta' + i\beta''. \quad (2.51)$$

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

$$\alpha_c = 2\beta'', \quad (2.52)$$

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

2.4.3 Implementation

The above method was implemented in MATLAB to find the loss and field patterns for various waveguide designs. Solving the transcendental equation (2.50) for a complex root β is a non-trivial task. We used a globally convergent Newton-Raphson method for solving nonlinear sets of equations[28]. While this method is generally successful at finding some solution, a good initial guess is necessary to ensure that the correct root corresponding to the mode of interest is found. For the lower order modes, the loss is expected to add only a small perturbation to the guide. Thus, to obtain an estimate for β , we treat the imaginary

components of the permittivities ϵ_i'' as perturbations to the real unperturbed system defined by the real parts ϵ_i' [29]. First, we find the exact solution to the unperturbed problem using the bisection method to find the real root $\beta^{(0)}$. Second, the variational method is used to obtain the imaginary part of β [30]. If the unperturbed field pattern is used as the variational wavefunction, we obtain

$$\beta^2 \cong \beta^{(0)2} + i \sum_{i=1}^N \omega^2 \mu_0 \epsilon_i'' \Gamma_i, \quad (2.53)$$

where Γ_i is the optical confinement factor in each layer. This value is used as the initial guess for β in the Newton-Raphson complex root solver. For most cases, this initial guess is close enough to the root of interest that the complex root solver is well behaved and quickly finds the root. When it fails, it may be necessary work “by hand,” that is, start with a guide with a known solution and gradually change the parameters, constantly substituting the previous solution as the initial guess. Such an incremental process is helpful with the more difficult cases.

2.4.4 Calculation results

Once the complex root β is known, the mode solution can be found. In the majority of cases, we are only interested in finding the TM_0 mode, as higher order modes tend to be much lossier. The loss is given by the complex part of β according to (2.52). Once the field solution is obtained, the loss can be separated into components due to the upper and lower layers. This is done by taking the ratio of the power flow out of the active region interface to the total forward power flow through the active region:

$$\alpha_{\uparrow,\downarrow} = \frac{\langle S_z(z_{\uparrow,\downarrow}) \rangle}{\int_{\text{active}} \langle S_y(z) \rangle dz}, \quad (2.54)$$

where $\langle \mathbf{S} \rangle = (1/2)\text{Re}\{\mathbf{S}\}$ is the time average of the complex Poynting vector \mathbf{S} , and z_{\uparrow} and z_{\downarrow} are the location of the upper and lower interfaces. This method can be used to separate the loss due to each cladding layer. The field solutions are also useful for finding the confinement factor Γ , which indicates the fraction of the mode power that couples with the gain medium. Traditionally, the confinement factor is the ratio of the electromagnetic

energy present in the gain medium to the total energy. However, we define our confinement factor to reflect the fact that our MQW gain medium will only couple with z polarized light:

$$\Gamma = \frac{\int_{active} |E_z|^2 dz}{\int_{-\infty}^{\infty} |E|^2 dz}. \quad (2.55)$$

The modal gain is given by $\gamma = \Gamma\gamma_{bulk}$. Hence, the goal for designing a waveguide structure consists of maximizing the quantity $\Gamma\gamma_{bulk} - \alpha$. The actual design of waveguide structures suitable for quantum cascade lasers is discussed in Section 3.4.

2.5 Drude model

The Drude theory for conductivity is used to obtain the complex permittivities of the semiconductor and metal layers in the waveguide calculation. This theory assumes each independent electron moves with some momentum relaxation time τ . This can lead 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 the wavelength of light λ is longer than the electron mean free path. This assumption doesn't always hold true in high purity, low temperature metals; in these cases a nonlocal theory of conductivity must be used. The other major deviation from the Drude model appears in the frequency dependence of τ . Acoustic and optical phonon scattering mechanisms have separate frequency dependences, as does impurity scattering. Our photon energies ($\hbar\omega = 120$ meV) are much lower than the bandgap ($E_G = 1.423$ meV), so no interband absorption occurs. 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[31]. In the far infrared limit this scattering rate reduces to a frequency independent constant that corresponds to the DC mobility. Jensen's calculations show that scattering is even slightly reduced as frequency increases in the mid-infrared, meaning that use of the DC mobility τ slightly overstates the loss.

The Drude model, described in numerous texts[32], predicts a frequency dependent

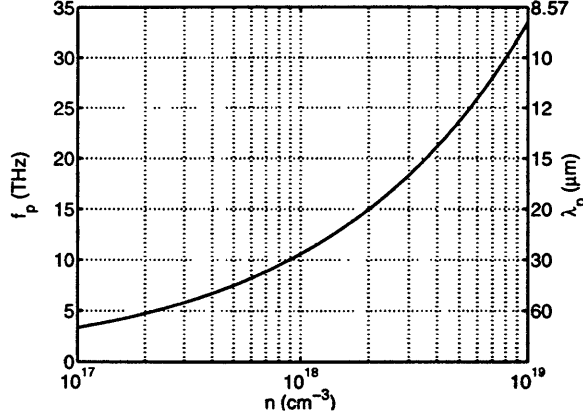


Figure 2-2: Plasma frequency for bulk GaAs as a function of free carrier density n using $\epsilon(\lambda = 10 \mu\text{m}) = 10.71\epsilon_0$. Note that our QCL is designed to operate at approximately 30 THz ($\lambda \sim 10 \mu\text{m}$).

conductivity

$$\sigma(\omega) = \frac{ne^2\tau}{m^*(1 - i\omega\tau)}, \quad (2.56)$$

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

$$\epsilon(\omega) = \epsilon_{core} \left(1 - \frac{\omega_p^2\tau^2}{1 + (\omega\tau)^2} + i\frac{\omega_p^2\tau}{\omega(1 + (\omega\tau)^2)} \right), \quad (2.57)$$

where ϵ_{core} is the core permittivity excluding free carrier contributions, and

$$\omega_p^2 = ne^2/m^*\epsilon_{core} \quad (2.58)$$

defines the plasma frequency for the material. The dependence of the plasma frequency on the free carrier density (doping) is shown in Fig. 2-2.

The buffer layers, which serve as contacts and waveguide layers, contain the majority of the doping in our structure. For n-type, Si doping, bulk GaAs cannot be generally doped higher than $5 \times 10^{18} \text{ cm}^{-3}$ [33]. We observe that $\omega_p < \omega$ for all possible N_D in the frequency range of our interest ($\lambda \sim 10 \mu\text{m}$). For these doping levels, the DC drift mobility of electrons in GaAs is at least $2500 \text{ cm}^2/\text{V s}$, corresponding to a Drude scattering time of

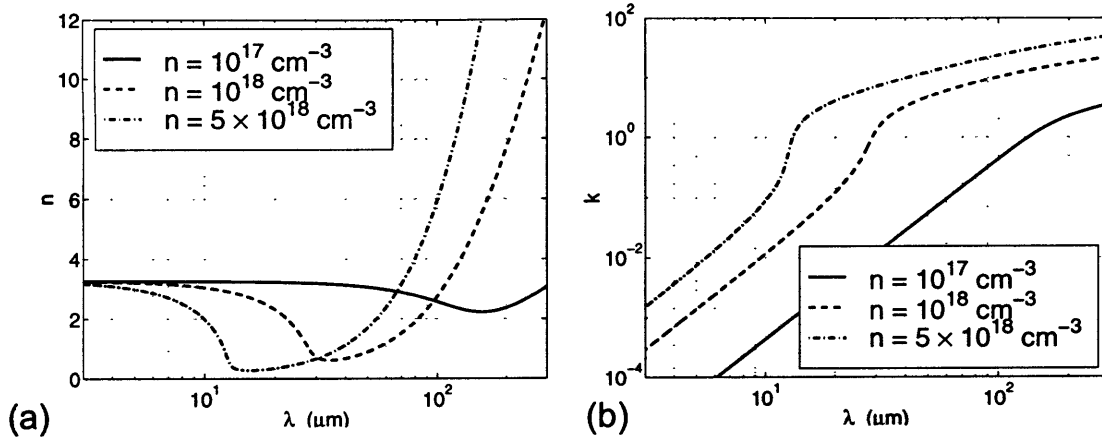


Figure 2-3: Complex refractive index $N = \sqrt{\epsilon/\epsilon_0} = n + ik$ versus wavelength for three free carrier densities. The calculation was performed using a momentum relaxation time of $\tau = 0.1$ ps and a permittivity $\epsilon(\lambda = 10 \mu\text{m}) = 10.71\epsilon_0$

$\tau = 0.1$ ps at 300 K[34]. The scattering rate decreases with reduced doping and at lower temperatures, but we use the value $\tau = 0.1$ ps in our calculations as a worst case. For our lasing wavelength $\lambda \sim 10 \mu\text{m}$ ($f \sim 30$ THz), and typical buffer doping $N_D = 10^{18} \text{ cm}^{-3}$, we can examine (2.57) in the limit $1 \ll \omega_p\tau \ll \omega\tau$. In this region, we have

$$\epsilon = \epsilon_{core} \left(1 - \frac{\omega_p^2}{\omega^2} + i \frac{\omega_p^2}{\omega^3\tau} \right). \quad (2.59)$$

The effect of the free carriers on the complex refractive index is shown in Fig. 2-3. The complex refractive index is given by $N = \sqrt{\epsilon/\epsilon_0} = n + ik$, where n is the real refractive index and k is the extinction coefficient. The reduction of the refractive index in the vicinity of the plasma frequency can be exploited to provide optical confinement by using doped cladding layers.

This will give us a bulk free carrier loss of

$$\alpha_{fc} = \frac{\omega_p^2}{\omega^2} \frac{1}{c\tau} \sqrt{\epsilon_{core}} = \frac{e^2\lambda^2}{4\pi^2 m^* c^3 \epsilon_0 \tau} \sqrt{\epsilon_{core}} \quad (2.60)$$

This bulk free carrier loss is plotted in Fig. 2-4 for several doping densities. The free carrier loss increases sharply as the radiation frequency moves to below the plasma frequency, after which it levels out. For our design wavelength of $\lambda \sim 10 \mu\text{m}$, the loss is in the λ^2 region

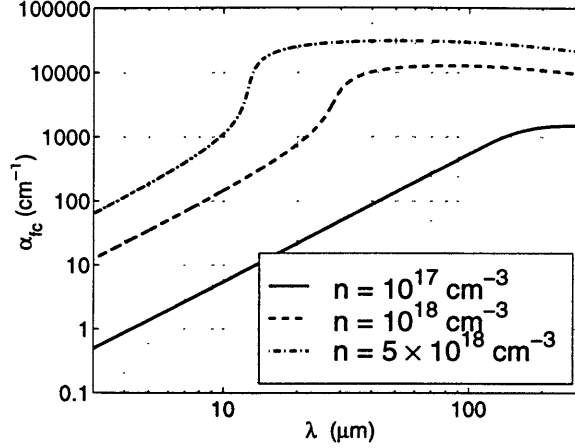


Figure 2-4: *Free carrier loss in GaAs calculated with Drude model using a momentum relaxation time of $\tau = 0.1$ ps and a permittivity of $\epsilon(\lambda = 10 \mu\text{m}) = 10.71\epsilon_0$*

for all possibly bulk dopings. This plot is shown to merely give an approximate view of the loss above and below the plasma frequency. In practice, the wavelength dependence of the various scattering mechanisms will change the ω^{-2} dependence of the loss.

The optical effects of a plasma can be used for optical confinement in a waveguide. Below the plasma frequency ω_p , $\text{Re}\{\epsilon\} < 0$ and carriers screen the field, and the light wave cannot propagate. This is the case for reflection from a metal, where plasma frequencies are typically in the ultraviolet. A heavily doped semiconductor may also have a plasma frequency high enough to provide this effect for far-infrared radiation. Above the plasma frequency, the real part of the permittivity is reduced by the factor $1 - \omega_p^2/\omega^2$ by free carrier effects, and the material can provide quasi-dielectric confinement. At very high frequencies, one is driving the electron gas faster than it can respond, and free carrier contributions to optical properties diminish, hence the ω^{-2} dependence of the loss.

2.6 Non-radiative relaxation (LO-phonon scattering)

Longitudinal-optical phonon scattering is the principal non-radiative relaxation mechanism in a QCL, as described below in Section 3.3.2. It is possible to engineer the subbands to achieve a population inversion by exploiting LO-phonon scattering. The formalism used below to describe this scattering is borrowed from Smet[21, 22]. In the deriva-

tion, 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 [35]. These results compare favorably with a fully microscopic treatment of LO-phonon scattering in a 56 Å GaAs/AlAs quantum well performed by Rucker *et al.* [36]. They found that the LO-phonon emission rates were between the bulk rates for the well and barrier materials, tending towards the bulk values of GaAs and AlAs for the limiting cases of wide wells and narrow wells respectively. In our calculations we merely use bulk phonon modes for GaAs, which provides a reasonable estimate for the scattering times. Electron screening is neglected, although an account is given by Smet in Ref. [21] The optical phonon branch is taken as dispersionless at the frequency ω_{LO} . The scattering rate for an electron initially in state $|i, \mathbf{k}_i\rangle$ (subband i , in-plane wavevector \mathbf{k}_i) to the state $|f, \mathbf{k}_f\rangle$ through an interaction potential H' is evaluated using Fermi's golden rule

$$W_{i \rightarrow f}(\mathbf{k}_i, \mathbf{k}_f) = \frac{2\pi}{\hbar} \left| \langle f, \mathbf{k}_f | H' | i, \mathbf{k}_i \rangle \right|^2 \delta(E_f(\mathbf{k}_f) - E_i(\mathbf{k}_i) \pm \hbar\omega_{LO}). \quad (2.61)$$

The envelope wavefunctions $\psi_i(z)$ for each subband are obtained using SEQUAL as described in Section 3.2.2. The electron-phonon interaction Hamiltonian takes the form of

$$H' = \sum_{\mathbf{q}} \left[\alpha(\mathbf{q}) \left(e^{i\mathbf{q}\cdot\mathbf{r}} b_{\mathbf{q}} + e^{-i\mathbf{q}\cdot\mathbf{r}} b_{\mathbf{q}}^\dagger \right) \right], \quad (2.62)$$

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 by

$$|\alpha(\mathbf{q})|^2 = 2\pi\hbar\omega_{LO} \frac{e^2}{q^2} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right), \quad (2.63)$$

where ϵ_0 and ϵ_∞ are the static and high frequency permittivities. The screening effects of other electrons are taken into account through q_s , the inverse of the 3D screening length. The q^{-2} dependence of this term results in reduced scattering rates for large in-plane momentum transfers, corresponding to scattering between subbands with a large energy separation.

ration.

The matrix element is given by

$$\begin{aligned} \left| \langle f, \mathbf{k}_f | H' | i, \mathbf{k}_i \rangle \right|^2 &= \frac{2\pi e^2 \hbar \omega_{LO}}{V} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right) \frac{1}{q_z^2 + q_\perp^2} |A_{i \rightarrow f}(q_z)|^2 \\ &\times \delta_{\mathbf{k}_i, \mathbf{k}_f \mp \mathbf{q}_\perp} (n_{\omega_{LO}} + 1/2 \mp 1/2), \end{aligned} \quad (2.64)$$

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

$$A_{i \rightarrow f}(q_z) = \int_{-\infty}^{\infty} dz \psi_f^*(z) \psi_i^*(z) e^{\pm q_z z} \quad (2.65)$$

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.

$$k_f^2 = k_i^2 + \frac{2m^*(E_f(0) - E_i(0) \mp \hbar\omega_{LO})}{\hbar^2} \quad (2.66)$$

After summation over the phonon modes \mathbf{q} , the expression (2.61) can be integrated over these final states \mathbf{k}_f to yield the total scattering rate from an initial wavevector:

$$\begin{aligned} W_{i \rightarrow f}^{abs}(\mathbf{k}_i) &= \frac{m^* e^2 \omega_{LO}}{2\hbar^2} (1 - f_f[E_f(k_f)]) \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right) n_{\omega_{LO}} \int_0^{2\pi} d\theta B_{i \rightarrow f}(q_\perp), \\ W_{i \rightarrow f}^{em}(\mathbf{k}_i) &= \frac{m^* e^2 \omega_{LO}}{2\hbar^2} (1 - f_f[E_f(k_f)]) \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right) (n_{\omega_{LO}} + 1) \int_0^{2\pi} d\theta B_{i \rightarrow f}(q_\perp), \end{aligned} \quad (2.67)$$

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

$$B_{i \rightarrow f} = \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} dz' \psi_f^*(z) \psi_i(z) \psi_i^*(z') \psi_f(z') \frac{1}{q_\perp} e^{-q_\perp |z - z'|}. \quad (2.68)$$

The effects of state blocking are taken into account in (2.67) using the Fermi function $f[E_f(k_f)]$ at the final energy. The angle θ is the angle between the in-plane wavevectors \mathbf{k}_i and \mathbf{k}_f . Energy conservation and the in-plane momentum conservation rule allows us to write the phonon wavevector q_\perp in terms of \mathbf{k}_i and \mathbf{k}_f :

$$q_\perp^2 = |\mathbf{k}_i - \mathbf{k}_f|^2 = k_i^2 + k_f^2 - 2k_i k_f \cos \theta. \quad (2.69)$$

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

$$\frac{1}{\tau_{i \rightarrow f}} = \frac{1}{(2\pi/L)^2} \int d^2\mathbf{k}_f f_i[E_i(k_i)] W_{i \rightarrow f}(\mathbf{k}_i), \quad (2.70)$$

where $f_i[E_i(k_i)]$ is the Fermi function at the initial energy. 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 due to population inversion, thus state blocking was not an issue. Intracsubband scattering times can also be calculated using this formalism, simply by setting $i = f$, $\psi_i(z) = \psi_f(z)$, etc.

Chapter 3

Design

3.1 Introduction

The design for the quantum cascade laser can be divided into two parts: the design of the multiple quantum well (MQW) structure that makes up the gain medium, and the design of the electromagnetic cladding used to form an optical cavity necessary for the achievement of a laser.

The design of the quantum well structure involves choosing the thickness of the epitaxial layers and their material composition in terms of the molar Al fraction. The simplest structure to grow is a binary system, where the wells are composed of GaAs, and the barriers are made up of $\text{Al}_x\text{Ga}_{1-x}\text{As}$. The barrier height is determined by x , the molar fraction of aluminum. One advantage of the GaAs/AlGaAs system is that GaAs is lattice matched to $\text{Al}_x\text{Ga}_{1-x}\text{As}$ for all values of x . This adds an extra degree of flexibility when designing a QCL structure. One can vary the well depths or barrier heights, or even use analog grading of the conduction band edge to optimize efficiency. However, utilizing this added flexibility places additional demands on the MBE growth process. Growing analog structures requires precision temperature control of the Al cells, while using multiple well depths requires multiple Ga or Al cells. Due to the nature of available growth facilities, and a desire to keep the growth process as simple as possible, we used a digital design based on GaAs/ $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$ quantum wells.

The choice of $x = 0.4$ is chosen to maximize the barrier height of the quantum wells.

At approximately $x = 0.45$, the Γ point energy moves above the X -point energy, and $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ceases to be a direct bandgap material. We stay well below this crossover point to avoid the fast $\Gamma - X$ scattering that can take place. At the Al concentration $x = 0.4$, the barrier height is $\Delta E_c = 324$ meV, assuming that 65% of the bandgap difference appears in the conduction band.

3.2 SEQUAL

The primary tool used to design the MQW gain structure is a program called SEQUAL. SEQUAL (Semiconductor Electrostatics by QUantum AnaLysis) is a device simulation program obtained from Purdue University that computes the transport characteristics for a given one-dimensional quantum device. We used this program to iteratively solve Schrödinger's equation and Poisson's equation for a series of heterostructures under bias to obtain a self-consistent solution for the electrostatic potential and space charge. SEQUAL first solves Schrödinger's equation for the MQW system to determine the quasi-bound and bound states, and then populates them with electrons according to some specified distribution. Poisson's equation is then solved for the resulting space charge, which yields modifications to the original electrostatic potential. The shape of the potential wells is consequently changed, and Schrodinger's equation is re-solved with this new potential profile. This process is repeated until the solution converges.

3.2.1 Theoretical formulation and propagating states

SEQUAL makes several simplifying assumptions for its calculations. Electrons are assumed to be majority carriers, and the effects of holes are ignored. In each material layer, the band structure is described by an effective mass, which is assumed to change abruptly at each material interface. However, this is not strictly true, as SEQUAL was recently modified to take into account the effects of nonparabolicity in the band structure. This perturbation is discussed in Section 3.2.4. Except for the self consistent potential, electron-electron interactions are neglected. Transport is assumed to be ballistic. The effects of scattering on the electron populations must be calculated externally via rate equations and

inserted into SEQUAL.

The device to be simulated in SEQUAL is input in terms of material layers in the longitudinal direction z . For each layer we specify the conduction band edge $E_c(z)$ (including bias), effective mass $m^*(z)$, dielectric constant $\epsilon(z)$, and doping profile $N_D(z)$. As described in the SEQUAL manual, the analysis proceeds as follows[37]. Contacts at either end inject electrons into the device with a spectrum of wavenumbers \mathbf{k} . Some electrons are transmitted through the device potential profile E_c , and the remainder are reflected back to the contact. Electrons populate the entire continuum of these “propagating states,” although not with equal probability. The presence of quantum wells results in the appearance of wavefunction resonances at certain energies, where transmission through the structure approached unity. These are known as “quasi-bound states,” and a substantial number of electrons can resonantly accumulate in each level before being transmitted through the device. Since scattering is forbidden, states below the band edge of the contact are inaccessible to electrons.

These propagating states of electrons are quantum mechanically described by the time-independent Schrödinger-like effective mass equation for an envelope function $\psi_{\mathbf{k}}(z)$:

$$\frac{\partial}{\partial z} \left(\frac{m_c^*}{m^*(z)} \frac{\partial}{\partial z} \psi_{\mathbf{k}}(z) \right) + \frac{2m_c^*}{\hbar^2} \left[E_z + E_t \left(1 - \frac{m_c^*}{m^*(z)} \right) - E_c(z) \right] \psi_{\mathbf{k}}(z) = 0 \quad (3.1)$$

where E_z is the longitudinal electron energy, and

$$E_t = \frac{\hbar^2}{2m_c^*} (k_x^2 + k_y^2) \quad (3.2)$$

is the transverse electron energy. This form of Schrödinger’s equation accounts for spatial variation of the effective mass $m^*(z)$ with respect to the effective mass of the contact m_c^* .

SEQUAL can solve for an arbitrary potential profile by subdividing the device into small nodes over which the potential is roughly constant and well described by its average value. Inside each region of constant potential, the solution is given by some superposition of plane waves. To preserve charge and current continuity, the solution requires that $\psi_{\mathbf{k}}(z)$ and $\frac{1}{m^*(z)} \frac{\partial}{\partial z} \psi_{\mathbf{k}}(z)$ be continuous at each node junction.

The electron density in the device due to propagating states is given by sum of the

contribution from each contact

$$n(z) = n^{l \rightarrow r}(z) + n^{r \rightarrow l}(z). \quad (3.3)$$

Without scattering, the two contributions can be accounted for separately. The electron densities are obtained by integrating the magnitude squared of the wavefunction over all k :

$$n^{l \rightarrow r}(z) = \int_0^\infty \frac{dk_z}{2\pi} |\psi_{\mathbf{k}}^{l \rightarrow r}(z)|^2 \sigma^{l \rightarrow r}(k_z) \quad (3.4)$$

$$n^{r \rightarrow l}(z) = \int_0^\infty \frac{dk_z}{2\pi} |\psi_{\mathbf{k}}^{r \rightarrow l}(z)|^2 \sigma^{r \rightarrow l}(k_z) \quad (3.5)$$

where the factors $\sigma(k_z)$ represent the integration over transverse wavevectors k_x and k_y :

$$\sigma^{l \rightarrow r}(k_z) = \frac{m_c^* k_B T}{\pi \hbar^2} \ln \left[1 + \exp\left[\frac{E_F - E_c(0) - \frac{\hbar^2 k_z^2}{2m_c^*}}{k_B T}\right] \right] \quad (3.6)$$

$$\sigma^{r \rightarrow l}(k_z) = \frac{m_c^* k_B T}{\pi \hbar^2} \ln \left[1 + \exp\left[\frac{E_F - E_c(N) - \frac{\hbar^2 k_z^2}{2m_c^*}}{k_B T}\right] \right]. \quad (3.7)$$

E_F is the Fermi level in each injecting contact, $E_c(0)$ and $E_c(N)$ are the conduction band edge energies at each contact, and T is the temperature. It is assumed in this derivation that the wavefunction is approximately constant in the transverse directions and can therefore be pulled out of the integral. SEQUAL can also use a similar expression to (3.5) to compute the current in the ballistic case. However, for our structure, scattering cannot be neglected in current calculations and we do not use SEQUAL's results for the current density.

It is also useful to examine the transmission coefficients in each direction. These are proportional to the squared-magnitude of the wavefunction at the opposite end of the device:

$$T^{l \rightarrow r}(k_z) \propto |\psi_{\mathbf{k}}^{l \rightarrow r}(N)|^2, \quad T^{r \rightarrow l}(k_z) \propto |\psi_{\mathbf{k}}^{r \rightarrow l}(0)|^2 \quad (3.8)$$

Quasi-bound (resonance) states can be identified by locating the Lorentzian peaks in this transmission function expressed as a function of energy. The energy-time uncertainty prin-

ciple can be used to estimate the tunneling lifetime of each state. This is given by

$$\tau_{tun} \cong \frac{\hbar}{\Delta E} \quad (3.9)$$

where ΔE is the FWHM of the resonance peak in energy space.

SEQUAL performs the numerical integration as follows. First, a uniform grid of wavevectors is used to find local maxima in the transmission function (3.8) corresponding to the quasi-bound states. These maxima are used to define integration intervals over which the wavefunctions are known to be smooth and well behaved. It is important to locate each maximum point exactly, as some resonance peaks are extremely sharp and would be missed in a simple examination of the grid values. Successive orders of Gauss-Legendre quadrature are applied over each interval until the result is specified to the desired accuracy, generally three significant digits.

3.2.2 Bound states

SEQUAL also solves for the states that cannot be populated by either contact, i.e. “bound states.” These states reside below the conduction band edges of either contact and are stationary rather than propagating. In an actual device, such states could be populated by electrons inelastically scattered from higher energies. Unless otherwise specified, SEQUAL will populate these states according to Fermi-Dirac statistics.

The bound states are obtained by finding the eigenvalues of Schrödinger’s equation (3.1) according to a finite-difference method. The wavefunction $\psi_i(z)$ at each eigenenergy ϵ_i is found and used to obtain the electron population in each state:

$$n_i(z) = \frac{m^*(z)k_B T}{\hbar^2 \pi} |\psi_i(z)|^2 \ln [1 + \exp[(E_F - \epsilon_i)/k_B T]]. \quad (3.10)$$

The use of Fermi-Dirac statistics is only valid in equilibrium. This is of course not true during the laser operation, so the scattering rates and rate equations must be used to calculate the population of each level, as described in Section 3.3.2.

In the case where a device consists of some MQW structure surrounded by thick bar-

riers, the resonance of the quasi-bound states will become very sharp, and the lifetimes very long. In the limiting case where the barriers become infinitely thick, these states will become true bound states. One can run a simulation where the contacts are replaced by barrier material, and SEQUAL will find bound states closely corresponding to the quasi-bound states. For propagating states SEQUAL only yields the squared-magnitude wavefunction $|\psi_{\mathbf{k}}(z)|^2$; for the bound state case, the wavefunction is real, and the program yields $\psi_i(z)$. The similarity between the quasi-bound and bound states can be exploited to obtain an approximation to the quasi-bound state wavefunction. This is necessary for calculating such quantities as dipole moments, and scattering rates. The validity of the approximation depends on the details of the structure and the true lifetime of the quasi-bound state.

3.2.3 Self-consistency

The preceding discussion assumes that the shape of the conduction band $E_c(z)$ is known. However, as the subbands are populated with electrons, space charge builds up and distorts the potential according to the Poisson equation

$$\frac{d}{dz} \left(\epsilon(z) \frac{d}{dz} \Phi(z) \right) = -q \left(N_D^+(z) - n(z) \right), \quad (3.11)$$

where $\Phi(z)$ is the electrostatic potential, and $N_D^+(z)$ is the ionized donor density. This results in a conduction band profile

$$E_c(z) = \Delta E_c(z) - q\Phi(z), \quad (3.12)$$

where $\Delta E_c(z)$ describes the conduction band edge offset due to the material layers.

For any device, the solutions for the electron density and electrostatic potential should be self-consistent. This is achieved by iteratively solving the Schrödinger and Poisson equations until the solution converges.

3.2.4 Nonparabolicity

Theory

By default, SEQUAL uses the effective mass approximation to describe conduction band shape around the Γ point as parabolic. While the validity of this approximation clearly depends on the actual shape of the conduction band, this approximation can fail when dealing with the upper subbands in a quantum well which require larger values of k_z . B. Riely recently modified SEQUAL to better handle the effects of band nonparabolicity. An empirical two-band model was employed, allowing the effects of non-parabolicity to be incorporated via an energy dependent effective mass $m^*(z, E)$, as described by Nelson *et al.* and Leavitt[38, 39]. This treatment allows the use of the traditional one-band effective mass Schrödinger equation (3.1), solved iteratively to find the eigenvalues.

Nonparabolicity near a band extremum can be described by the dispersion relation

$$E = \frac{\hbar^2 k^2}{2m^*} (1 - \gamma k^2) \quad (3.13)$$

where E and k are the energy and wave number of the charge carrier, and γ is the bulk nonparabolicity parameter. Nelson *et al.* base their work upon the Bastard derivation[40, 41] of an envelope function model which considers the interaction of the conduction band and the light-hole bands throughout a heterostructure. Such a model produces the same effective mass m^* for each band and the same nonparabolicity parameter γ for each band. The energy dependent effective mass is written as

$$m^*(E) = m^* \left(1 + \frac{2m^*\gamma(E - E_c)}{\hbar^2} \right), \quad (3.14)$$

where the nonparabolicity parameter is given by

$$\gamma = \frac{\hbar^2}{2m^*E_g} \quad (3.15)$$

and $E_g = E_c - E_v$ is the energy gap. For GaAs as well as most materials, this model is incomplete; each band has a different effective mass and nonparabolicity parameter. This

feature will only appear when more than two bands are included. However, the two-band treatment contains the relevant physics for treating nonparabolicity and can be used once modified to allow for these experimentally observed differences.

Nelson *et al.* take an “empirical” approach where m^* and γ are chosen separately for each band according to experimental results. The two-band Hamiltonian for treating the conduction band is written as

$$H = \begin{bmatrix} E_c & \frac{i\hbar^2 k}{2m^* \sqrt{\gamma}} \\ \frac{-i\hbar^2 k}{2m^* \sqrt{\gamma}} & E_c - \frac{\hbar^2}{2m^* \gamma} \end{bmatrix}. \quad (3.16)$$

In this “empirical” method, the energy gap $\hbar^2/2m^*\gamma$ will likely not be equal to the actual energy gap E_g . In fact the second band doesn’t correspond to the “true” light-hole band, but rather an “effective” band separated from the conduction band by an “effective” bandgap that represents the collective interactions of all the valence bands with the conduction band. If desired, a separate calculation can be used for the light-hole band, using a different effective mass, effective bandgap, and nonparabolicity parameter. At a material interface, such as in a quantum well, continuity of current requires $m^* \sqrt{\gamma}$ to remain constant across the interface.

Leavitt expanded this theory to heterostructures in the presence of an applied field. The energy dependent effective mass merely becomes position dependent as:

$$m^*(E, z) = m^*(z) \left(1 + \frac{2m^*(z)\gamma(z)(E - E_c(z))}{\hbar^2} \right), \quad (3.17)$$

where $E_c(z)$ contains the effect of the electrostatic potential as in (3.12). The z dependence in $m^*(z)$ and $\gamma(z)$ is due only to the material variations in the heterostructure layers. Now the Hamiltonian (3.16) can be altered and rewritten to yield Schrödinger’s equation:

$$\begin{bmatrix} E_c(z) - E & \frac{i\hbar^2}{2m^*(z)\sqrt{\gamma(z)}} \frac{d}{dz} \\ \frac{-i\hbar^2}{2m^*(z)\sqrt{\gamma(z)}} \frac{d}{dz} & E_c - \frac{\hbar^2}{2m^*(z)^2 \gamma(z)} m^*(z, E) \end{bmatrix} \begin{pmatrix} \psi_c(z) \\ \psi_v(z) \end{pmatrix} = 0, \quad (3.18)$$

where the total wavefunction of this solution is given by

$$\Psi = \begin{pmatrix} \psi_c(z) \\ \psi_v(z) \end{pmatrix}, \quad (3.19)$$

where $\psi_c(z)$ is the principal component from the conduction band and a small effective valence band component.

Taking the lower component of (3.18), we obtain an expression for the effective valence wavefunction in terms of the $\psi_c(z)$:

$$\psi_v(z) = -\frac{m^*(z)\sqrt{\gamma(z)} \frac{d\psi_c(z)}{dz}}{m^*(z, E)}. \quad (3.20)$$

Plugging this into the upper portion, we obtain a second order differential equation for $\psi_c(z)$:

$$\frac{d}{dz} \left(\frac{1}{m^*(z, E)} \frac{d\psi_c(z)}{dz} \right) + (E_c(z) - E)\psi_c(z) = 0. \quad (3.21)$$

Note, as $\gamma \rightarrow 0$, we obtain $m^*(z, E) \rightarrow m^*(z)$, (3.21) becomes the usual one-dimensional time-independent Schrödinger wave equation, and the effective valence component $\psi_v(z)$ to the wavefunction vanishes. We return to the parabolic case. Nonparabolicity was thus incorporated into SEQUAL by applying an energy dependent effective mass $m^*(z, E)$, and iteratively solving for the eigenenergy and the conduction band wavefunction $\psi_c(z)$.

Optical properties

The presence of nonparabolicity in the bands will also affect the optical properties of the subbands. In our attempt to collapse nonparabolicity into an single band treatment using $\psi_c(z)$, we employed an energy dependent effective mass. We therefore find ourselves unable to make use of the commutation relation (2.11) necessary to use dipole matrix formalism. For an optical transition between an initial and final subband indexed i and f , we therefore define the oscillator strength in terms of the momentum matrix elements as [23]

$$f_{i \rightarrow f} = \frac{2}{m_0} \frac{|\langle \Psi_f | p_z | \Psi_i \rangle|^2}{E_i - E_f}. \quad (3.22)$$

The effective mass has been brought inside the expectation value as an operator to account for its z dependence. As is appropriate for the case of a spatially dependent effective mass, we have used the definition for the bare oscillator strength $f_{i \rightarrow f}$ in (3.22), rather than the scaled oscillator strength $f'_{i \rightarrow f} = (m^*/m_0)f_{i \rightarrow f}$ in (2.17). This expression (3.22) reduces to (2.17), under the dipole approximation in a bulk material with parabolic bands.

As shown in Leavitt[39] and Sirtori *et al.* [23] momentum matrix element can be written in terms of only the conduction band component $\psi_c(z)$:

$$\langle \Psi_f | p_z | \Psi_i \rangle = \frac{1}{2} \left\langle \psi_c^{(f)} \left| p_z \frac{m_0}{m^*(z, E_f)} + \frac{m_0}{m^*(z, E_i)} p_z \right| \psi_c^{(i)} \right\rangle \quad (3.23)$$

$$= \frac{i\hbar}{2} \int \left[\frac{m_0}{m^*(z, E_f)} \frac{d\psi_c^{(f)*}(z)}{dz} \psi_c^{(i)}(z) - \frac{m_0}{m^*(z, E_i)} \psi_c^{(f)*}(z) \frac{d\psi_c^{(i)}(z)}{dz} \right] dz. \quad (3.24)$$

The use of (3.24) and (3.22) to calculate the oscillator strength, rather than the parabolic dipole matrix element method described in (2.17) will generally result in slight reductions in the oscillator strength for a given transition. However, if the nonparabolicity is not too great, the parabolic case can provide a reasonable approximation. Also, the inclusion of nonparabolicity relaxes the selection rule for intersubband transitions. No longer is the momentum (or dipole) matrix element equal to zero for the in-plane directions. However, the selection rule was experimentally found to hold to an accuracy of 0.2% for a 8.1 μm intersubband transition in a GaAs quantum well[42], and this deviation from the selection rule was neglected.

3.3 Multiple quantum well design

3.3.1 Overview

The design of the MQW gain medium for our laser was performed using the tools described above. SEQUAL was used to find the quasi-bound states for the structure under consideration, taking into account effects of space charge and nonparabolicity of the bands. The wavefunctions were used to find the scattering lifetimes of each level. This information allows us to find the populations of each subband and thus obtain a self-consistent picture

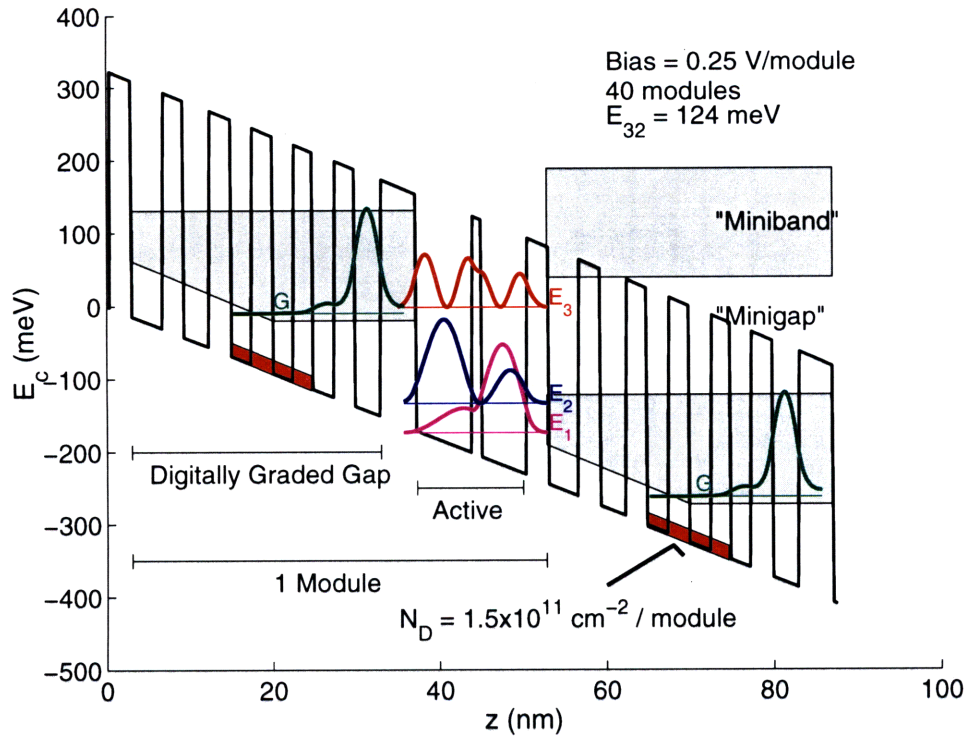


Figure 3-1: Conduction band profile for mid-infrared ($10\ \mu\text{m}$) quantum cascade laser.

of the structure. The wavefunctions were also used to calculate spontaneous and stimulated emission rates, which were in turn used to calculate the bulk gain of the structure.

The design of the conduction band was similar to the QCL designs by the Bell Labs group[43, 44]. A schematic of our conduction band under the design bias is shown in Fig. 3-1. The structure consists of 40 modules cascaded together. Each module consists of a double well active region, and a six well superlattice region known as the digitally graded gap (DGG). The laser is designed to operate as follows. Electrons are injected from the superlattice ground state G into the excited state E_3 . The “lasing” transition takes place between levels 3 and 2 with a calculated energy difference of $E_{32} = 124\ \text{meV}$ ($\lambda = 10.5\ \mu\text{m}$). E_2 is then quickly depopulated by fast LO-phonon scattering to E_1 , which ensures a population inversion between levels 3 and 2. Electrons are extracted from E_1 by rapid tunneling through the digitally graded gap, where they relax into G and are injected into the next module. The DGG superlattice region forms a highly transmissive “miniband” around the E_1 and E_2 energy range. It also acts as a Bragg reflector and forms a “minigap,” preventing electron escape from E_3 into the continuum. Carriers are provided by doping

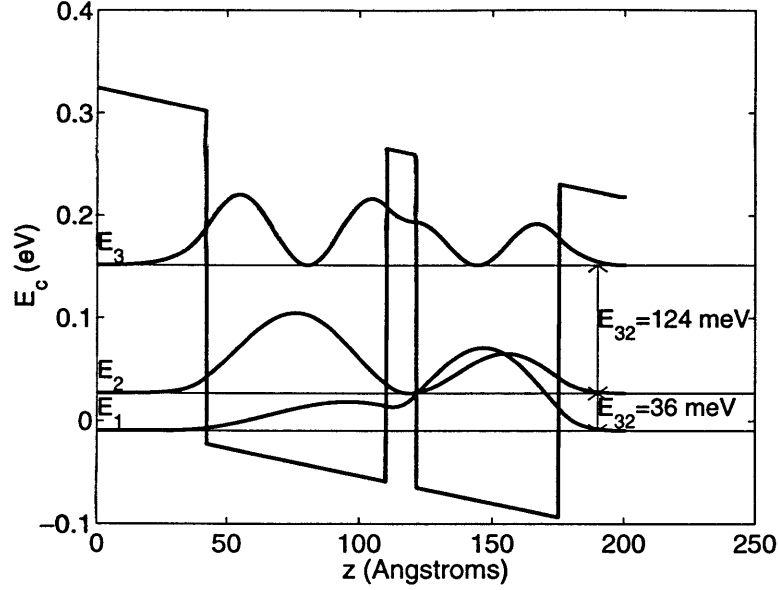


Figure 3-2: *Double well active region. The layer thicknesses are 42.4, 67.8, 11.3, 53.7, and 25.4 Å (15/24/4/19/9 ML). The intrawell lasing transition occurs between E_3 and E_2 . The envelope functions are obtained from (3.21) and the squared magnitude $|\psi_i(z)|^2$ plotted at their respective energies E_i .*

inside the DGG at a level to provide $N_D = 1.5 \times 10^{11} \text{ cm}^{-2}/\text{module}$. Doping the modules to provide carriers helps prevent the buildup of space charge in the MQW structure, thus avoiding large-scale band bending. It is also advantageous to have the dopants set back from the active region to prevent line broadening due to impurity scattering.

3.3.2 Active region

The active region is composed of a 24 ML and a 19 ML well separated by a 4 ML center barrier. Each GaAs monolayer is 2.83 Å. When the active region is under the design bias $V_{mod} = 0.265 \text{ V/module}$, the double well system contains three subbands at energies E_1 , E_2 , and E_3 , shown in Fig. 3-2. The wavefunctions $\psi_i(z)$ are obtained from (3.21) and plotted at their respective energies. The wells are designed to maximize the population inversion $\Delta N_{32} = N_3 - N_2$, and the dipole matrix element z_{32} (2.26). Population inversion is achieved through engineering the wavefunctions to control scattering rates and thus lifetimes of levels.

Simulation of the active region is accomplished by using SEQUAL to find the bound

Table 3.1: *Active region properties*

ij	E_{ij} (meV)	$z_{ij,eff}$ (Å)	$f_{i \rightarrow j}$ (w/ NP)	$\tau_{ij}^{(LO)}$ (ps)
21	36	30.2	8.76	0.19
32	124	25.8	21.8	1.29
31	160	3.68	0.57	1.65

states of a 5 layer structure. Although the states are actually “quasi-bound” states and should be treated using SEQUAL in transmission mode, this method will only yield $|\psi_i(z)|^2$, rather than $\psi_i(z)$ that is necessary to compute matrix elements. The bound states can be found by SEQUAL by setting the contacts to be composed of barrier material rather than well material. The resulting bound states yield a real wavefunction that closely corresponds to that of the traveling wave states. A summary of the properties of the active region is presented in Table 3.1. The values for the oscillator strength $f_{i \rightarrow j}$ are calculated according to (3.22) and (3.24), and thus include the effects of nonparabolicity and the spatially dependent effective mass. The effective dipole matrix elements were calculated starting from $f_{i \rightarrow j}$ using (2.17) and are included for reference.

LO-phonon scattering

For the energies involved in our structure, the dominant scattering mechanism is polar longitudinal optical (LO) phonon scattering. This is the standard assumption in all QCL design[6], and is experimentally justified by observing the temperature dependent electroluminescence from a quantum cascade structure[45]. Such an assumption can also be justified by examining Smet’s extensive calculation of intersubband and intrasubband scattering rates in single and multiple quantum wells[21]. Intrawell scattering times for intersubband LO phonon scattering are typically in the subpicosecond to picosecond range. For longitudinal deformation acoustic and piezoelectric phonon scattering, typical scattering times are hundreds of picoseconds at 4.2 K, and tens of picoseconds at higher temperatures. Intersubband scattering due to interface roughness and impurities (for our conditions) are calculated to remain above tens of picoseconds. Intersubband electron-electron scattering calculations are somewhat involved, but Smet’s calculations indicate that they will likely be greater than several picoseconds. For first order calculations of subband populations, it

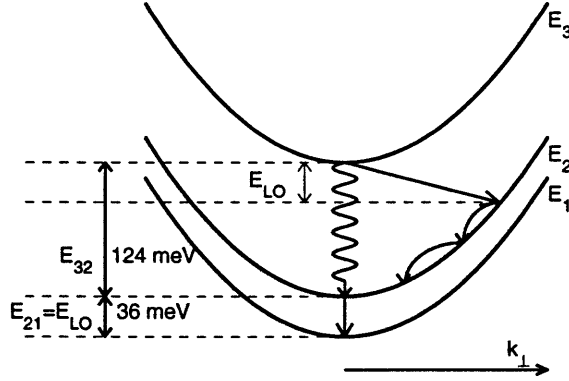


Figure 3-3: *Subband dispersion in active region. The wavy line indicates the calculated 124 meV optical transition between E_3 and E_2 . LO-phonon scattering is the dominant scattering mechanism; both intersubband and intrasubband transitions are indicated by the solid arrows. E_{32} is set at the LO phonon resonance (36 meV) to provide fast depopulation of E_2 .*

is sufficient to consider only LO-phonon scattering and tunneling.

The LO-phonon scattering paths are shown in Fig. 3-3. The scattering rates are calculated assuming zero initial in-plane momentum $k_{\perp,i} = 0$, as fast intrasubband scattering will quickly bring the electron to the subband minimum. This is not strictly true, as the electron bath will probably be somewhat heated and fill the subband according to some Fermi distribution. Even if cold, the subband will be filled up to the Fermi level one or two meV above the edge. Still, these effects will only slightly affect the results, and are thus omitted from the analysis. Also, state blocking effects in the final subband are omitted, as the lower subbands are expected to be empty due to population inversion. The energy separation E_{21} is set at the LO phonon resonance $\hbar\omega_{LO} = 36$ meV, resulting in a fast scattering time $\tau_{21}^{LO} = 0.19$ ps. The LO relaxation time between E_3 and E_2 is much longer due to the large in-plane momentum exchange: $\tau_{32}^{LO} = 1.29$ ps. Since the subband separation E_{32} is large, and the energy of the emitted phonon is fixed, there must be a substantial momentum transfer between the phonon and the electron. This slows the scattering rate, since the Fröhlich interaction strength is inversely proportional to the square of the in-plane momentum exchange (2.63). After the initial intersubband scattering event, the electron will very quickly relax to the bottom of the subband through a series of intrasubband LO-phonon scattering events. Electrons in E_3 will also make spontaneous radiative transitions with a

much slower rate ($\tau_{32}^{(rad)} = 87 \text{ ns}$)⁻¹. The overall lifetime of the upper state is given by

$$\tau_3 = \left(\tau_{32}^{-1} + \tau_{31}^{-1} + \tau_{3,esc}^{-1} \right)^{-1} = 0.71 \text{ ps}, \quad (3.25)$$

where $\tau_{3,esc}$ is the escape time of an electron in E_3 to continuum due to tunneling. It will be shown later that this time is quite long due to Bragg confinement from the DGG. We therefore have a quantum efficiency of $\eta_q = \tau_3 / \tau_{32}^{(rad)} \sim 10^{-5}$ for our LED. The vast majority of electrons that travel through the structure will emit phonons rather than photons.

Rate equations

We can estimate the population of each level by setting up a series of rate equations:

$$\frac{dN_3}{dt} = \frac{J}{e} - N_3 \left(\frac{1}{\tau_{32}} + \frac{1}{\tau_{31}} \right) \quad (3.26)$$

$$\frac{dN_2}{dt} = \frac{N_3}{\tau_{32}} - \frac{N_2}{\tau_{21}} \quad (3.27)$$

$$\frac{dN_1}{dt} = \frac{N_3}{\tau_{31}} + \frac{N_2}{\tau_{21}} - \frac{N_1}{\tau_1}, \quad (3.28)$$

where J is the current density at the operating point, and N_1, N_2, N_3 are the two-dimensional population densities of each level. This expression assumes that all the electrons that contribute to the current flow through E_3 , and out of E_1 . To conserve current flow between modules, we will also assume that $N_1 / \tau_1 = J / e$. Assuming a steady state, we can derive an expression for the population inversion:

$$\Delta N_{32} = N_3 \left(1 - \frac{\tau_{21}}{\tau_{32}} \right) = \frac{J \tau_3}{e} \left(1 - \frac{\tau_{21}}{\tau_{32}} \right). \quad (3.29)$$

After calculating the lifetimes, we can experimentally measure J to obtain an actual estimate of the inversion.

Well design

We designed the lasing transition to be primarily a vertical intrawell transition, as opposed to some designs possessing a diagonal interwell transition[6]. An intrawell transition will

likely have a narrower linewidth, since interface roughness scattering is less important. Also, improved overlap between the wavefunctions leads to a greater oscillator strength, and higher gain. An interwell transition is advantageous in that the upper state E_3 will have a longer lifetime due to slower LO phonon scattering through the barrier, and reduced electron escape to the continuum through tunneling. Interwell transitions also offer the advantage of tunable emission through Stark shift in the energy levels[12].

While the design philosophy was based on that of the Bell Labs QCLs[43, 44], material differences made some of the design more challenging. The GaAs/Al_{0.4}Ga_{0.6}As system has a barrier height of $\Delta E_c = 324$ meV, much smaller than the height $\Delta E_c = 520$ meV for the In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As material system. This sets a limit on our energy separations, although our designed energy is easily achievable without exceeding the barrier edge. More importantly for our structure, the lower barrier height puts an upper limit on the applied bias before the wavefunctions leak out of the wells into the continuum.

The active region well design had to meet the following requirements:

- (1) $E_{32} \approx 125$ meV ($\lambda \approx 10$ μ m).
- (2) $E_{21} \approx 36$ meV (LO-phonon resonance).
- (3) Majority of $\psi_3(z)$ and $\psi_2(z)$ in left well.
- (4) Keep E_3 below barrier edge to prevent leakage.
- (5) $E_{43} > 36$ meV (only E_3 populated).

These design criteria proved somewhat challenging to meet while still using reasonable values for device bias and barrier height. When designing the structure, it is conceptually helpful to consider the system to be made of two single quantum wells coupled by a middle barrier. As the coupling is increased by shrinking the width of the barrier, the wavefunctions will increasingly mix together to form symmetric and antisymmetric eigenfunctions.

Since this is an intrawell transition, satisfaction of condition (1) determines the width of the left well. This energy difference E_{32} is largely insensitive to bias. Therefore, at the designed bias, $\psi_1(z)$ and $\psi_4(z)$ will mostly be localized in the right well. Note that $\psi_4(z)$ is not shown in Fig. 3-2 as E_4 has been pushed above the barrier height, and therefore no longer appears in bound state calculations. Our only concern with this fourth subband

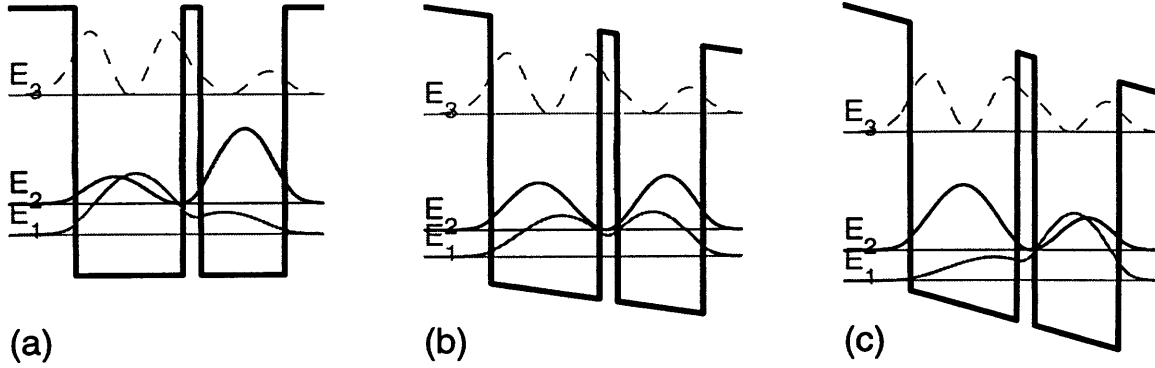


Figure 3-4: Evolution of states E_1 and E_2 under bias. Shown are three bias points: (a) zero bias, (b) anticrossing point $V = 0.138$ V, (c) operating point $V = 0.265$ V. The E_3 level is shown for reference, although it was not calculated using nonparabolicity.

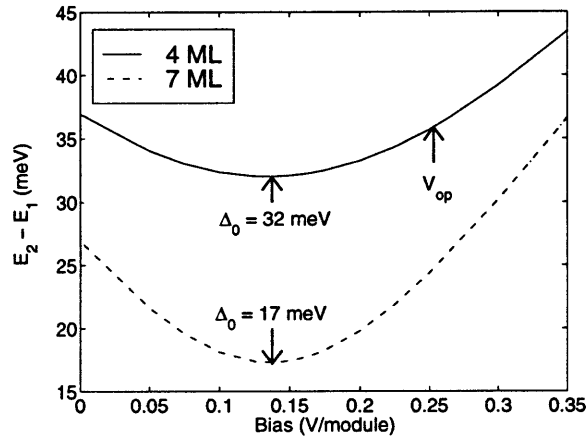


Figure 3-5: Energy difference E_{21} as a function of bias in two structures. The actual design structure has a center barrier 4 ML thick and a large Δ_0 , while a similar structure with a 7 ML center barrier has a small Δ_0 . The operating voltage where E_{21} is indicated by V_{op} .

is that it not participate in the laser operation. The right well must be made narrower than the left so as to push E_4 above E_3 and out of the well. The resulting situation can be seen in Fig. 3-4. At zero bias (a), $\psi_1(z)$ is localized in the left well and $\psi_2(z)$ in the right. As the bias is increased (b), the levels anticross and both wavefunctions (symmetric and antisymmetric) are evenly split between the wells. This is the point when the energy difference E_{21} is at its minimum. As the bias is increased to the operating point (c), $\psi_1(z)$ is now mostly in the right well and $\psi_2(z)$ in the left.

If there was no interaction between the wells, a very high bias would need to be applied to the structure to achieve the desired separation E_{21} . The anticrossing between E_2 and

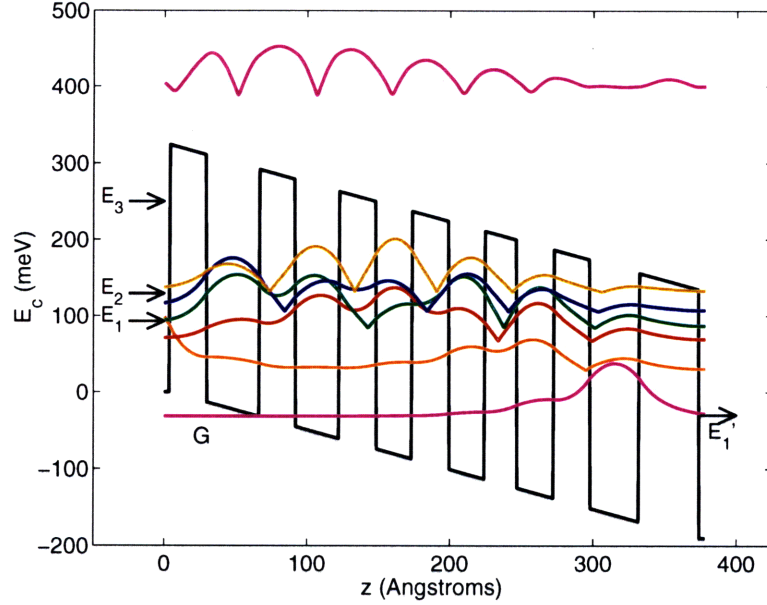


Figure 3-6: *Digitally graded gap.* The layer thicknesses are 25.4, 36.7, 25.4, 31.1, 25.4, 25.4, 25.4, 25.4, 22.6, 25.4, 25.4, 28.3 and 42.4 Å (9/13/9/11/9/9/9/8/9/9/10/15 ML). The six-transmission state squared-magnitude wavefunctions that make up the miniband are shown at their respective energies. Also shown is the first of the continuum states above the minigap. The lowest state G is aligned with E_3 in the next module.

E_1 is exploited to obtain the necessary $E_{21} = 36$ meV at a reasonable bias. As a bias is applied to the double well structure, the energy difference between the two levels reaches a minimum at some bias. This minimum energy difference is called the anticrossing gap Δ_0 , which is a measure of the coupling between the two wells. Tightly coupled wells will have a large anticrossing gap, and vice versa. This can be seen in Fig. 3-5, where E_{21} is plotted for our structure with a 4 ML center barrier, and for another similar structure with a thicker 7 ML barrier. The loosely coupled structure with the thick barrier has $\Delta_0 = 17$ meV and would need to be biased up to $V_{mod} = 0.35$ V/module to achieve the desired separation. The actual design structure can be biased as low as $V_{mod} = 0.25$ V due to the larger splitting from state coupling. The tradeoff is the wavefunction mixing that takes place around the anticrossing point, which results in a delocalized wavefunction, and a transition that is less intrawell in character.

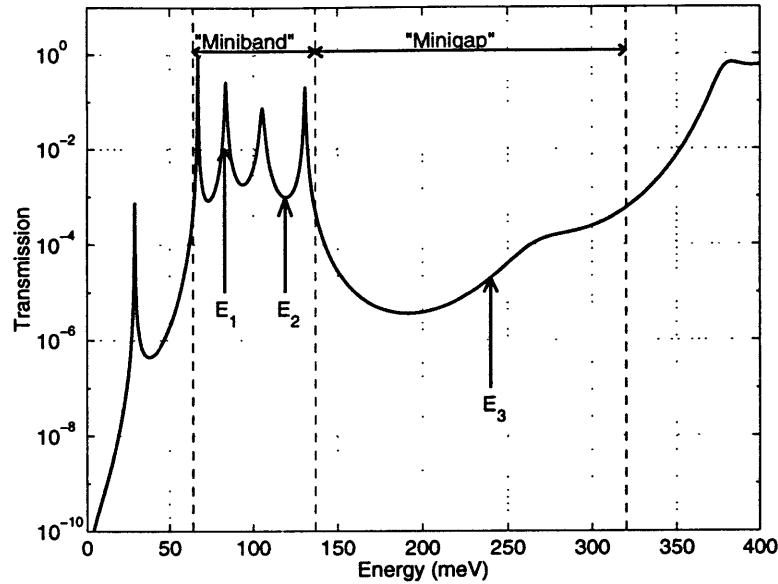


Figure 3-7: Transmission through six-well digitally graded gap under design bias shown in Fig. 3-6. The energy bands of high and low transmission are labeled the “miniband” and “minigap” respectively. The energies of the three states from the previous active region are indicated by arrows.

3.3.3 Digitally graded gap

The purpose of the digitally graded gap (DGG) is to allow electrons to quickly exit from levels E_1 and E_2 , while blocking tunneling from E_3 . Electrons collect in the ground state G where they are injected into E_3 of the next module. The results of the SEQUAL transmission state simulation are shown in Fig. 3-6, with the transmission coefficient of an electron tunneling through the structure shown in Fig. 3-7. Transmission peaks indicate quasi-bound states, and the FWHM of each peak indicates the tunneling lifetime of the state (3.9). The DGG is designed so several states cluster together near E_2 and E_1 of the previous module. In real devices, the quasi-bound states will be additionally broadened by impurity and phonon scattering, creating a highly transmissive “miniband.” Electrons will quickly exit from the previous module, scatter down to state G by fast LO-phonon scattering, and tunnel into E'_3 in the next module.

Electrons are prevented from escaping from E_3 into the continuum by Bragg reflection caused by the DGG. Transmission at E_3 is 10^4 times less likely than from E_1 . The depth of the transmission dip depends on the number of wells; if fewer than six are used,

Bragg confinement suffers. To determine the escape lifetimes $\tau_{3,esc}$, one entire module (active + DGG) was simulated. However, since we were only interested in escape into the continuum, the injection barrier was artificially widened from 15 ML to 45 ML, effectively blocking tunneling back into the previous module. The resonance peak for E_3 had a FWHM of $0.45 \mu\text{eV}$, corresponding to a $\tau_{3,esc} \approx 1.5 \text{ ns}$. This time is much greater than the LO-phonon lifetimes of this level, and thus will not affect τ_3 (3.25).

3.3.4 Injection and exit barriers

The injection and exit barriers are the interfaces between the active region and the digitally graded gap. These barriers must be thick enough to isolate the active region, but still allow fast tunneling in and out. We used two equivalent methods to calculate the tunneling times through

the barriers. The first method, mentioned above, uses SEQUAL to find the transmission as a function of energy through a structure. The FWHM of the resonance can be used to find the lifetime of a quasi-bound state (3.9). The problem with this method is that it gives the tunneling rate out of the state in all directions, and is not well suited to finding a tunneling time through a single barrier.

WKB tunneling time

An equivalent approach is to use a semiclassical method based on the Wentzel-Kramers-Brillouin (WKB) approximation[46]. The WKB approximation is useful for finding the tunneling probability through a barrier, provided that the energy of the particle is small compared with that of the barrier. The theory approximates an potential barrier $V(z)$ as a series of square barriers, under which the transmission probability falls off exponentially, while neglecting reflection effects inside the barrier. The tunneling probability is given by

$$T \sim \exp \left| -2 \int_a^b k(z) dz \right|, \quad (3.30)$$

where

$$k(z) = \sqrt{\frac{2m^*}{\hbar^2} |V(z) - E|}. \quad (3.31)$$

This tunneling probability can be used with a further semiclassical approximation to obtain a tunneling time. For a quasi-bound state, the wavepacket energy is peaked at a certain energy. One can visualize this case as a particle bouncing back and forth between the walls with some frequency given by the kinetic energy E , slowly leaking into the continuum. The tunneling lifetime can be estimated from the tunneling probability at the barrier multiplied by the rate at which the particle strikes the barrier.

$$\frac{1}{\tau} = \frac{\sqrt{\frac{2E}{m^*}}}{2W} T(E), \quad (3.32)$$

where W is the well width. This approach works best for highly excited states in a well, where the semiclassical picture is most valid. However, using this approach to calculate tunneling times for even the ground state can give a reasonable estimate for the tunneling time.

Injection barrier

The 15 ML injection barrier is used to isolate the active region from the previous DGG superlattice region. This barrier should be thick enough so that $\psi_3(z)$ does not spread into the previous well. Having a thick barrier also minimizes coupling between E_3 and G , which keeps the anticrossing gap small. A large anticrossing gap will prevent close alignment of G and E_3 and will increase wavefunction mixing. We would like the anticrossing gap to be smaller than the expected linewidth of the transition, so that transitions from G to E_2 when biased at the anticrossing gap will not broaden the transition.

The anticrossing gap between G and E_3 is 9 meV and occurs at $V_{mod} = 0.234$ V, slightly less than the design voltage $V_{mod} = 0.265$ V. Due to nonparabolicity, at the anticrossing point, the wavefunctions are still highly localized.

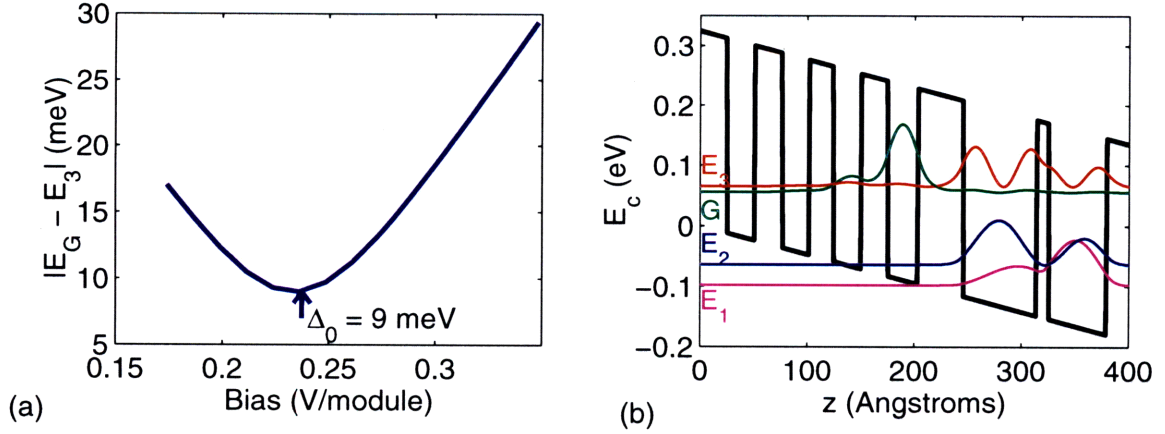


Figure 3-8: *Injection barrier anticrossing. (a) Energy difference $|E_G - E_3|$. (b) Interface module at anticrossing voltage $V_{mod} = 0.234$ V.*

Table 3.2: *Exit barrier WKB tunneling times*

Exit barrier thickness (ML)	WKB tunneling time (ps)
6	0.17
9	0.48
12	1.3

Exit barrier

The function of the 9 ML exit barrier is essentially the same as the injection barrier. It must allow for fast tunneling out of E_1 and E_2 , without a great deal of coupling between the active region states and the DGG states. The WKB method described above was used to find tunneling times for the exit barrier at three thicknesses. These tunneling times provide upper limits on the actual times, since they do not take into account additional transmission provided by the resonant structure of the DGG. The value of 9 ML was fast enough to empty E_1 more quickly than τ_3 .

3.3.5 Self-consistency

The electrons that supply the current in the device are supplied by doping two wells and barriers in the center of the DGG (Fig. 3-1). This 100 Å region was doped at $N_D = 1.5 \times 10^{17} \text{ cm}^{-3}$, which works out to a two-dimensional doping of $N_D^{(2D)} = 1.5 \times 10^{11} \text{ cm}^{-2}/\text{module}$. While SEQUAL has the capability of accounting for the effects of space charge on the po-

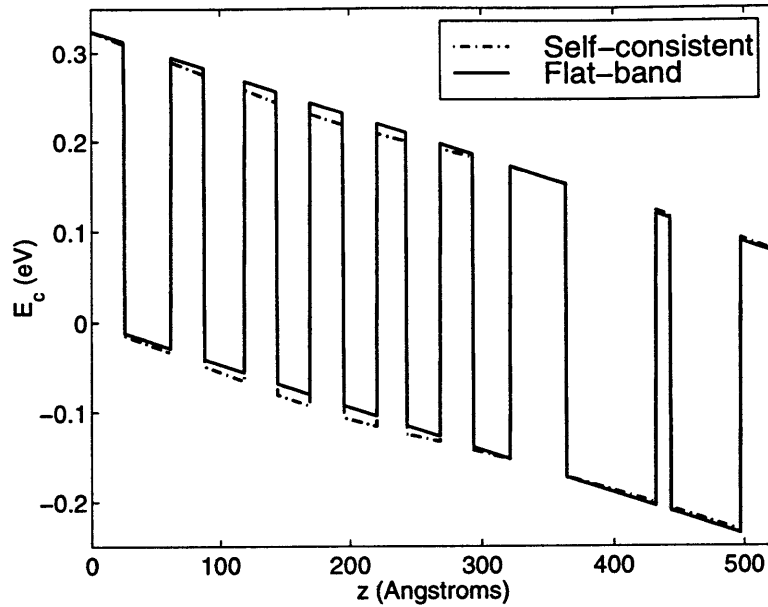


Figure 3-9: *Self-consistent band structure (dashed) calculated with the electron population of $1.5 \times 10^{11} \text{ cm}^{-2}$ split 50/50 between the injection level G and the excited state at E_3 . The flat-band structure is shown for comparison.*

tential, the computational cost is high. The level of doping was chosen to be low enough to cause very little band-bending. Consequently, many of the calculations above were performed with no band-bending, and the full self-consistent solution was only computed at the end for verification. The distribution of charge carriers among the levels is determined by the full set rate equations including all the levels in the DGG. Since many of the scattering rates are unknown, this full solution is not feasible. However, if the active region functions as designed, the majority of the electrons will be in state G and E_3 . The band structure with the electron population split 50/50 between the two levels is shown in Fig. 3-9. The band structure is largely unchanged inside the active region. Changes are small, as the energy levels shifted by only one or two meV.

3.3.6 General Discussion

The effect of band nonparabolicity was treated only as a perturbation of the parabolic design. Nelson *et al.* observed that in a GaAs/AlGaAs quantum well, the effect of nonparabolicity was to raise the energy of the ground state slightly ($\sim 1 \text{ meV}$), and to lower the excited states by variable amounts. This negligible increase in the ground state energy

Table 3.3: *Effect of nonparabolicity*

ij	E_{ij} (meV)(w/o NP)	E_{ij} (meV)(w/ NP)	f_{ij} (w/o NP)	f_{ij} (w/ NP)
21	36.7	36.5	9.00	8.76
32	142.8	124.3	23.2	21.8
31	179.4	160.8	0.534	0.574

was observed regardless of the energy of the ground state. This is related to the lowering of $m^*(z, E)$ inside the barriers. When the ground state is low in energy, the wavefunction is mostly in the well where the nonparabolicity adjustment is small. In a narrow well, where the ground state energy is high, the electron spends much of its time in the barriers, where the nonparabolicity is small. This same effect was observed in our MQW simulation. It was generally observed that any “ground-like” state, made up of mixed ground states from individual wells, would remain largely unaffected by nonparabolicity. In our device, only E_3 is an excited state, and is lowered by about 20 meV. For most of the design process, nonparabolicity was only applied to calculate the E_3 energy, with full nonparabolicity calculations being applied at the end of the design process for fine adjustment of the structure. The effects of considering nonparabolicity for the active region are shown in Table 3.3. The oscillator strength for the $3 \rightarrow 2$ lasing transition is also reduced slightly.

The MQW structure discussed above will compose the gain medium for the QCL. Each module is 497.2 Å (176 ML) long. The design bias, $V_{mod} = 0.265$ V/module ($E = 54.3$ kV/cm), is the point where we obtain $E_{21} > 36$ meV, and E_3 is in rough alignment with G . This is the point when LO-phonon scattering from E_2 to E_1 is most efficient. However, simulation of the injection barrier with nonparabolicity shows that most efficient injection will occur at the anticrossing point between E_3 and G , $V_{mod} = 0.234$ V/module ($E = 47.1$ kV/cm). This is the point that will correspond to the peak electroluminescence power, and the conductance plateau during transport measurements. It is desirable for these two bias points to coincide, and the present separation is due to a simulation error during the design.

The design also incorporates a feature found useful by Faist *et al.* to improve laser operation. If the separation between the ground state G and the preceding E_2 level is too small, thermal backfilling of E_2 will occur and the population inversion will drop. This

phenomenon becomes more severe at higher temperatures. In one sample[47], Faist was able to reduce the threshold current at 50 K from 500 mA to 400 mA by increasing the separation from ~ 70 meV to ~ 100 meV. In our structure, the separation between G and E_2 is approximately 130 meV at operating bias, which should provide sufficient protection from thermal backfilling.

The peak bulk gain due to this MQW structure can be calculated using the expression (2.27):

$$\gamma(\nu_0) = \frac{\Delta N e^2 f_{i \rightarrow f}}{2\pi c n \epsilon_0 m_0 \Delta \nu}. \quad (3.33)$$

If the oscillator strength $f_{i \rightarrow f}$ is calculated according to (3.22) and (3.24), this expression for the peak gain takes into account nonparabolicity and the spatially dependent effective mass in the structure. To find the linewidth and population inversion, we borrow some experimental results from the next section. A linewidth of 15 meV was used for this calculation, which is a somewhat conservative estimate based on the observed 11 meV linewidth described in Section 4.3.1. The peak population inversion can be estimated (3.29), taking $J_{op} = 4$ kA/cm², and the state lifetimes as calculated in Section 3.3.2. We obtain $\Delta N_{32}^{(2D)} \approx 1.5 \times 10^{10}$ cm⁻² over each 497 Å long module. This is approximately 1/10 of the module doping. As calculated above, the oscillator strength is $f_{32} = 21.8$. The refractive index at the operating wavelength $\lambda = 10$ μm is taken to be $n = 3.27$. Using these values, we calculate that the peak bulk gain is $\gamma_{bulk} = 96$ cm⁻¹ for the A40 design.

3.4 Waveguide design

The second major step for building a laser is to design the resonant electromagnetic cavity to provide optical feedback to the gain medium. As is done with most semiconductor lasers, the end mirrors are formed by cleaving the wafer along a crystalline axis, and using the air/semiconductor interface as a mirror. However, lateral and vertical confinement must also be provided, which is traditionally done by forming a dielectric waveguide, where the active region is surrounded by layers of lower dielectric constant. This is the method used by Bell Labs for their InGaAs/InAlAs QCLs. This material system is grown on a

InP substrate, which has a smaller refractive index than either InGaAs or InAlAs. For the upper side confinement, either InAlAs[47] or InP[11] (both of which have a smaller ϵ than the core) are grown above the active region, providing confinement and conduction through light doping. Further steps are taken to prevent coupling of the field with the lossy interface mode at the semiconductor-metal contact interface.

The design of a waveguide structure for a GaAs/AlGaAs based QCL is more difficult. Because $\epsilon_{GaAs} > \epsilon_{AlGaAs}$, the intrinsic dielectric constant of the GaAs substrate is actually greater than the active region. However, the n^+ substrate, doped at $N_D \cong 10^{18} \text{ cm}^{-3}$, provides confinement due to the free carrier contribution to ϵ given by (2.59), but accompanied by a great deal of loss. It is difficult to use the barrier material $\text{Al}_x\text{Ga}_{1-x}\text{As}$, because of the difficulty in achieving conduction through this layer. In $\text{Al}_x\text{Ga}_{1-x}\text{As}$, for $0.23 \leq x \leq 0.9$, DX centers appear below the conduction band in n -type AlGaAs[48]. These centers act as deep donor levels which prevent carrier ionization, restricting current flow. Strasser *et al.* have attempted this method of waveguiding and have experienced the expected problems with carrier freeze-out in the cladding regions at low temperatures[18]. Li *et al.* is considering pure AlAs cladding regions, which avoids the deep DX level problem. This method has potential, however, such a method requires complicated material grading, and a large Al MBE source that are beyond our current capability.

This section explores two possible waveguide structures for our mid-infrared quantum cascade laser. The first structure makes use of the metal contact layer above the MQW structure for confinement above, and the use a heavily doped layer in addition to the doped substrate for confinement. The second structure is borrowed from the far-infrared laser project being undertaken by Bin Xu. This structure involves a layer of metal both immediately above and below the MQW layer. This metal-metal waveguide structure is more difficult to fabricate, but offers the promise of better confinement.

3.4.1 A40 design

The A40 structure displayed in Fig. 4-1 was only designed for observing spontaneous emission, and was not optimized for laser waveguiding. Nevertheless, analysis of the structure

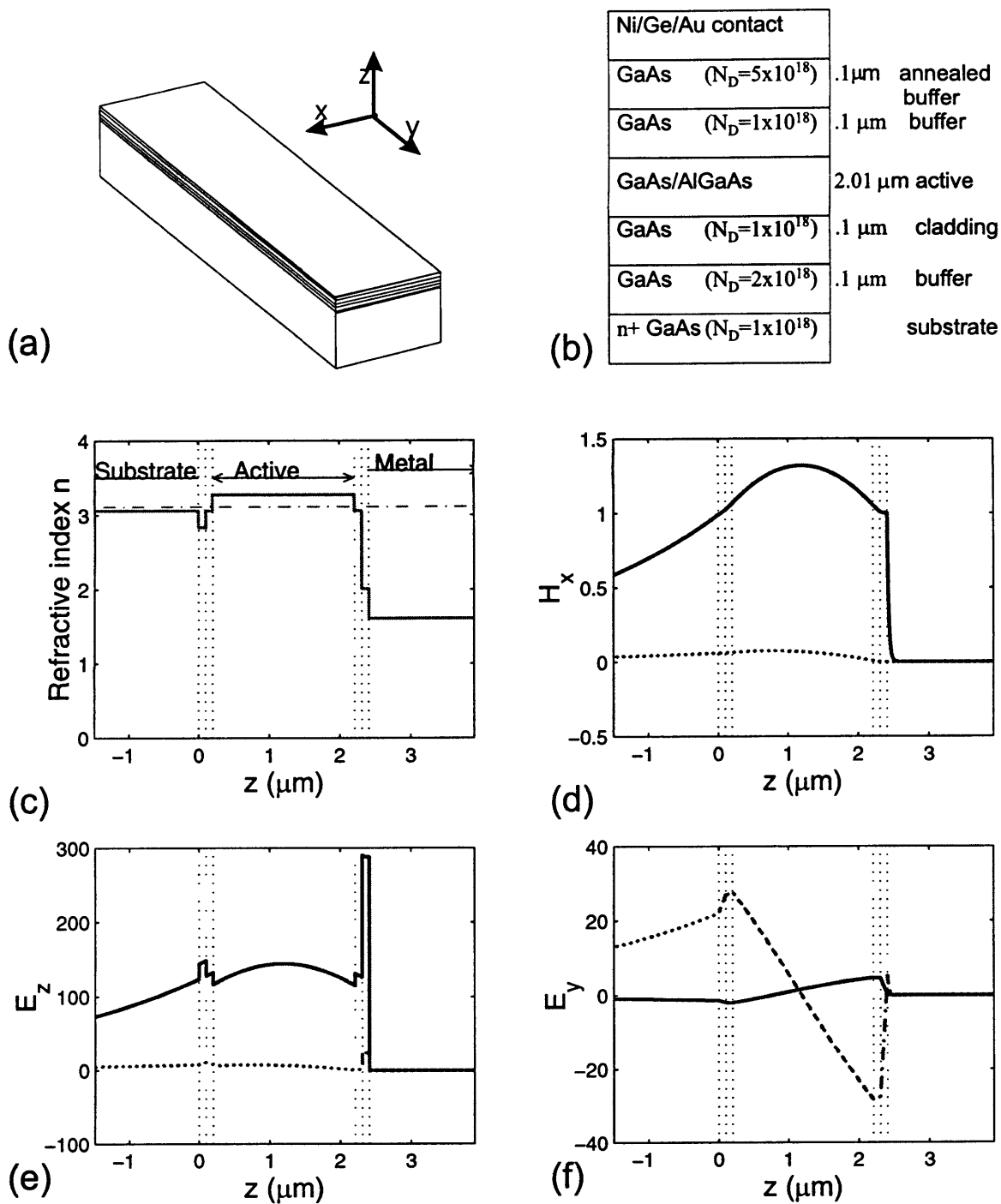


Figure 3-10: Calculated bound TM_0 mode for A40 design with a loss of $\alpha_c = 149 \text{ cm}^{-1}$ and a confinement of $\Gamma = 0.58$. The axis orientation is shown in (a) and the simulated structure is shown in (b). The high doping density in the buffer layer adjacent to the metal contact is meant to simulate the effect of annealing. The real part of the refractive index profile is displayed in (c). The horizontal line gives the real part of the mode effective index $n_{eff} = \beta/k_0$. The field components are shown in (d), (e) and (f), with the real and imaginary parts given by the solid and dotted lines respectively.

reveals that some confinement is actually obtained. The simulated structure and TM_0 mode profile is shown in Fig. 3-10. For this calculation, the permittivity of the active region was taken as an average of the well and barrier materials at the operating wavelength, yielding $\epsilon_{active} = 10.279\epsilon_0$. The calculated cavity loss is $\alpha_c = 149 \text{ cm}^{-1}$ and quantum well confinement factor is $\Gamma = 0.58$. Note that this and all mentioned confinement factors include the selective coupling of quantum wells with z polarized light (2.55). This calculation has a large uncertainty due to the alloyed Ni/Ge/Au ohmic contact. During the rapid thermal anneal, Ge dopants diffuse into the bulk semiconductor to form a heavily doped region aiding electron tunneling through the surface barrier. However, this diffusion process results in a relatively uncontrolled spread of dopants as far as 1000 \AA into the material. For this reason, the A40 device was simulated with the upper buffer layer doped at $N_D = 5 \times 10^{18} \text{ cm}^{-3}$, the maximum level of n -type doping found in bulk GaAs.

This heavily doped region is the source of the majority of the loss. In this structure, 140 cm^{-1} of the loss is due to the upper layers, while only 10 cm^{-1} is due to the lower layer. For the doping level $N_D = 5 \times 10^{18} \text{ cm}^{-3}$, the plasma frequency is about $f_p = 24 \text{ THz}$, quite close to our laser frequency of 30 THz (Fig. 2-2). Near the plasma frequency, the real part of the permittivity ϵ approaches zero. As a result, the electric field E_z , which is inversely proportional to ϵ becomes very strong in this annealed region, causing a great deal of free carrier loss due to impurity scattering.

3.4.2 Plasma waveguide

The A40 structure, while quite lossy, does in fact provide confinement due to the metal contact and the free carrier reduction of ϵ in the substrate. This structure could provide a reasonable waveguide if the loss due to the annealed dopants was reduced.

Non-alloyed ohmic contact

One solution is to use a non-alloyed ohmic contact. These types of contacts have traditionally been very difficult to fabricate to n -type (Si doped) GaAs due to the fact that $(N_D - N_A)$ is limited by amphoteric doping to about $\sim 5 \times 10^{18} \text{ cm}^{-3}$. High doping concentrations

$\sim 10^{20} \text{ cm}^{-3}$ can be obtained within a thin ($\sim 30 \text{ \AA}$) surface layer due to Fermi level pinning by midgap surface states. However, this highly doped region is quickly oxidized and usually removed during the oxide etch that precedes metal deposition steps.

Patkar *et al.* solved this problem by growing a thin ($\sim 30 \text{ \AA}$) passivating low temperature grown (LTG) undoped GaAs layer above the heavily doped n^+ GaAs[33, 49]. This LTG layer has a high number of midgap states due to the incorporation of excess As in the layer. These states will pin the Fermi level in the n^+ GaAs at midgap just as the surface states did. Since LTG GaAs oxidizes much more slowly than n^+ GaAs, the passivating layer will protect the doped GaAs during the oxide etch[50]. A nonreactive metal layer such as Ti/Au or Ag can be deposited to form an ohmic contact. Transport occurs as the electrons tunnel through the midgap states in the LTG GaAs layer and through the narrow space charge region. Ohmic contacts with ρ_c of mid $10^{-7} \Omega \text{ cm}^2$ have been produced[33], at least as good as the contact resistances for the traditional alloyed Ni/Ge/Au contact. The advantage for waveguiding is that the doping can be quite well controlled, and need be high only in a thin layer adjacent to the surface.

Proposed structures

This structure, shown in Fig. 3-11(a), can be used as a basis to evaluate the properties of a plasma-metal waveguide. Confinement on the lower side is provided by a $2 \mu\text{m}$ GaAs doped cladding layer, along with the n^+ substrate. We simulated the substrate to have $N_D = 10^{18} \text{ cm}^{-3}$, which lies on the low side of the quoted specification of $1.1 - 2.6 \times 10^{18} \text{ cm}^{-3}$. Confinement on the upper side is provided by the gold contact, described by the Drude model with a momentum relaxation time of $\tau = 0.1 \text{ ps}$, roughly corresponding to the DC conductivity at 100 K. The nonalloyed ohmic contact is simulated with a 70 \AA region doped at the maximum bulk level followed by a 30 \AA surface layer doped at $5 \times 10^{19} \text{ cm}^{-3}$. In reality, these contact layers are too thin to be treated using the bulk Drude model without consideration of quantum effects. It is merely done to give an approximation of the loss. There is a 900 \AA buffer layer between the active region and the non-alloyed ohmic contact, which must be doped to provide electrical contact.

The effect of the width of the active region on the cavity loss α_c and confinement fac-

Ti/Au contact	
LTG-GaAs	3.5 nm cap
GaAs ($N_D=5 \times 10^{19}$)	10 nm contact
GaAs ($N_{D, upper}$)	0.1 μm buffer
GaAs/AlGaAs	d_{ac} (2.5 μm) active
GaAs ($N_{D, lower}$)	2 μm cladding
n+ GaAs ($N_D=1 \times 10^{18}$)	substrate

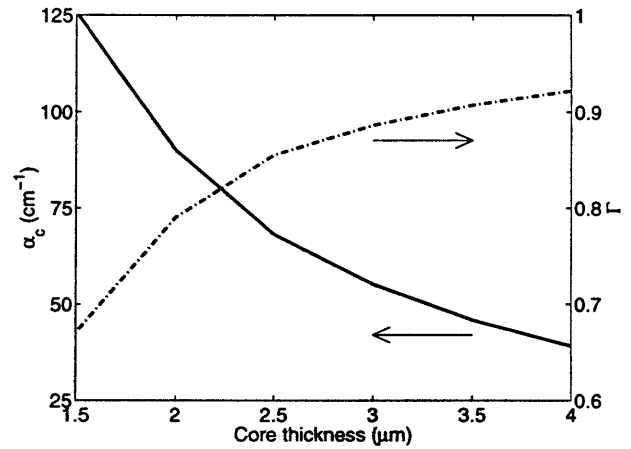


Figure 3-11: (a) Proposed plasma-metal waveguide structure with non-alloyed ohmic contact. (b) Variation of cavity loss α_c and confinement factor Γ with thickness of the core active region. For this plot, the doping of the upper buffer layer and lower cladding layer are both taken as $N_D = 2 \times 10^{18} \text{ cm}^{-3}$.

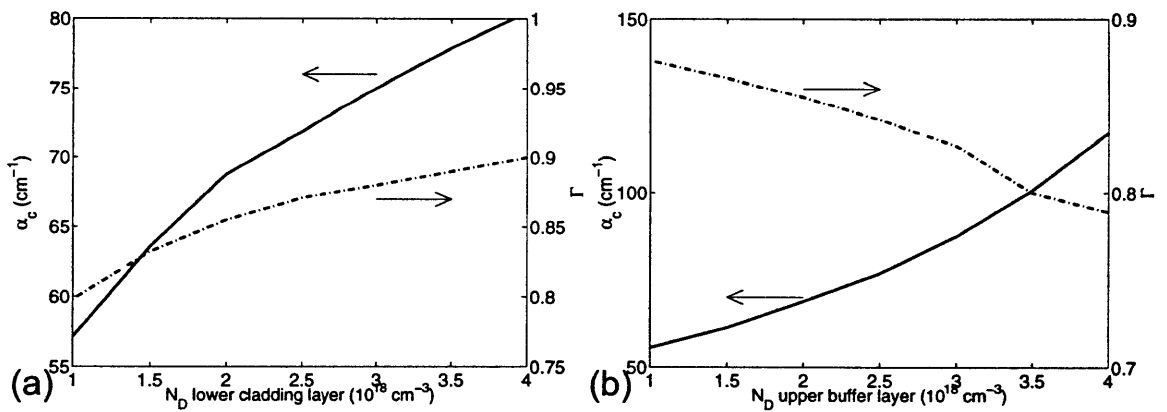


Figure 3-12: Variation of the cavity loss α_c and confinement factor Γ of proposed plasma-metal waveguide for different doping levels of the (a) lower cladding layer and (b) upper buffer layer. For each plot, the doping of the other layer was taken to be $N_D = 2 \times 10^{18} \text{ cm}^{-3}$ and the thickness of the core region is 2.5 μm .

tor Γ is shown in Fig. 3-11(b). In this plot, the lower cladding and upper buffer regions are simulated at $N_D = 5 \times 10^{18} \text{ cm}^{-3}$. Since the radiation wavelength inside GaAs is roughly $3 \mu\text{m}$, the loss depends strongly on the active region width for those plotted. As the waveguide becomes thinner than the wavelength, more of the field is pushed into the lossy cladding layers. It is thus beneficial to design the active region as thick as MBE growth time constraints and limits on device bias allow.

The dependence of the loss and confinement on buffer doping are shown in Fig. 3-12. As expected, increasing the doping increases the free carrier loss. However, it also reduces the refractive index of the cladding layer, improving confinement. Increasing the doping in the upper buffer region also increases loss, albeit much more sharply than for the thicker lower cladding layer. This is due to the reason given above. Due to the high reflectivity of the metal, much of the energy is reflected back into the cavity and the field strength is very high near the metal surface. Also, since E_z is inversely proportional to ϵ , the large reduction of the permittivity in the buffer layer and highly doped contact layer will increase the E -field strength in a very lossy region. Since the upper buffer region (not the contact layer) is not intended to provide confinement, but only transport, it only need be doped high enough to provide current and make contact with the active region.

Proposed plasma-metal structure: PL50

Based on these examinations of waveguide characteristics, we propose a plasma-metal waveguide structure for a GaAs based quantum cascade laser. The simulated profile and field pattern are shown in Fig. 3-13. We choose the thickness of the active region to be $2.52 \mu\text{m}$, corresponding to 50 modules. The lower $2 \mu\text{m}$ cladding layer as well as the upper 900 \AA buffer layer is doped at $2 \times 10^{18} \text{ cm}^{-3}$, chosen so that the Fermi level will be aligned with the digitally graded gap states of the first and last modules. The calculated cavity loss is $\alpha_c = 64 \text{ cm}^{-1}$ with a confinement factor of $\Gamma = 0.85$. The loss can be reduced by reducing the doping in the upper buffer layer, at the risk of increasing contact resistance into the laser structure. Since bulk gain can be designed to be in the hundreds of inverse centimeters, this waveguide structure shows promise for development of a GaAs based QCL.

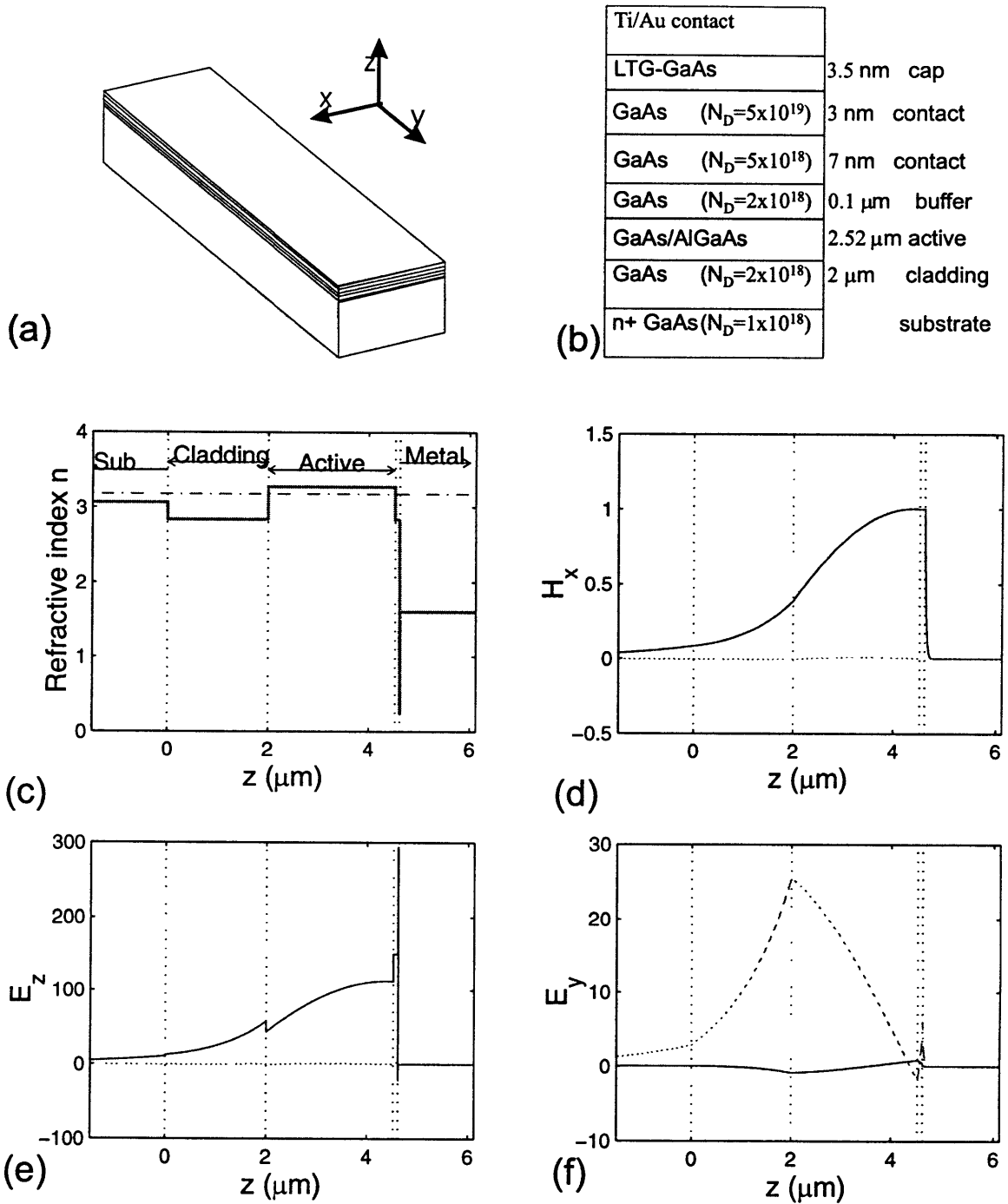


Figure 3-13: Bound TM_0 mode for proposed PL50 design, with cavity loss $\alpha_c = 64 \text{ cm}^{-1}$ and confinement $\Gamma = 0.85$. The axis orientation is shown in (a) and the simulated structure with non-alloyed ohmic contact is shown in (b). The real part of the refractive index profile is displayed in (c). The horizontal line gives the real part of the mode effective index $n_{eff} = \beta/k_0$. The field components are shown in (d), (e) and (f), with the real and imaginary parts given by the solid and dotted lines respectively.

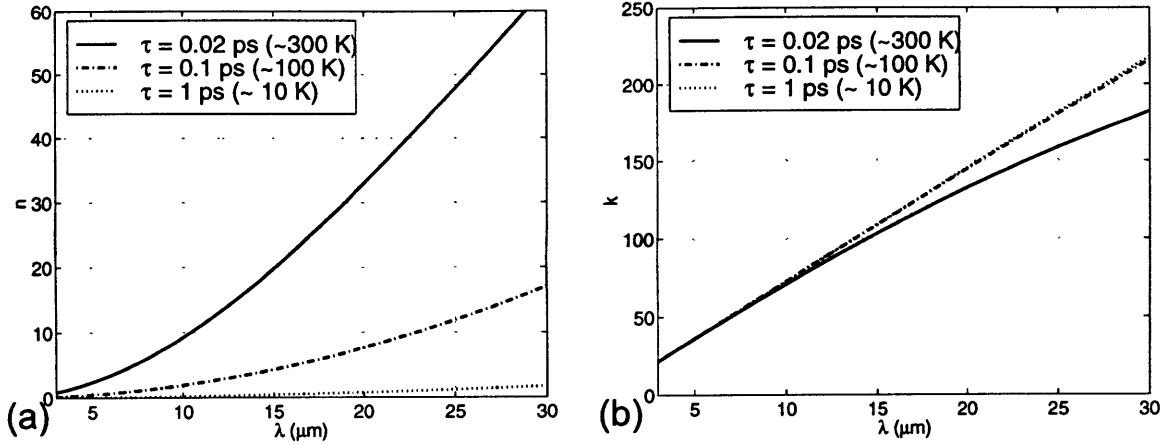


Figure 3-14: Drude model complex refractive index $N = n + ik$ for gold. The temperature dependence is illustrated with typical momentum relaxation times τ , taken from DC conductivity data[51].

For this structure, a great deal of the loss is caused by the close proximity of the metal contact to the active region. The resulting high field that appears in the metal skin and the highly doped contact region is the reason for the extreme sensitivity of the loss to dopants. The loss is also very dependent of the conductivity of the Au contact. In these simulations, a momentum relaxation time of $\tau = 0.1$ ps was used, roughly corresponding to the DC conductivity of gold at 100 K. The added phonon scattering due to higher temperature operation, or the added impurity scattering due to deposition impurities will reduce the conductivity. The Drude model behavior of the complex refractive index for various τ is given in Fig. 3-14. It is observed that for our frequencies of interest, the extinction coefficient k is largely independent of the scattering time. However, as scattering increases, the refractive index n rises rapidly. This reflects the fact that the skin depth for a metal is proportional to $\sigma^{-1/2}$, i.e. the increased scattering impedes the screening of the field. Thus for the same extinction coefficient, the optical loss will be greater. A calculation reveals that for $\tau = 0.02$ ps, corresponding to roughly 300 K, the loss has risen to $\alpha_c = 88 \text{ cm}^{-1}$. To improve on the Drude model approximation, the calculation was repeated for experimentally gathered Au data at 300 K[52]. While the Drude model predicted $n = 9.24$ and $k = 70.8$, experimentally obtained reflectivity data gave $n = 12.24$ and $k = 54.7$ for $\lambda \sim 10 \mu\text{m}$. The difference was significant; the calculated loss for the experimental data was $\alpha_c = 128 \text{ cm}^{-1}$. Unfortunately, there was no available low temperature data for the

optical constants of Au. However, it can be inferred that the loss might be slightly larger than calculated than with the Drude model. This design is promising at low temperatures, but its sensitive temperature performance makes it undesirable for a room temperature laser.

3.4.3 Metal-metal waveguide

A metal-metal waveguide is currently being considered for confinement for a far-infrared cascade laser. Such a waveguide, with metal directly above and below the active region, would have a confinement factor of almost unity. However, the above analysis of the plasma-metal waveguide indicates that it is unlikely that there would be any net advantage to such a scheme. The majority of loss from the PL50 waveguide is from the heavily doped layer adjacent to the metal, and the metal itself. With metal on both sides, the field would be forced even more strongly into the metal on both sides. Simulations indicate that the resulting single sided loss is even greater than the single sided loss from the metal side in the PL50 structure. Additionally, the temperature dependence of the metal loss would be more severe, since both sides are metal. Perhaps the biggest problem with a metal-metal waveguide is that additional problem remains that there is currently no way to fabricate non-alloyed ohmic contacts on both sides of the interface. Hence, a very lossy alloyed contact would have to be made on one side of the device.

These issues are not as critical in a far-infrared laser. In the THz range, the radiation frequency is far below the plasma frequency of the doped buffer layers, and the permittivity becomes negative and large. Hence the electric field strength, which is inversely proportional to the permittivity, will be much smaller than for the mid-infrared case.

Based on these simulations, even the plasma-metal waveguide is not ideal. For long term laser development, this author believes that a way must be found to use dielectric confinement with the minimum of doping needed for transport used. The most likely way to accomplish this is by using AlAs cladding layers, graded to match the GaAs wells at the interfaces. The AlAs cladding layers would then only need to be doped to provide conduction, rather than confinement as well. The lower doping levels would reduce the loss and separate the confined mode from the metal contact. This will likely require a

complicated MBE growth in a machine with a large Al source. Such growth is currently beyond our capability.

Chapter 4

Experiments

4.1 Setup

The first step towards building any laser is observation of spontaneous emission from the gain medium. For this purpose, a MQW structure was grown by molecular beam epitaxy (MBE) and fabricated into suitable test structures. The growth profile for this structure, described in Section 3.3, hereafter referred to as A40, is displayed in Fig. 4-1. This structure was designed for the purpose of observing electroluminescence and testing the electrical transport properties, and thus was grown without deliberate waveguiding, but merely doped contacts. These measurements were done at cryogenic temperatures (4.2 – 77 K), to improve device performance and ease interpretation of the data. DC current measurements were performed using a dipstick immersed in liquid Helium.

The optical measurements were performed with the device mounted on a copper plate cooled by liquid Helium. The dewar was fitted with a output window made from 2 mm thick ZnSe coated for high transmission between 3 – 13 μm . A custom device mount made from oxygen free copper was designed to hold the device to the cold plate. This mount is pictured in Fig. 4-2. The device was affixed to a copper chip carrier with conductive epoxy on the substrate side, providing good thermal contact to the copper heat sink, and a low resistance return path for current. Electrical contact was made to the top side of the emission structures with aluminum wire bonds. The chip carrier was screwed onto the larger mount, allowing devices to be easily changed. A temperature sensor was mounted

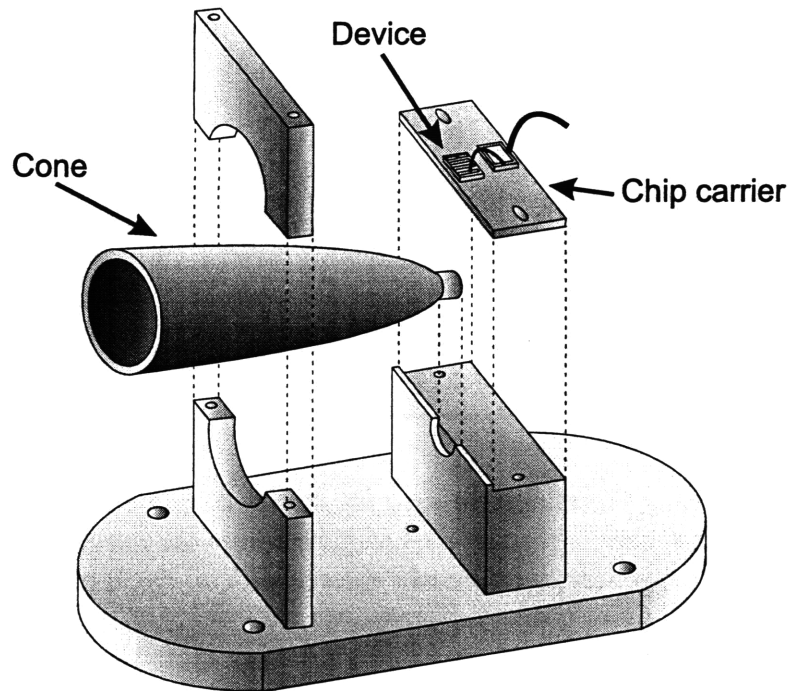


Figure 4-2: *Copper mount used to hold electroformed cone and chip carrier with device. The mount was screwed to the cold plate, providing good heat sinking for the device. All metal-metal surfaces were plated with indium to ensure maximum thermal conductivity. A temperature sensor (not shown) was attached to the mount just below the chip carrier.*

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. Leads were then soldered to the chip carrier to provide the connection with the external power supply. Also, high resistance wires were affixed before and after the chip carrier to allow for 4-terminal bias measurements. This allowed us to find the voltage applied across the active region, assuming there was negligible voltage dropped across the chip carrier, wire bonds, and ohmic contacts to the device.

Since the free space wavelength of the emitted radiation ($\sim 10 \mu\text{m}$) was larger than the thickness of the waveguide ($\sim 3 \mu\text{m}$), the radiation rapidly diverged upon exit from the device. 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 instead. A Winston cone is an off axis parabolic condenser usually used for collection of far-infrared and sub-millimeter radiation for detection[53, 54]. A schematic diagram is shown in Fig. 4-3. For a device placed at the smaller aperture of the cone, radiation emitted at any angle will emerge

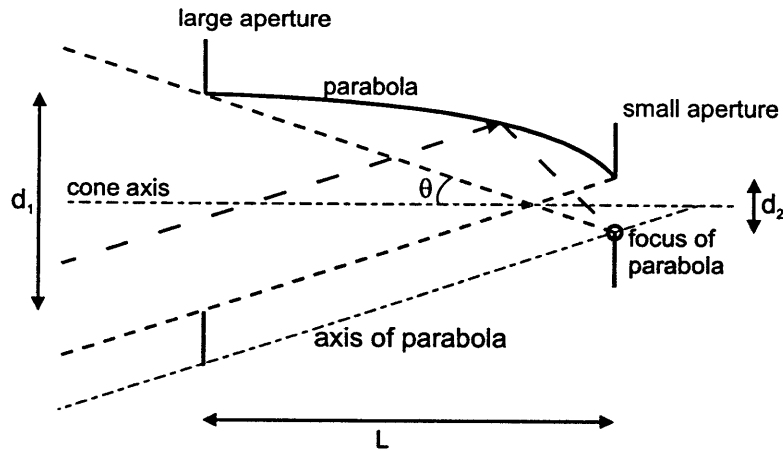


Figure 4-3: Diagram of Winston cone used for output coupling. The cone is formed by rotating a parabolic section about an axis (different from the parabolic axis). Radiation originating within the smaller aperture will emerge with an angle of divergence less than $\theta_{max} = \theta$.

from the larger aperture at an angle θ smaller than some maximum angle

$$\tan \theta_{max} = \frac{d_1 + d_2}{2L}, \quad (4.1)$$

where d_1 and d_2 are the diameters of the larger and smaller apertures, and L is the length of the cone. Radiation emerging from 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.

A cone was chosen rather than a lens to ease optical alignment. Due to the high degree of divergence, a very fast lens placed close to the device would be needed. Such a lens would be very sensitive to alignment, which is impossible to adjust when mounted inside a closed dewar.

Due to the expected high current densities of the devices ($\sim 5 - 10 \text{ kA/cm}^2$) it was necessary to supply the device with fast, high current voltage pulses to avoid heating of the device. Such a bias supply was constructed with the help of B. Riely. The pulses are generated by using a power MOSFET to switch an adjustable voltage. For short, high current pulses, a large $300 \mu\text{F}$ capacitor provides the current. For longer, lower current

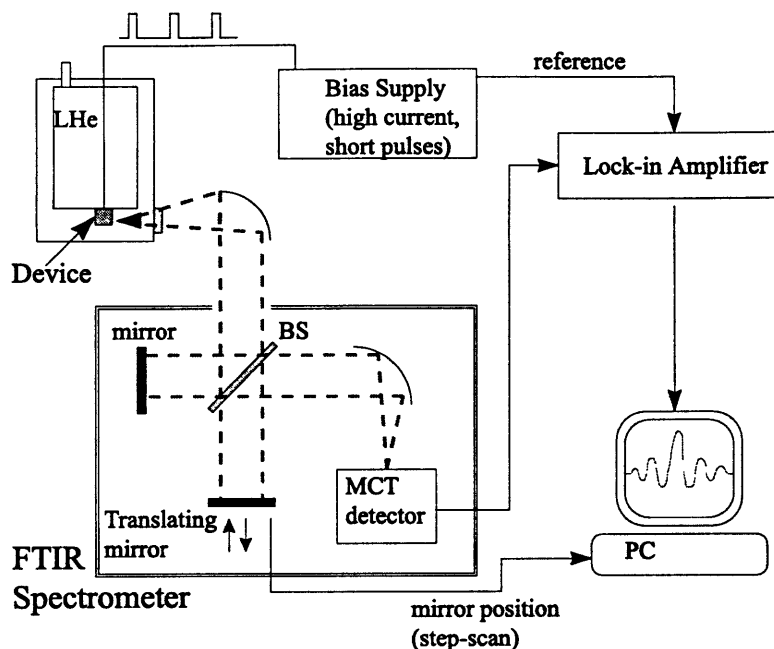


Figure 4-4: *Experimental setup for emission measurements.*

pulses, the voltage regulator can source the necessary current. While the biasing box itself provided sharp pulses with widths less than 500 ns, the poor response of the connector cable and dewar wiring limited the minimum pulse width to about 10 μ s.

While using a MOSFET switch can provide fast, high power pulses, this approach somewhat limits the flexibility of the biasing circuit. A pulse generator based around a power operational amplifier can provide pulses of either polarity, and can more easily be adapted to act as a current source rather than a voltage source. Our decision to use the MOSFET approach was based on the difficulty of finding a suitable high speed, high power op-amp at reasonable cost.

To prevent excessive heating of the device, measurements of emitted optical power were performed in pulsed mode. Typically, pulse widths of 15 – 50 μ s were used at duty cycles of 1 – 2%. Mid-infrared radiation was detected with a mercury-cadmium-telluride (MCT) photodetector cooled to 77 K (peak responsivity : 2462 V/W). The experimental setup is shown in Fig. 4-4. The emission spectrum was resolved using a Nicolet Fourier transform infrared spectrometer (FTIR) operated in step-scan mode with a lock-in amplifier. The FTIR operates by using a Michelson interferometer to trace out an interferogram. The Fourier transform of the interferogram yields the power spectrum of the emitted radi-

ation. When the FTIR is operated in step-scan mode, the translating mirror pauses at each step to allow averaging or other processing of the detector signal. Among other things, this technique allows for small signal detection, such as lock-in detection or boxcar averaging, necessary for observing the small spontaneous emission power above background radiation. For this experiment, the detector signal was fed into a lock-in amplifier referenced with the electrical chopping frequency. The step time of the interferometer was set to be equal to the lock-in integration time, and the lock-in output signal was collected as an interferogram and processed in software. A boxcar averager was used to simultaneously measure bias and current through the device using the 4-terminal scheme in concert with a current sensing resistor. All displayed spectra were collected with 4 meV resolution and Fourier transformed with one level of zero padding using a Happ-Genzel windowing function.

4.2 Fabrication

The molecular beam epitaxy of the multiple quantum well structure was performed at Purdue university by Dr. Michael Melloch. The final fabrication of the wafer into testable structures was performed here at MIT in the Microsystems Technology Laboratory (MTL). The contact mask used was originally designed by Jurgen Smet for testing of far-infrared emitters. The mask contains 2.5 mm long bar structures with widths of 34, 40, and 83 μm suitable for forming laser cavities, as well as $100 \times 100 \mu\text{m}^2$ square diode mesas for measuring I-V characteristics. The processing sequence performed on the A40 wafer was designed by Bin Xu, and is shown in Fig. 4-5. Before processing, the wafer consists of $\sim 2.5 \mu\text{m}$ of MBE quantum wells grown on a n^+ substrate (not shown to scale). The first photostep (b) patterns photoresist to define the areas to be protected from the mesa etch. A wet etch in $\text{NH}_4\text{O}_4/\text{H}_2\text{O}_2/\text{H}_2\text{O}$ solution in the ratio 5:3:240 is used to etch a 2 μm mesa. For a (100) oriented wafer, this etch solution yields an undercut in the $[\bar{1}\bar{1}0]$ direction, and an sloped profile in the $[110]$ direction. The mask is oriented in such a way that the bar structures have shallow sloped sidewalls so that later, metal will be able to climb the sidewalls for the bonding paths. After the mesa etch, a 1600 Å insulating layer of SiO_2 is deposited. It is

important not to make the SiO₂ layer much thicker than 2000 Å, as stress will build up and damage the surface. In this fabrication run, the oxide was deposited using an electron beam evaporator. In fact, this method yields somewhat low quality oxide, with a large pinhole defect density and a high etch rate. It is preferable to use sputtering or low-temperature PECVD. In the second photostep (e),(f) a window is opened in the oxide with which to make a metal contact. The etch is performed with a 7:1 Buffered Oxide Etch (BOE) solution. Due to the low quality of the deposited oxide, the 1600 Å layer was etched away in just a few seconds.

Finally the third photostep (g) is performed. This is a metal lift off step used to lay down the metal for the ohmic contact. An image reversal photostep was performed to give the resist the necessary undercut. Then, an electron beam evaporator was used to deposit a metal layer composed of Ni/Ge/Au/Ni/Au in thicknesses 50/660/1330/400/1000 Å. The photoresist strip and metal lift-off was performed for 10 minutes in an ultrasound acetone bath. The same metal layer sequence was then deposited on the back side of the wafer. Following a rapid thermal anneal for 30 s at 380°C, these metal layers form the NiGeAu alloyed ohmic contact to the device and the n^+ GaAs substrate. This common method for forming ohmic contacts to n^+ GaAs is discussed extensively in the literature. Some assorted references are given[55, 56, 57]. Following the anneal, the contact was verified by using the TLM test structure. The wafer was then diced and wire bonded to the copper chip carrier.

The description above is of the ideal processing sequence. During the actual fabrication of the A40 device, there were several deviations from the ideal. During the first photostep, the mesa was only etched to 1.3 μm rather than 2 μm. The wafer had to be removed from the etching solution early due to unanticipated etching of the photoresist. It is possible that part of the 2000 Å buffer layer was also etched away before the wafer could be removed. Also, it was found that the NiGeAu contact adhered poorly to the oxide layer. Wire bonding to the bonding paths was thus impossible, so all bonds were made directly to the mesas, where the NiGeAu contact adhered well directly to the GaAs. In future steps, the bonding path can be deposited in a separate photostep than the metal for the ohmic contact. This will allow the use of a Ti adhesion layer between the oxide and the Au.

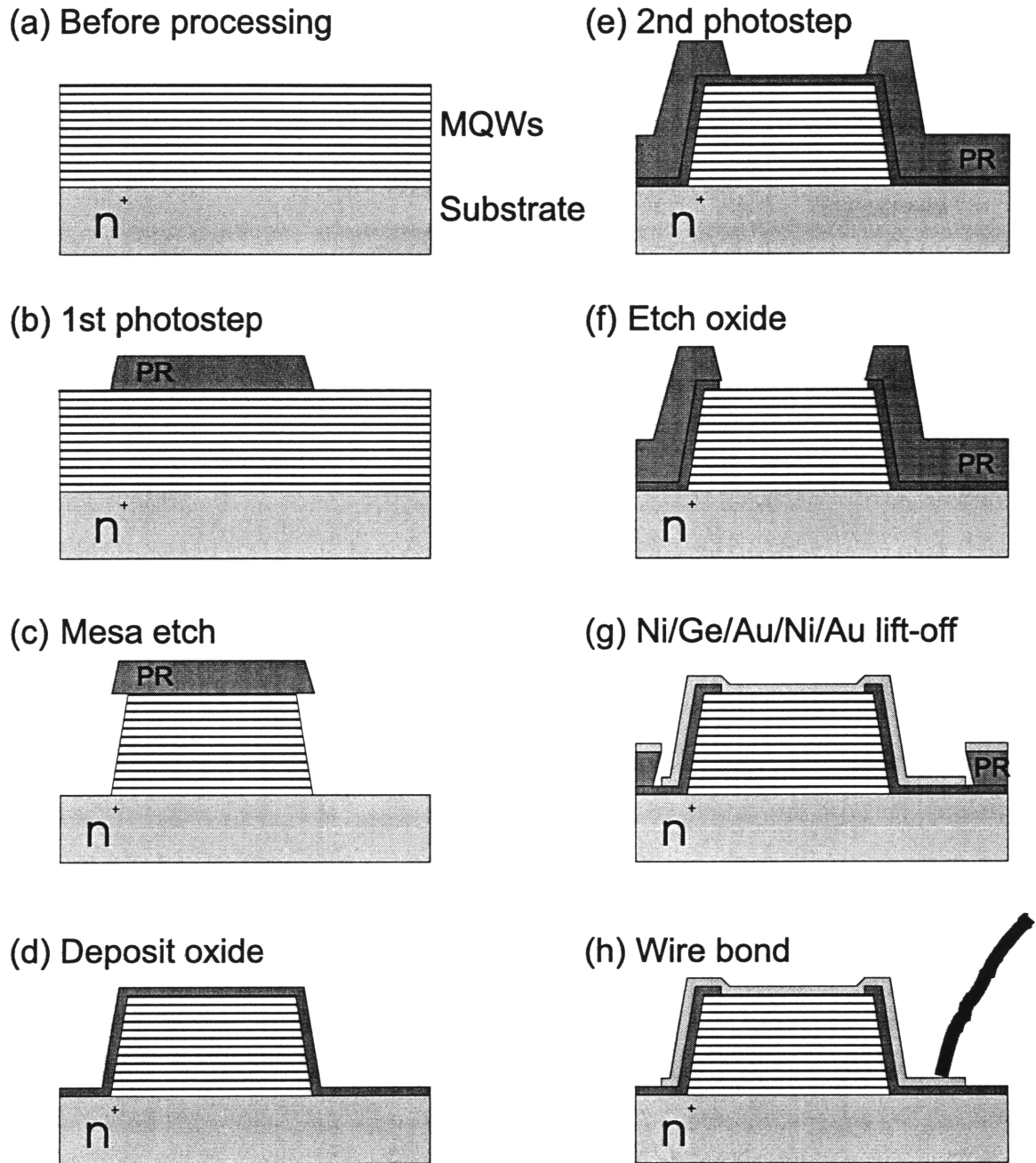


Figure 4-5: Fabrication sequence used for A40 device.

Table 4.1: *Emission devices tested*

Device	Width (μm)	Length (μm)	Comments
A	33	~ 800	Low output power, poor data quality
B	83	~ 450	Good data, narrow linewidth
C	83	~ 700	Cleaved both ends, good data
D	83	~ 660	Cleaved both ends, wider linewidth

The fabrication of the proposed laser structure described in Section 3.4.2 must be performed differently due to the presence of the delicate LTG GaAs layer on top of the wafer. This newly developed fabrication sequence is described in Chapter 5.

4.3 Measurements (A40)

Extensive measurements were performed on our A40 design, including DC-IV measurements of two diode structures, and pulsed optical and current measurements of four edge emitting bar structures. The DC-IV did not yield good data due to excessive heating, and thus discussion of those results will be omitted.

4.3.1 Emission measurements

Four emission devices, described in Table 4.1, were tested and observed to emit spontaneous emission due to intersubband transitions. Presentation of the data will focus on devices B and C, as they yielded the best quality data. Device A, due to its narrower width, emitted less spontaneous emission power, and thus the collected data was much noisier. Device D performed adequately, although the observed linewidths for the emission were somewhat wider than for the other two measured devices. Note that the lengths of the device are measured somewhat crudely, so that they are only accurate within $50 \mu\text{m}$.

Power and current versus bias - Device B

The emitted optical power from Device B was measured simultaneously with the current (Fig. 4-6). The device was operated in pulsed mode using a duty cycle of 1% to avoid excessive heating of the device. This data was collected using a pulse width of $25 \mu\text{s}$ with

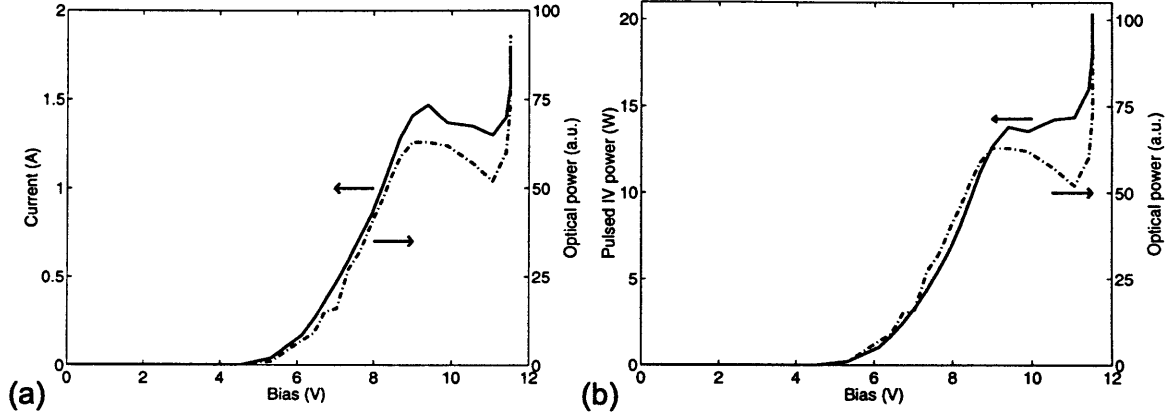


Figure 4-6: Device B: (a) Pulsed current and optical power versus device bias. (b) Pulsed ohmic power $P = IV$ compared with measured optical power.

an electrical chopping frequency of 400 Hz. The temperature of the heat sink remained between 5 K and 8 K during the measurement. Since the heat sink has a finite, and presumably constant, capability to dissipate heat from the device, the measured temperature is a good indicator of the average amount of power dumped into the device.

The current, seen in Fig. 4-6(a), behaves as expected for our design. Essentially no current flows until $V_D = 5$ V, when the energy levels begin to come into alignment. The current flow peaks at $V_{D,op} \cong 9.5$ V, which is in good agreement with the predicted maximum injection voltage of $N_{mod} \times V_{mod,inj} = 40 \times 0.234$ V = 9.36 V (Section 3.3.6). This corresponds to the designed operating point of the device. The current density at this point is $J_{op} \cong 4$ kA/cm². Beyond $V_{D,max}$, a negative differential resistance is observed, as the injection level G becomes misaligned with E_3 . This is a clear indication that the transport through the device is governed by resonant tunneling through the multiple quantum well structure. The presence of the upward jump in current and optical power at high bias will be discussed later in Section 4.3.2.

The emitted optical power also behaves as expected. The spontaneous emission power is expected to be proportional to the number of excited electron N_3 , so as expected, the curve shape tracks that of the current, which is also proportional to N_3 (3.29), and is approximately linear up to $V_{D,op}$. Beyond that, the emitted optical power drops as the levels become misaligned. This data is consistent with intersubband emission being the source of the optical power. Fig. 4-6(b) replots the emitted optical power compared with instan-

taneous ohmic power calculated as $P_{ohmic} = IV_D$. The ohmic power is seen to increase above $V_{D,op}$ even as the emitted power decreases. If blackbody radiation were the major source of observed optical power, the optical power would increase along with the ohmic power. Observation of the heat sink temperature during the measurement confirms that the average power dissipation increases while the emitted power drops. The presence of the upward jump in current and optical power at high bias *is not* due to intersubband emission and will be discussed later in Section 4.3.2.

We can approximately predict the amount of spontaneous emission observed to be

$$P \approx \eta_{coll}\eta_{out} \times \eta_q \times \hbar\omega \times N_m \left(\frac{W}{\alpha_c}\right) \frac{J}{e}, \quad (4.2)$$

where J_{op} is the current density, W is the device width, α_c is the cavity loss, N_m is the number of modules, and $\eta_q \approx 10^{-5}$ is the quantum efficiency. This model assumes that only electrons within one extinction length α_c^{-1} of the edge will contribute to the observed power. η_{out} is the output efficiency, which accounts for the amount of emitted radiation that is coupled out of the device. Since spontaneous emission will emit uniformly in all directions in a plane, and the cone only couples light from one side of the cone, a factor of about $\sim 1/4$ should be accounted for, as well as the 27% facet reflectivity. η_{coll} is the efficiency of the optical collection system that couples light through the FTIR. If we make a conservative estimate of the loss at $\alpha_c = 200 \text{ cm}^{-1}$, we obtain an estimate $P = 1 \text{ } \mu\text{W}$, not considering collection efficiency. The peak emitted power in Fig. 4-6 is roughly several hundred nanowatts, calculated using the lock-in amplifier sensitivity of 0.1 mV and a duty cycle of 1%. This corresponds well with the estimate. Due to uncertainties in converting the lock-in amplitude into a definite input voltage, a more specific estimate cannot be obtained.

Emission spectrum - Device B

The emission spectrum of device B was measured at several different voltages. Fig. 4-7(a) shows three typical spectra at different biases, taken under otherwise identical conditions using 25 μs pulsewidth with 1% duty cycle. The emission narrows and becomes more powerful as the bias approaches the operating point. The spectrum with the narrowest observed

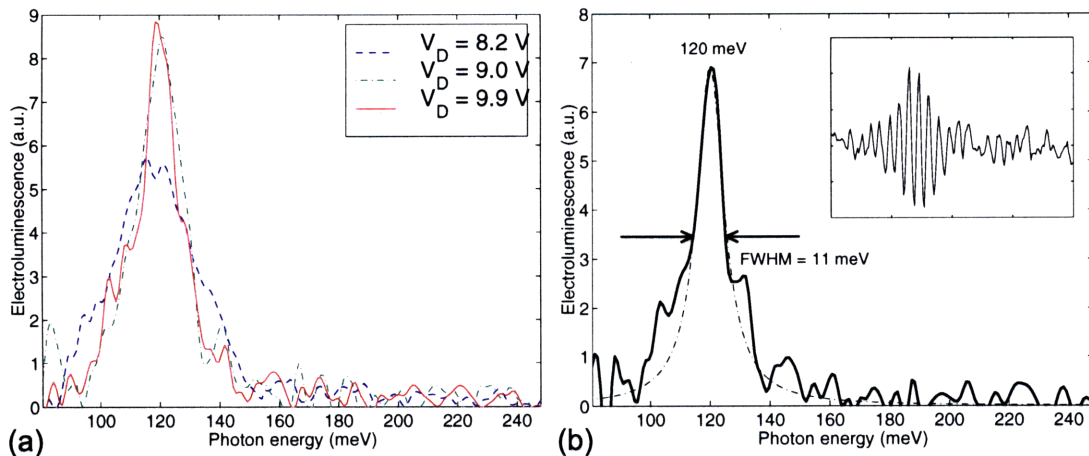


Figure 4-7: *Emission spectra from Device B. (a) Three characteristic spectra at different biases measured with 25 μ s pulse width, 1% duty cycle, at 14 K. (b) Emission spectrum and its Lorentzian fit with narrowest observed linewidth collected using 15 μ s pulse width at $V_D = 9.9$ V. Inset shows the corresponding interferogram.*

linewidth, 11 meV, is shown in Fig. 4-7(b). Lorentzian fits of these spectra show that the peak is at $\hbar\omega_0 = 120 \pm 0.5$ meV, reasonably close to the design value of $E_{32} = 124$ meV. These spectra, in conjunction with the dependence of the optical power on bias, demonstrate that we are observing intersubband emission. The linewidth of the spectra is much too narrow to be from blackbody radiation.

Emission spectrum - Device C

Below we present the emission results for device C. The principal difference between device B and C is that device C is larger and thus has a higher current. Also, device C was cleaved on both ends of the bar rather than just one. Due to the high cavity loss of the A40 design, lasing was not expected, and was not observed. The narrowest observed linewidth for this device was 15 meV compared with 11 meV for B. However, more emission spectra were collected from this device, permitting a more systematic look at the device operation. The optical power versus bias is plotted in Fig. 4-8. The power shows the same general behavior as that measured from device B. The emission peak occurs at $V_D = 10.2$ V, slightly higher than observed with device B, but still in reasonable agreement with the design. Also plotted on the same axes is the average heat sink temperature during the measurement. Past

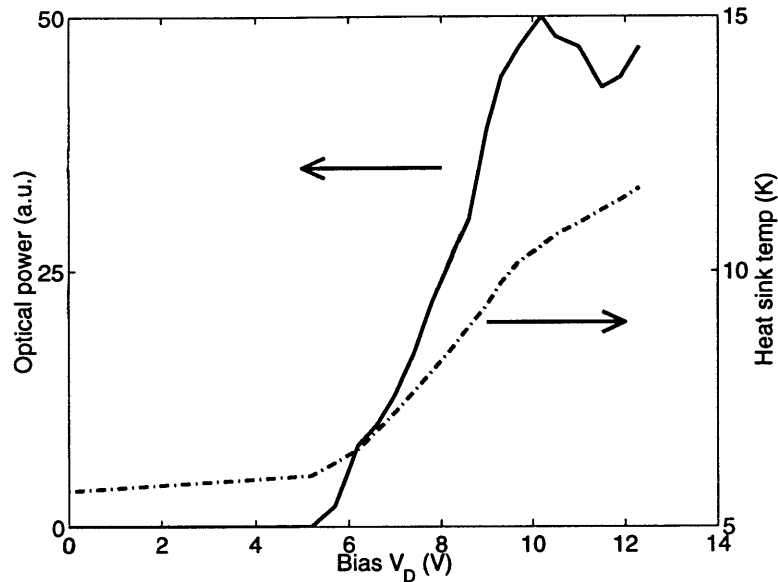


Figure 4-8: *Optical power vs. bias for device C. Also plotted is the heat sink temperature as a gauge of average power dissipation during the measurement. The device was biased with 25 μ s voltage pulses with a 1% duty cycle.*

the optical power peak, the average dissipated power continues to increase, indicating that the bulk of emitted power is due to intersubband transitions rather than heating.

The narrowest observed spectrum had a linewidth of 15 meV and occurred at $V_D = 11.3$ V, somewhat beyond the power peak. This spectrum is shown in Fig. 4-9, with both a Lorentzian and Gaussian fit. Although the data is somewhat noisy, the Lorentzian is clearly the better fit. The origin of this broadening is discussed in Section 5.1.1. These fits are representative of the spectra taken near operating bias.

Spectra were measured at a series of biases from 8 – 12 V. Characteristic spectra and their corresponding Lorentzian fits are shown in Fig. 4-10. Also plotted in the insets are the center frequencies and linewidths for the fitted curves for all collected spectra from device C. Some of the spectra were noisy, hence the scatter in the peak frequency data.

Despite the noise, however, this device displayed a smaller linewidths at higher bias. The reason for this trend is not immediately clear. It seems unlikely that this effect is due to gain narrowing below lasing threshold. If such were the case, the gain narrowing would be accompanied by a superlinear current dependence of the emitted power. Examination of Fig. 4-8 reveals that the power had an approximately a linear voltage dependence.

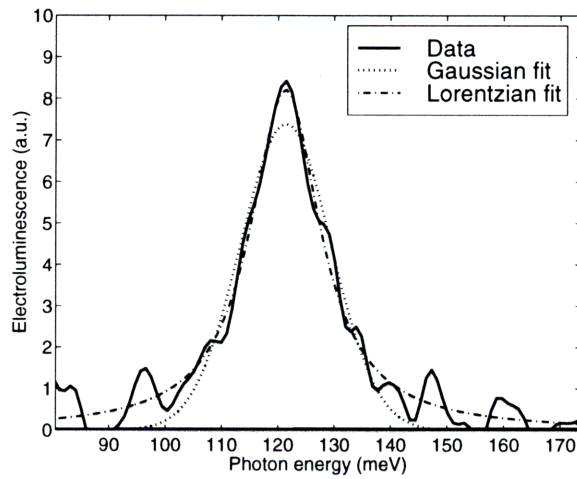


Figure 4-9: Narrowest emission spectra from Device C at $V_D = 11.3$ V, $25 \mu\text{s}$ pulse width, 1% duty cycle. Lorentzian and Gaussian fits to this data are shown for comparison.

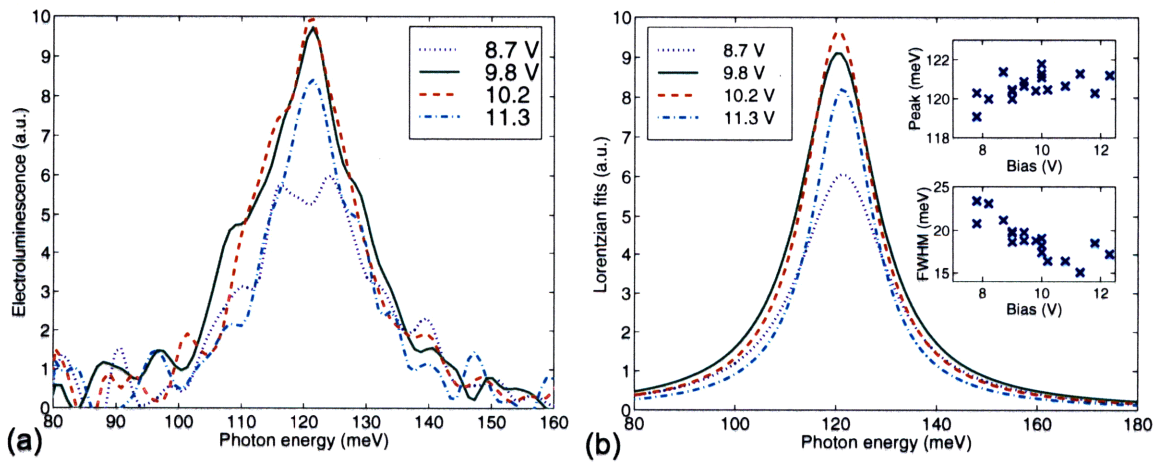


Figure 4-10: Emission spectra from Device C. (a) Four characteristic spectra at different biases measured with $25 \mu\text{s}$ pulse width, 1% duty cycle, at 9-10 K. (b) Lorentzian fits to the four spectra. The two insets show the various peak positions and linewidths measured for all similarly measured device C data.

Unfortunately, due to experimental difficulties, the current was not being simultaneously measured, so the exact current dependence cannot be confirmed. However, other current measurements confirm this approximately linear current dependence, as the electrons populate E_3 , so gain narrowing can likely be ruled out. Also, if the linewidth narrowing were due to the presence of gain, the linewidth wouldn't continue to decrease past the operating point of $V_{D,op} = 10$ V.

One possible explanation is the presence of a background blackbody spectrum. During the application of the bias, the lattice temperature of the device is raised considerably, likely to at least several tens of Kelvin, and possibly to over one hundred Kelvin. Provided the thermal time constant of the device is shorter than the electrical chopping frequency (400 Hz), the blackbody radiation will be detected by the lock-in amplifier, and contaminate the signal. At such low temperatures, the blackbody peak is located in the far-infrared; however, the tail will extend into the spectral range of interest. Several of the spectra exhibit “shoulders” above which a narrower peak extends. It is possible that these shoulders are due to the tail of the blackbody radiation, suppressed on the low-energy side by the reduced response of the detector preceding its 85 meV cutoff. Hence, for very low-power spectra collected at low bias, the broader blackbody spectrum would contribute more towards the measured linewidth. As the intersubband emission power increased, the narrower peak would extend past the background. Once again though, it is not clear how this mechanism would allow the spectrum to narrow past $V_{D,op}$.

4.3.2 Heating

Both device B and C displayed an upward jump in emitted optical power for high biases beyond the operating point (Fig. 4-6,4-8). This jump in output power is not due to intersubband emission, but rather runaway blackbody radiation due to ohmic heating of the device. The spectrum for the emitted radiation for device B at $V_D = 12.5$ V is shown in Fig. 4-11. The emission looks very similar to the broad spectrum of a measured 300 K blackbody spectrum, meaning this high power emission is due to heating of the device during the 25 μ s current pulse. However, the similarity of the spectra *should not* be taken to mean that

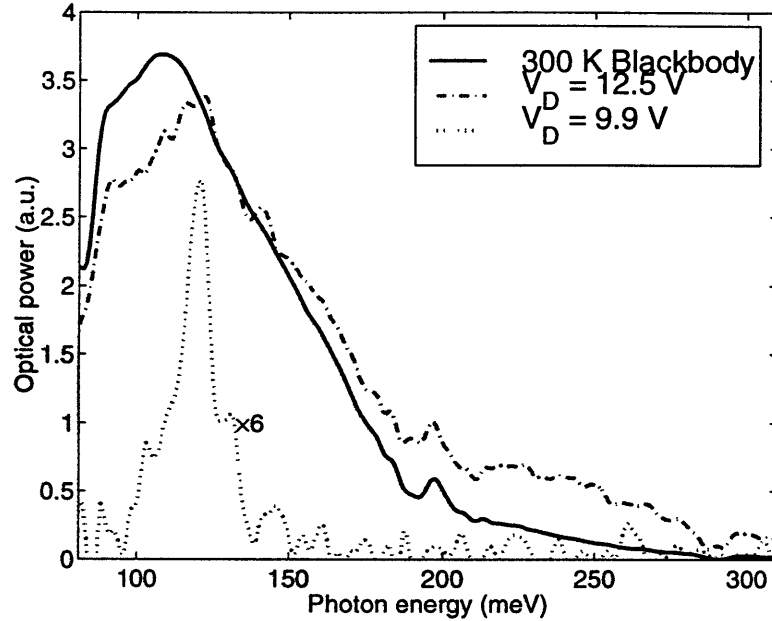


Figure 4-11: *High bias emission spectra from Device B. The emission spectrum for $V_D = 12.5\text{V}$ is shown with a measured 300 K blackbody spectrum for comparison. Note that the peak in the blackbody spectrum is not the true peak, which is located at 75 meV, but rather due to the detector response near the roll-off at 80 meV. The narrow intersubband emission spectrum is shown for comparison.*

the device temperature is 300 K. It is difficult to extract the device temperature from this measurement, as the true blackbody peak (75 meV for 300 K) is below the detector cut-off (80 meV). The observed peak in the 300 K blackbody spectrum is caused by the detector response near the cut-off. The peak of the $V_D = 12.5\text{ V}$ spectrum appears at a higher energy than the pure blackbody spectrum due to the additional intersubband emission present.

The reason this heating power increase appears as a threshold behavior in the optical power characteristic is due to the temperature dependence of the device resistance. As temperature increases, the resistance at a given bias remains constant until some threshold level, when $k_B T$ becomes large enough to excite electrons from the upper state E_3 into the continuum. Beyond this threshold temperature, which is measured to be at least 50 K, the resistance drops rapidly. When the device is biased high enough, ohmic heating will increase the lattice temperature to the threshold point, after which the current will increase and heat the device further. The result is a runaway process that results in a sharp increase in blackbody emission power.

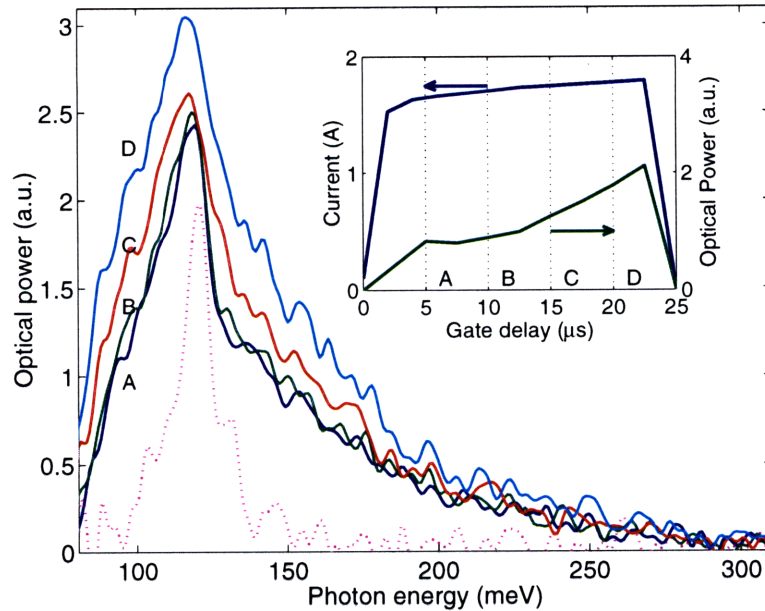


Figure 4-12: *Time-resolved high bias emission spectra from Device B. A gated boxcar averager was used to collect the spectra during 4 sequential 5 μs periods of the 25 μs pulse. A narrow intersubband emission spectrum (scaled differently) is shown for linewidth comparison. The inset shows the change in the current and emitted optical power over the course of the pulse.*

This process was probed by examining the current and optical power behavior during the course of the 25 μs pulse using a gated boxcar averager (Fig. 4-12, inset). The bias remains constant while the current and optical power increase. The gate delay on boxcar was adjusted to collect spectra over four separate 5 μs intervals during the pulse. As shown in Fig. 4-12, blackbody radiation becomes more and more prominent, overwhelming the spontaneous emission, broadening the spectrum, and shifting the observed spectral peak towards lower energies. This measurement illustrates the importance of good heat sinking. Unfortunately, with the current experimental setup, we are unable to bias the devices with pulses shorter than 10 μs . In many measurements, Faist used pulses as short as 20 ns to eliminate lattice heating [6]. Also, reducing the device size will improve matters by reducing the total current. In these measurements, we primarily tested 83 μm bar structures. Narrower structures provided too little output power. Since laser structures must be made much longer to reduce facet loss, they must be made as narrow as possible to minimize total current. With the available mask, the narrowest bar structures available are 33 μm wide. A new mask has been designed by B. Riely, and will be used on the next processing

run to provide us with structures down to 6 μm wide, thus reducing heating.

Chapter 5

Discussion

5.1 A40 Results

The results of the measurements performed on the A40 design were quite promising. The I-V data is consistent with transport through the multiple quantum well structure by resonant tunneling. As shown in Fig. 4-6(a), current switched on at some bias point where the subbands began to align, and quickly built up to a peak corresponding to peak electron injection into level E_3 . The current dropped once biased past the designed operating point due to misalignment of the subbands. The current peak was observed approximately where expected, at $V_D \approx 9.5$ V for device B compared to the predicted value of $V_{D,inj} = 9.36$ V.

Also, spontaneous emission due to intersubband transitions was observed at a center frequency of $\hbar\omega_0 = 120$ meV, very close to the calculated value of $E_{32} = 124$ meV for the transition. The frequency was largely independent of bias, consistent with the primarily vertical transition that was designed. This result gives us confidence in our ability to design multiple quantum well structures taking account of nonparabolicity.

As shown in Fig. 4-6 the current density at the operating point is $J_{op} = 4$ kA/cm². Using (3.29) with τ_3 , we can estimate the excited state population $N_3 \approx 1.75 \times 10^{10}$ cm⁻², with a population inversion $\Delta N_{32} \approx 1.5 \times 10^{10}$ cm⁻². This is a somewhat lower population than expected. The design of the MQW structure was intended that τ_3 would be the longest lifetime in the system, with the other levels in the active region and digitally graded gap scattering much more quickly due to LO-phonon resonance. It is possible that the total

transit time through the digitally graded gap is being slowed, either due to misalignment of the energy levels or reduction of phase breaking due to impurity scattering from the setback doping. Also, the injection barrier could be too large, inhibiting tunneling into the active region.

5.1.1 Linewidth broadening

Linewidths as narrow as 11 meV FWHM were observed for the emission spectra from device B, although they were somewhat larger for device C (15 meV). Although the data was somewhat noisy, the emission peaks at operating bias were best fit by Lorentzian functions. Transition broadening results from the finite phase coherence lifetime of each state, which requires an uncertainty in the energy level according to $\Delta E \cong \hbar/\tau$. For homogeneous broadening, all electrons involved in transition experience the same interaction potential, and are thus broadened according to a Lorentzian lineshape. The total transition broadening $\Delta E = \Delta E_3 + \Delta E_2$ is determined by the sum of the linewidths from each state, i.e. the sum of all the phase-breaking rates that interrupt the coherent atom-photon interaction(2.20). This phase breaking can be inelastic, involving a transition to a new state, in a different subband, within the same subband, or it can be elastic, breaking the phase without changing the energy.

There are several line broadening mechanisms for our intersubband transition. Electrons can escape the subband due to tunneling, or scattering by LO-phonons. Other inelastic scattering events such as LO-phonon scattering, acoustic phonon scattering, and electron-electron scattering can occur within the subband, breaking phase but leaving the state lifetime unaffected. Elastic scattering such as impurity and interface roughness are also often significant factors in intersubband linewidths. Also, if there is a significant electron distribution over the subband, nonparabolicity will asymmetrically broaden a transition. A MBE growth related issue is that of many-module non-uniformity, i.e. the case where each module emits at a slightly different frequency.

Let us consider each of these effects on our intersubband transition. The lifetimes of E_3 and E_2 are 0.7 ps and 0.2 ps respectively, which are dominated by LO-phonon emission.

The lower state lifetime can also be lowered by tunneling into the digitally graded region, but this is likely slower than the LO-phonon resonance. The total broadening resulting from these lifetimes, 4.2 meV, is insufficient to account for all of the broadening.

It is difficult to calculate the phase breaking rate of intrasubband processes, although they almost certainly play a significant role. Electron-electron scattering is essential in maintaining thermal equilibrium in the electron gas, and electron-phonon scattering maintains equilibrium between the electron gas and the lattice. Let us consider the two states separately for the moment. During operation, the excited E_3 state has a significant electron population with a Fermi distribution set by some electron temperature not necessarily the same as the lattice temperature. This electron population will engage in fast (~ 0.1 ps) intrasubband electron-electron scattering. At the operating point, the excited state population was estimated to be approximately $N_3 \approx 2 - 3 \times 10^{10} \text{ cm}^{-2}$. If the electrons were at zero temperature, only the small fraction of electrons at the quasi-Fermi level would participate in the scattering due to state blocking. However, the electron temperature in the subbands is likely elevated to at least several tens of Kelvin. For the estimated excited state population, the quasi-Fermi level lies below the subband edge, leaving many subbands states available for electrons to scatter into. To obtain some average phase breaking rate, one must average the fast (subpicosecond) electron-electron collision rate with the fraction of electrons in the subband that undergo such scattering events. This is quite a difficult calculation and will not be undertaken here. Acoustic phonon scattering is expected to be at least 10 ps and thus unimportant at cryogenic temperatures. Intrasubband LO-phonon scattering is also unimportant as a phase breaking mechanism. In the excited state the majority electron population always resides in the lowest 5-10 meV, making intrasubband LO-phonon scattering impossible.

For the depopulated E_2 state, electron-electron scattering is much less important. The same arguments stated above hold for the unimportance of intrasubband acoustic and LO-phonon scattering. Although most electrons scattered into E_2 from E_3 will emit several LO-phonons in the process of cooling to the bottom of the subband, these events also do not contribute to the linewidth. Because the radiative transition is vertical, and the nonparabolicity small, only scattering events within the lowest 5 meV of E_2 are relevant.

It is unlikely that impurity scattering is a major cause of broadening in our system. The MQW well module was designed in such a way that the dopants were set back over 100 Å in the digitally graded gap for the purpose of separating them from the double well active region (Fig. 3-1). The relative importance of impurity scattering on the intersubband electroluminescence linewidth was investigated by Faist *et al.* [58]. Comparing the emission of two devices, one doped in the active region, and one with setback doping, Faist *et al.* observed the electroluminescence spectrum change from a Gaussian lineshape (FWHM = 50 meV) to a much narrower Lorentzian spectrum (FWHM = 21 meV) after dopants were set back from the active region. Impurity scattering is expected to cause inhomogeneous broadening due to the large distance between impurities (~ 200 Å) and it can be reduced by using set back doping. Faist attributed the remaining linewidth broadening to interface roughness scattering, which acts as a homogeneous broadening mechanism, due to the shorter length scale of the interface roughness[59]. This length scale is given by the average terrace width due to MBE growth, estimated to be about 50 – 70 Å[60]. This is much smaller than the minimum deBroglie wavelength ($\lambda_b = 2\pi/\sqrt{2\pi N_3} \approx 600$ Å) of the electron making the transition. Although the noise of our collected data makes an exact fit impossible, our observed electroluminescence at operating voltage is definitely better fit with a Lorentzian than with a Gaussian, suggesting homogeneous broadening.

Broadening due to interface roughness scattering is likely a factor in our system. This broadening is caused by growth imperfections, but its importance can be reduced by careful engineering of the transition. During our design, we attempted to obtain a primarily vertical intrawell transition so that the transition would not occur through a barrier, increasing the effect of interface roughness. However, interface roughness is still a factor, since a portion of the wavefunction extends into the second well across the thin 4 ML central barrier (see Fig. 3-2). The small single monolayer width fluctuations in the central barrier will have an exaggerated effect since the fractional width fluctuations are greater than any other barrier.

It is unlikely that nonparabolicity is a large source of broadening. As described by Gelmont *et al.* , such asymmetric broadening occurs only at high electron temperature. Gelmont calculated approximately 5 meV broadening for a 300 meV transition in a InGaAs quantum well with an electron temperature of 100 K. Although our electron temperature

might be as high as 100 K, nonparabolicity is less pronounced in GaAs, and our energy level separation is only 120 meV; the expected broadening would be reduced accordingly. It is possible this effect makes a small contribution, but the noise present in observed spectra makes a quantitative analysis impossible.

Many-module non-uniformity is also a possible source of broadening. The result would be certain modules would emit with Lorentzian lineshapes at different frequencies. The observed lineshape would be the sum of all these lineshapes. Such a situation would likely not yield a Lorentzian lineshape, so this explanation is likely not a major effect.

Our observed linewidths compare well with others seen in quantum wells. Two other groups also pursuing a GaAs quantum cascade laser have reported spontaneous emission linewidths of 15 meV (Strasser *et al.*) and 14 meV (Li *et al.*) [18, 19]. However, Li's structure is based on a diagonal transition, rather than a vertical intrawell transition, rendering it more susceptible to interface roughness broadening. Gauthier-Lafaye *et al.* observed a 8 meV photoluminescence in their optically pumped GaAs/AlGaAs double well structure at 77 K. Our linewidth is similar to that measured by Faist *et al.* in a vertical transition structure (12.4 meV) [43].

In summary, only about 4 – 5 meV of the narrowest observed 11 meV broadening can be accounted for by the finite state lifetime due to LO-phonon intersubband scattering. The rest of the observed broadening is likely due to phase breaking intrasubband electron-electron collisions and interface roughness scattering. These collisions would have to occur at a total rate of $(0.1 \text{ ps})^{-1}$ to account for the extra broadening, which is a reasonable number. The observed broadening appears to be more closely Lorentzian shaped than Gaussian, implying that the transition is homogeneously broadened. This suggests nonparabolicity and impurity scattering are not important broadening mechanisms. The measured linewidth compares favorably with those measured in similar structures in GaAs and InGaAs. Since the peak gain is inversely proportional to the linewidth of the transition, the narrow measured linewidth promises lower threshold currents for a quantum cascade laser developed using this MQW structure as a gain medium.

5.2 Laser design PL50

The measured results described above show promise for the development of a GaAs based quantum cascade laser. With a few small modifications, the tested MQW structure can be used as the gain medium and surrounded by the appropriate waveguide cladding. We propose increasing the number of modules to 50, making the active region thickness $2.5 \mu\text{m}$. The operating voltage will then be roughly 12 V. Also, to increase the population inversion, the doping per module is increased to $N_{D,mod} = 2 \times 10^{11} \text{ cm}^{-2}$. The additional charge results in only a small amount of band bending, and negligible shift in the energy levels. Redesign of the MQW structure should be unnecessary, and will be kept the same as in the A40 structure. This MQW core active region can be placed inside the newly proposed waveguide structure described in Section 3.4.2. This structure has contacts doped at $N_D = 2 \times 10^{18} \text{ cm}^{-3}$, which should improve conduction into the active region. The MBE growth diagram for this proposed structure, called PL50, is shown in Fig. 5-1. This device is grown with a non-alloyed ohmic contact as described by Patkar *et al.* [33]. The top 70 Å is Si doped at $N_D = 5 \times 10^{19} \text{ cm}^{-3}$. The upper 30 Å of this region will remain at this doping as the Fermi level is pinned by midgap defect states in the low temperature grown (LTG) cap layer. The rest of the heavily doped layer will reduce to $N_D - N_A \approx 5 \times 10^{18} \text{ cm}^{-3}$ as Si donors become acceptors in the bulk. The Ti/Au contact layer is deposited on top of the LTG GaAs; conduction occurs by tunneling through midgap states. This use of a nonalloyed ohmic contact reduced the loss due to the diffusion of dopants into the bulk during the annealing of a traditional Ni/Ge/Au ohmic contact. Non alloyed ohmic contacts don't possess the rough surface morphology of their alloyed counterparts, and are therefore more suitable for all types of optoelectronic devices. Unfortunately, a rapid thermal anneal step is still necessary for the backside contact, which might cause some minor diffusion of impurities in from the surface.

We have developed a modified processing sequence for use with PL50 laser design, which is depicted in Fig. 5-2. In order to protect the thin LTG GaAs layer, the Ti/Au metal layer for the contacts is deposited in the first step (a), and serves also as a mask for the mesa etch. This self aligned process allows the fabrication of much narrower bar structures. A

dry plasma etch step in an electron-cyclotron resonance reactive ion etching (ECR-RIE) machine has been substituted for the wet etch (b). This step allows us to define mesas with vertical sidewalls and no undercut of the metal mask. The rest of the process is largely the same as for the A40 design. 2000 Å of oxide is deposited, and vias are etched to connect with the upper contact (d), (e), (f). Ti/Au bonding paths are then electron-beam evaporated at an angle to ensure coverage of one sidewall.

The gain of our proposed design can be calculated using (2.26), and is $\gamma_{bulk} = 127 \text{ cm}^{-1}$. For this gain calculation, the conservative linewidth of 15 meV was used, due to the fact that not all tested devices displayed the narrowest 11 meV linewidth. For the TM_0 mode calculated in Section 3.4.2, $\Gamma = 0.85$, yielding a peak modal gain as high as $\gamma = 108 \text{ cm}^{-1}$. This waveguide structure has a calculated cavity loss of $\alpha_c = 64 \text{ cm}^{-1}$. If the ends are cleaved to form a 2 mm long Fabry-Perot waveguide ($R = 0.27$), the facet loss will be $\alpha_f = 6.5 \text{ cm}^{-1}$. There will be some additional reduction in the confinement factor due to lateral confinement. According to this calculation lasing should be attainable from this structure with a threshold current density of $J_{th} \approx 2.7 \text{ kA/cm}^2$.

The high bias results discussed in Section 4.3.2 reveal the importance of good heat sinking for the device. Due to the higher doping, the PL50 laser design will presumably have 4/3 of the current density of the A40. Also, in order to minimize facet loss, the laser structures tested will be much longer than the devices reported upon here. To reduce the total current consumed, it will be necessary to use much narrower test structures. A new mask has been designed by Brian Riely, and will be used on the next processing run to provide us with structures as narrow as 6 μm , thus reducing heating.

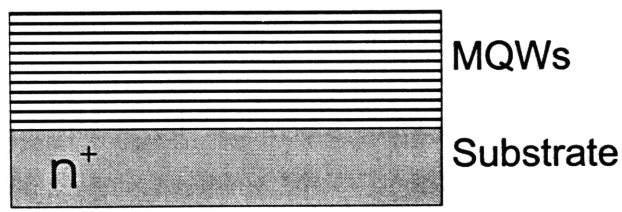
5.3 Conclusion

As of this writing, the following steps have been completed. A mid-infrared GaAs/AlGaAs quantum cascade emitter has been designed, fabricated, and tested. It was found to behave as expected and is suitable for use for the gain medium of a quantum cascade laser. A waveguiding structure was designed and simulated. The proposed PL50 laser structure, incorporating the MQW gain region from the tested A40 design, and the newly designed

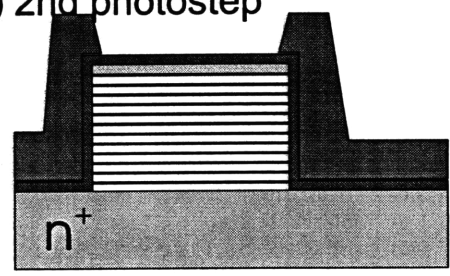
250° C	i-GaAs	undoped	35 A	
	GaAs	$5 \times 10^{19} \text{ cm}^{-3}$	70 A	
	GaAs	$2 \times 10^{18} \text{ cm}^{-3}$	900 A	
	Al _{0.4} Ga _{0.6} As	undoped	25.4 A (9 ML)	
	GaAs	undoped	36.7 A (13 ML)	
	Al _{0.4} Ga _{0.6} As	undoped	25.4 A (9 ML)	
	GaAs	undoped	31.1 A (11 ML)	
	Al _{0.4} Ga _{0.6} As	undoped	25.4 A (9 ML)	
	GaAs	$2 \times 10^{17} \text{ cm}^{-3}$	25.4 A (9 ML)	
	Al _{0.4} Ga _{0.6} As	$2 \times 10^{17} \text{ cm}^{-3}$	25.4 A (9 ML)	
	GaAs	$2 \times 10^{17} \text{ cm}^{-3}$	25.4 A (9 ML)	
	Al _{0.4} Ga _{0.6} As	$2 \times 10^{17} \text{ cm}^{-3}$	22.6 A (8 ML)	
	GaAs	undoped	25.4 A (9 ML)	
	Al _{0.4} Ga _{0.6} As	undoped	25.4 A (9 ML)	
	GaAs	undoped	28.3 A (10 ML)	
	Al _{0.4} Ga _{0.6} As	undoped	42.4 A (15 ML)	
	Repeat 50 times	GaAs	undoped	67.8 A (24 ML)
		Al _{0.4} Ga _{0.6} As	undoped	11.3 A (4 ML)
		GaAs	undoped	53.7 A (19 ML)
Al _{0.4} Ga _{0.6} As		undoped	25.4 A (9 ML)	
GaAs		undoped	36.7 A (13 ML)	
Al _{0.4} Ga _{0.6} As		undoped	25.4 A (9 ML)	
GaAs		undoped	31.1 A (11 ML)	
Al _{0.4} Ga _{0.6} As		undoped	25.4 A (9 ML)	
GaAs		$2 \times 10^{17} \text{ cm}^{-3}$	25.4 A (9 ML)	
Al _{0.4} Ga _{0.6} As		$2 \times 10^{17} \text{ cm}^{-3}$	25.4 A (9 ML)	
GaAs		$2 \times 10^{17} \text{ cm}^{-3}$	25.4 A (9 ML)	
Al _{0.4} Ga _{0.6} As		$2 \times 10^{17} \text{ cm}^{-3}$	22.6 A (8 ML)	
GaAs		undoped	25.4 A (9 ML)	
Al _{0.4} Ga _{0.6} As		undoped	25.4 A (9 ML)	
GaAs		undoped	28.3 A (10 ML)	
Al _{0.4} Ga _{0.6} As		undoped	42.4 A (15 ML)	
GaAs		$2 \times 10^{18} \text{ cm}^{-3}$	2 μm	
	n+ GaAs substrate			

Figure 5-1: MBE growth diagram for proposed PL50 laser structure.

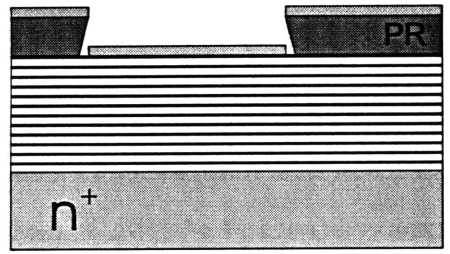
(a) Before processing



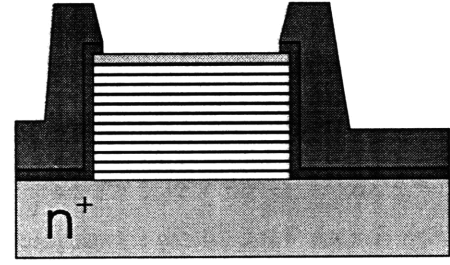
(e) 2nd photostep



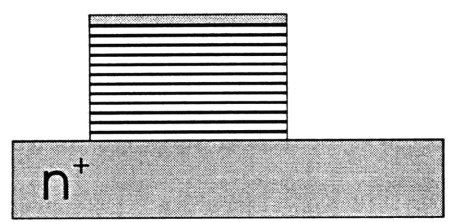
(b) 1st photostep - Ti/Au lift-off



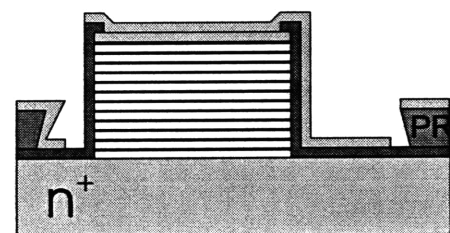
(f) Etch oxide



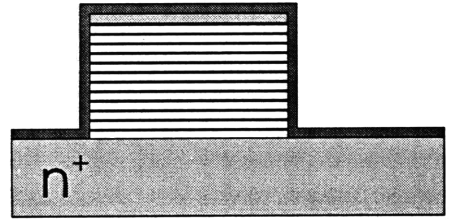
(c) Mesa dry etch - Au mask



(g) Angled Ti/Au lift-off



(d) Deposit oxide



(h) Wire bond

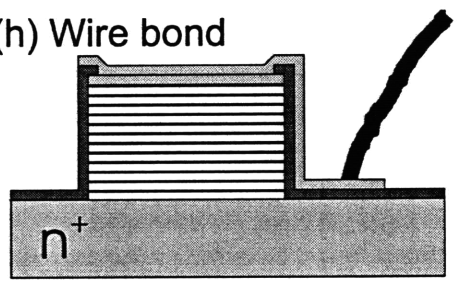


Figure 5-2: Processing sequence for PL50 laser design.

waveguide has been fabricated and is currently awaiting testing. Calculations indicate that lasing should be possible from this device. Additional accomplishments include development of an experimental setup for testing the device under cryogenic conditions. A device holder and heatsink was designed to hold the device inside the dewar aligned with the output optics. A bias supply was successfully designed to provide the device with short, high current voltage pulses. Also, several software tools were developed to aid the design of the laser, including a one-dimensional slab waveguide mode solver, and modifications to our multiple quantum well numerical simulation program SEQUAL to increase its capability.

Work on this project is continuing in an attempt to produce a structure that lases. Calculations indicate that the plasma-metal waveguide can be successfully used for a laser operated at lower temperatures. However, this design is still quite lossy, and the rapid increase of loss with temperature is very undesirable. For long term laser development, a waveguiding method utilizing dielectric confinement rather than plasma confinement must be found. One likely candidate, currently being explored by Li *et al.* [19], is the use of large AlAs cladding layers, graded to form a smooth interface with the GaAs wells.

Once a laser has been successfully fabricated, the structure can be optimized to operate at higher power at higher temperatures, with lower threshold currents. The eventual goal is to introduce tunability via oscillator strength tuning [12], making the GaAs based QCL suitable for use in a high power room temperature heterodyne remote sensing spectroscopy system.

Bibliography

- [1] M. Tacke. New developments and applications of tunable ir lead alt lasers. *Infrared Phys. Technol.*, 36:447–463, 1995.
- [2] L. Esaki and R. Tsu. Superlattice and negative differential conductivity in semiconductors. *IBM J. Res. Dev.*, 14:61, 1970.
- [3] 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.
- [4] 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.
- [5] S. Luryi. Hot-electron injection and resonant-tunneling heterojunction devices. In F. Capasso and G. Margaritondo, editors, *Heterojunction Band Discontinuities: Physics and Device Applications*, pages 489–564. North Holland Publishing, 1987.
- [6] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho. Quantum cascade laser. *Science*, 264:553, 1994.
- [7] Q. Hu and S. Feng. Feasibility of far-infrared lasers using multiple semiconductor quantum wells. *Appl. Phys. Lett.*, 59:2923–2925, 1991.
- [8] Bin Xu, Qing Hu, and Michael R. Melloch. Electrically pumped tunable terahertz emitter based on intersubband transition. *Appl. Phys. Lett.*, 71:440, 1997.
- [9] C. Sirtori, J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchison, and Alfred Y. Cho. Pulsed and continuous-wave operation of long wavelength infrared ($\lambda = 9.3 \mu\text{m}$) quantum cascade lasers. *IEEE Journal of Quantum Electronics*, 33:89, 1997.
- [10] G. Scamarcio, F. Capasso, C. Sirtori, J. Faist, A. L. Hutchison, D. L. Sivco, and A. Y. Cho. High-power infrared (8-micrometer wavelength) superlattice lasers. *Science*, 276:773, 1997.
- [11] J. Faist, A. Tredicucci, Federico Capasso, C. Sirtori, D. L. Sivco, J. N. Baillargeon, A. L. Hutchinson, and A. Y. Cho. High-power continuous-wave quantum cascade lasers. *IEEE J. Quantum Electron.*, 34:336, 1998.

- [12] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho. Laser action from the ground state by oscillator strength tuning. *Nature*, 387, 1997.
- [13] K. Namjou, S. Cai, E. A. Whittaker, J. Faist, C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho. Sensitive absorption spectroscopy with a room-temperature distributed-feedback quantum-cascade laser. *Optics Lett.*, 23:219–221, 1998.
- [14] S. Slivken, C. Jelen, A. Rybaltowski, J. Diaz, and M. Razeghi. Gas-source molecular beam epitaxy growth of an 8.5 μm quantum cascade laser. *Appl Phys. Lett.*, 71:2593–2595, 1997.
- [15] O. Gautheri-Lafaye, P. Boucaud, F. H. Julien, S. Sauvage, S. Cabaret, J.-M. Lourtioz, V. Thierry-Mieg, and R. Planel. Long wavelength ($\approx 15.5 \mu\text{m}$) unipolar semiconductor laser in GaAs quantum wells. *Appl. Phys. Lett.*, 71:3619–3621, 1997.
- [16] C.-H. Lin, R. Q. Yang, D. Zhang, P. C. Chang, Y. Zhou, S. S. Pei, J. I. Malin, C. L. Felix, J. R. Meyer, C. A. Hoffman, and J. F. Pinto. *IEE Electron. Lett.*, 33:598, 1997.
- [17] R. Q. Yang, B. H. Yang, D. Zhang, C.-H. Lin, S. J. Murry, H. Wu, and S. S. Pei. High power mid-infrared interband cascade lasers based on type-II quantum wells. *Appl Phys. Lett.*, 71:2409–2411, 1997.
- [18] G. Strasser, P. Kruck, M. Helm, J. N. Heyman, L. Hvozدارa, and E. Gornik. Mid-infrared electroluminescence in GaAs/AlGaAs structures. *Appl. Phys. Lett.*, 71:2892, 1997.
- [19] Y. B. Li, J. W. Cockburn, M. S. Skolnick, M. J. Birkett, J. P. Duck, R. Grey, and G. Hill. Intersubband electroluminescence in GaAs/AlGaAs quantum cascade structures. *IEE Electronics Lett.*, 33:1874–5, 1997.
- [20] A. Yariv. *Quantum Electronics*. John Wiley & Sons, New York, third edition, 1989.
- [21] 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.
- [22] 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, 1996.
- [23] 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.
- [24] S. D. Brorson, H. Yokoyama, and E. P. Ippen. Spontaneous emission rate alteration in optical waveguide structures. *IEEE J. of Quantum Electron.*, 26:1492, 1990.
- [25] A. E. Siegman. *Lasers*. University Science Books, Mill Valley, California, first edition, 1986.

- [26] J. T. Verdeyen. *Laser Electronics*. Prentice Hall, Englewood Cliffs, NJ, third edition edition, 1995.
- [27] 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.
- [28] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flanery. *Numerical Recipes in C*, section 9.7, pages 383–393. Cambridge University Press, Cambridge, second edition, 1992.
- [29] B. Riely. Personal communication.
- [30] S. L. Chuang. *Physics of Optoelectronic Devices*, section 7.6, pages 273–278. Wiley Series in Pure and Applied Optics. John Wiley and Sons, New York, first edition, 1995.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] J. S. Blakemore. Semiconducting and other major properties of gallium arsenide. *J. Appl. Phys.*, 53:R123, 1982.
- [35] 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.
- [36] H. Rucker, E. Molinari, and P. Lugli. Microscopic calculation of the electron-phonon interaction in quantum wells. *Phys. Rev. B*, 45:6747, 1992.
- [37] M. McLennan. *SEQUAL Users Manual*. Purdue University, Lafayette, 1989.
- [38] D. F. Nelson, R. C. Miller, and D. A. Kleinman. Band nonparabolicity effects in semiconductor quantum wells. *Phys. Rev. B*, 35:7770, 1987.
- [39] R. P. Leavitt. Empirical two-band model for quantum wells and superlattices in an electric field. *Phys. Rev. B*, 44:11270, 1991.
- [40] G. Bastard. Superlattice band structure in the envelope-function approximation. *Phys. Rev. B*, 24:5693, 1981.

- [41] G. Bastard. Theoretical investigations of superlattice band structure in the envelope-function approximation. *Phys. Rev. B*, 25:7584, 1982.
- [42] H. C. Liu, M. Buchanan, and Z. R. Wasilewski. How good is the polarization selection rule for intersubband transitions. *Appl. Phys. Lett.*, 72:1682–1684, 1998.
- [43] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho. Vertical transition quantum cascade laser with bragg confined excited state. *Appl. Phys. Lett.*, 66:538, 1995.
- [44] C. Sirtori, J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchison, and A. Y. Cho. Long wavelength infrared ($\lambda = 11\mu\text{m}$) quantum cascade lasers. *Appl. Phys. Lett.*, 69:2810, 1996.
- [45] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, S. N. G. Chu, and A. Y. Cho. Mid-infrared field-tunable intersubband electroluminescence at room temperature by photon-assisted tunneling in coupled-quantum wells. *Appl. Phys. Lett.*, 64:1144, 1994.
- [46] E. E. Mendez. Physics of resonant tunneling in semiconductors. In E. E. Mendez and K. von Klitzing, editors, *Physics and applications of quantum wells and superlattices*, number 170 in NATO ASI series. Series B, Physics.
- [47] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho. Continuous wave operation of a vertical transition quantum cascade laser above $T = 80\text{ K}$. *Appl. Phys. Lett.*, 67:3057, 1995.
- [48] Pearson and Shah. *High speed semiconductor devices*. Wiley, New York, 1990.
- [49] 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.
- [50] T.-B. Ng, D. B. Janes, D. McInturff, and J. M. Woodall. Inhibited oxidation in low-temperature grown GaAs surface layers observed by photoelectron spectroscopy. *Appl. Phys. Lett.*, 69:3551–3553, 1996.
- [51] D. R. Lide, editor. *CRC Handbook of Chemistry and Physics*, pages 12–33. CRC Press, Boca Raton, 72nd edition, 1991.
- [52] D. W. Lynch and W. R. Hunter. Gold. In E. D. Palik, editor, *Handbook of optical constants of solids*, Academic Press Handbook Series, chapter 4, pages 286–295. Academic Press, Orlando, 1985.
- [53] W. T. Welford and R. Winston. *High collection nonimaging optics*. Academic Press, San Diego, first edition, 1989.

- [54] 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.
- [55] V. L. Rideout. A review of the theory and technology for ohmic contacts to group III-V compound semiconductors. *Solid-State Electronics*, 18:541–550, 1975.
- [56] M. N. Yoder. Ohmic contacts in GaAs. *Solid-State Electronics*, 23:117–119, 1980.
- [57] N. Braslau. Alloyed ohmic contacts to GaAs. *J. Vac. Sci. Technol.*, 19:803, 1981.
- [58] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, S. N. G. Chu, and A. Y. Cho. Narrowing of the intersubband electroluminescent spectrum in couple-quantum-well heterostructures. *Appl. Phys. Lett.*, 65:94, 1994.
- [59] J. Faist, C. Sirtori, F. Capasso, L. Pfeiffer, and K. W. West. Phonon limited intersubband lifetimes and linewidths in a two-dimensional electron gas. *Appl. Phys. Lett.*, 64:872–874, 1994.
- [60] H. Sakaki, T. Noda, K. Hirakawa, M. Tanaka, and T. Matsusue. Interface roughness scattering in GaAs/AlGaAs quantum wells. *Appl. Phys. Lett.*, 51:1934, 1987.