Development of Terahertz Quantum-Cascade Lasers

by

Sushil Kumar

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

April 2007
[June 2007]

Author .....
Department of Electrical Engineering and Computer Science
April 11, 2007

Certified by..
.....
Qing Hu
Professor
Thesis Supervisor

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

Development of Terahertz Quantum-Cascade Lasers

by

Sushil Kumar

Submitted to the Department of Electrical Engineering and Computer Science
on April 11, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The terahertz or the far-infrared frequency range of the electromagnetic spectrum ($\nu \sim 1 - 10$ THz, $\lambda \sim 300 - 30$ μm, $\hbar\omega \sim 4 - 40$ meV) has historically been technologically underdeveloped despite having many potential applications, primarily due to lack of suitable sources of coherent radiation. Following on the remarkable development of mid-infrared ($\lambda \sim 3 - 30$ μm) quantum-cascade lasers (QCLs) in the past decade, this thesis describes the development of electrically-pumped terahertz quantum-cascade lasers in GaAs/$Al_xGa_{1-x}As$ heterostructures that span a spectral range of $1.59 - 5.0$ THz ($\lambda \approx 190 - 60$ μm).

A quantum-cascade laser (QCL) emits photons due to electronic intersubband transitions in the quantum-wells of a semiconductor heterostructure. The operation of terahertz QCLs at frequencies below the *Reststrahlen* band in the semiconductor ($\sim 8 - 9$ THz in GaAs), is significantly more challenging as compared to that of the mid-infrared QCLs. Firstly, due to small energy separation between the laser levels various intersubband scattering mechanisms are activated, which make it difficult to selectively depopulate the lower laser level. Additionally, as electrons gain enough kinetic energy in the upper laser level thermally activated longitudinal-optical (LO) phonon scattering reduces the level lifetime and makes it difficult to sustain population inversion at higher temperatures. Secondly, waveguide design for terahertz mode confinement is also more challenging due to higher free-carrier losses in the semiconducting doped regions at the terahertz frequencies.

For successful designs reported in this work, the lower radiative state depopulation is achieved by a combination of resonant-tunneling and fast LO phonon scattering, which allow robust operation even at relatively high temperatures. An equally important enabling mechanism for these lasers is the development of metal-metal waveguides, which provide low waveguides losses, and strong mode confinement due to subwavelength mode localization in the vertical dimension. With these techniques some record performances for terahertz QCLs are demonstrated including the highest pulsed operating temperature of 169 K, the highest continuous-wave (cw) operating temperature of 117 K, and the highest optical power output (248 mW in pulsed and 138 mW in cw at 5 K) for any terahertz QCL.

Towards the bigger goal of realizing a 1-THz solid-state laser to ultimately bridge the gap between electronic and optical sources of electromagnetic radiation, QCLs with a unique one-well injection scheme, which minimizes intersubband absorption losses that occur at longer wavelengths, are developed. Based on this scheme a QCL operating at 1.59 THz ($\lambda \sim 189\ \mu$m) is realized, which is amongst the lowest frequency solid-state lasers that operate without the assistance of a magnetic field.

This thesis also reports on the development of distributed-feedback lasers in metal-metal waveguides to obtain single-mode operation, with greater output power and better beam quality. The subwavelength vertical dimension in these waveguides leads to a strongly coupled DFB action and a large reflection from the end-facets, and thus conventional coupled-mode theory is not directly applicable to the DFB design. A design technique with precise control of phase of reflection at the end-facets is developed with the aid of finite-element analysis, and with some additional unique design and fabrication methods, robust DFB operation has been obtained. Single-mode surface-emitting terahertz QCLs operating up to $\sim$ 150 K are demonstrated, with different grating devices spanning a range of approximately 0.35 THz around $\nu \sim 3$ THz using the same gain medium. A single-lobed far-field radiation pattern, higher output power due to surface-emission, and a relatively small degradation in temperature performance compared to the Fabry-Perot ridge lasers makes these DFB lasers well suited for practical applications that are being targeted by the terahertz quantum-cascade lasers.

Thesis Supervisor: Qing Hu
Title: Professor

Acknowledgments

It is my pleasure to thank the following people with whose support this work was made possible.

I would like to gratefully acknowledge Prof. Qing Hu for supervising this work. I have learned immensely from him as a teacher, and consider myself lucky to have learned the art of conducting research under his guidance. The foremost credit for my overall development as a researcher goes to him.

I also thank Prof. Terry Orlando for his academic advising, and Prof. Erich Ippen and Prof. Franz Kärtner for agreeing to serve on my thesis committee.

I wish to thank Benjamin Williams for being a wonderful colleague and guide during all these years. He trained me with many of the valuable skills that I have today. He has always been very approachable and has never hesitated to help me whenever the need was there. I consider myself lucky to have worked along with him and I hope to emulate his zeal towards approaching problems in research.

It is a pleasure to thank Hans Callebaut for being a friend and a fine colleague, and for providing good company during the often lengthy hours of laboratory work. Many a times he went out of his way to explain the nuances of many theoretical concepts during our discussions. I have always felt the most comfortable discussing my newest ideas with him.

I am grateful for the past and present students in this research group including Juan Montoya, Kostas Konistis, Stephen Kohen, Alan Lee, Qi Qin, and Allen Hsu for being a bunch of enthusiastic and motivated colleagues who have kept the intellectual environment in this group very stimulating. It has been a pleasure to work with them and to learn from them. I would specifically like to thank Stephen Kohen for giving me an introductory crash course to FEMLAB, and to Alan Lee and Qi Qin for all their help with many of the experiments.

This work would not have been possible without the excellent MBE grown semiconductor wafers provided to us by Dr. John Reno from the Sandia National Laboratories. Much of the progress and success of this project rested on quality of those

MBE wafers. For that, and for answering all my MBE related questions, I wish to extend my gratitude to Dr. Reno. I also would like to thank Dr. Zbig Wasilewski from the National Research Council, Canada for a small but fruitful research collaboration over the past year. My knowledge about MBE has been enhanced significantly through a generous offering of detailed explanations by Dr. Wasilewski.

Finally, I wish to acknowledge the prompt help and support from the excellent staff at the Research Laboratory of Electronics, the Microsystems Technology Laboratories, and the graduate office in the Department of Electrical Engineering and Computer Science of MIT during the course of this work.

I dedicate this work to my parents, teachers, and my brother

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and background

The terahertz frequency range of the electromagnetic spectrum (Fig. 1-1) has remained relatively unexplored as compared to the neighboring millimeter-wave and infrared spectral ranges, primarily due to the lack of convenient and efficient radiation sources. A multitude of applications for terahertz imaging and spectroscopy exist, be it in the fields of astronomy, remote sensing and monitoring of earth's atmosphere, biomedical imaging, detection of concealed weapons and drugs, end-point detection in plasma-etching processes, free-space optical wireless communications, non-invasive inspection of semiconductor wafers, and so-forth [5, 6, 7]. Interest in spectroscopy [145, 146] mainly arises from the fact that many molecules, for example, carbon-monoxide and carbon-dioxide, water, nitrogen, oxygen, to name a few, have strong characteristic rotational and vibrational absorption features in the terahertz, as opposed to the neighboring spectral regions. Also the dielectric resonances that occur due to various conformational and binding states in heavier molecules such as proteins and DNA could be probed in the terahertz region. Additionally, terahertz spectroscopy is relevant in condensed matter physics to study collective effects in materials such as superconductors, and charge density plasmas since the energy scales involved are of the order of 1 meV (1 THz $\sim$ 4 meV). For imaging [71, 116, 45, 189], terahertz radiation is specifically useful since many materials such as paper, plastics,

and ceramics, which are opaque at visible frequencies, are transmissive in the terahertz and microwave region. In comparison to microwaves, terahertz radiation provide a much better spatial resolution due to its shorter wavelength. When compared to imaging with high energy X-rays, imaging with terahertz radiation is non-invasive, and can provide much better contrast in terms of identification of different materials due to their different absorption and refraction indices in the terahertz. Towards this goal, it is worth mentioning that atmospheric water absorption in the terahertz is an important consideration for terahertz imaging over a distance. The frequency band of $1.3 - 1.5$ THz is very attractive for such an application because it offers the lowest atmospheric water absorption in the frequency range of $1 - 5$ THz.

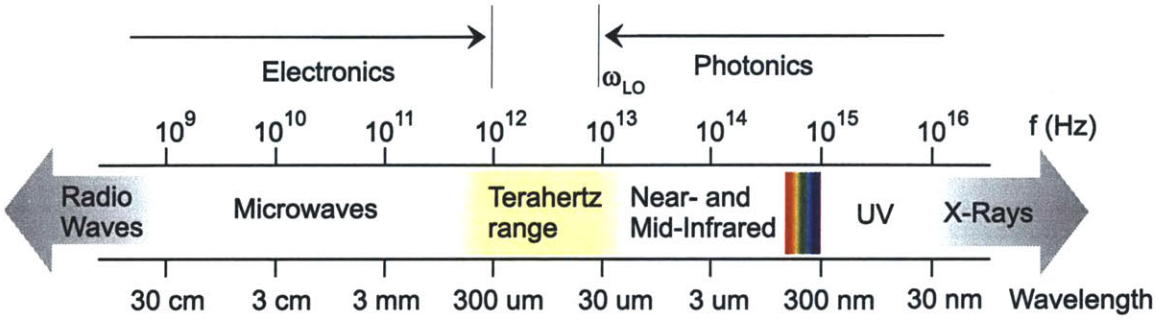


Figure 1-1: The "terahertz gap" in the electromagnetic spectrum.

Sensitive terahertz detectors such as liquid-helium cooled bolometers and photoconductors have existed since the 1960s as detailed in the review articles by Richards [133] and Haller [67]. The introduction of a micromachined monolithic silicon bolometer by Downey *et al.* in 1984 [46] sparked off a fascinating development in the room-temperature microbolometer focal-plane array technology for the infrared [136], which has now been shown to be useful for real-time imaging in the terahertz as well [99]. Even though terahertz detectors have long existed, there has been a dearth of coherent radiation sources in the terahertz. At the lower end of the terahertz spectrum, electronic sources such as transistors, Gunn oscillators, and Schottky-diode frequency multipliers can reach up to hundreds of gigahertz. The output power of these devices falls as $1/\nu^4$ or faster since their operation is dependent on transport of charge carriers, which is limited by their transit time and also by the parasitic capacitances in

both the device and the circuit. Consequently, output power levels obtained from current state of the art electronic devices operating near 1 THz is of the order of 10 μW [164, 115], which may be very small for practical applications. In contrast, for operation at much higher frequencies, photonic devices such as semiconductor laser diodes emit light due to quantum-mechanical electric-dipole oscillations, which are not limited by carrier transit-time effects. However, such devices can not emit light at frequencies below their electronic bandgap energy. While most such devices operate at near-infrared and optical frequencies, devices made from narrow bandgap materials such as lead-salts could go as low as $\sim$ 15 THz in frequency ($\lambda \sim 20$ μm) [153].

Thus, the scientifically important spectral range from $1-10$ THz is inaccessible to conventional semiconductor devices. Excluding the quantum- cascade lasers that are the subject of this thesis, the solid-state terahertz laser sources that currently exist are based on impurity-doped germanium or silicon [74]. In germanium, lasing could be obtained by hole population inversion between the light and heavy hole bands of an electrically pumped p-doped crystal when subjected to a crossed electric and magnetic field. Lasing has also been obtained between light and heavy hole impurity states of a strained p-Ge crystal without any magnetic field. In silicon, lasing originates from impurity state transitions of group-V donors in an optically-pumped n-doped crystal, where the population inversion is established either by charge accumulation in long-lived bound excited states of neutral donors or by resonant electron-phonon depopulation of certain impurity states. The aforementioned lasers are, however, of limited practical use because they need to be cooled to near liquid-helium temperatures for operation and only pulsed operation is possible due to their high power consumption and low efficiency. In comparison to lasers, broadband solid-state sources of terahertz radiation are more widely available. While such sources are inherently low power (average power < 1 μW), and may not be used as a local oscillator, they are useful nevertheless because of room temperature operation and for their ability to be detected coherently. Techniques such as generation of terahertz bandwidth time- domain pulses in high resistivity semiconductors [39], non- linear generation by

optical-rectification in crystals such as ZnTe [80], or non-linear generation by optical parametric conversion in materials such as $LiNbO_3$ [82], have been used for various raster-scanned imaging and broadband spectroscopic applications.

Until recently the only viable continuous-wave terahertz lasers for use in laboratory environment were the optically pumped molecular-gas lasers. These lasers achieve inversion between vibrational/rotational transitions of molecules that are excited by a pump laser (usually a CO_2 laser operating at $\lambda \sim 10$ μm), and are capable of room-temperature operation [76]. However, depending on the type of gas used, there is only a limited number of laser frequencies available. Moreover, these lasers are expensive, bulky, and power- hungry, which makes them less convenient for most applications. Despite these facts, these lasers have been used for important applications. For example, NASA's Aura satellite (formerly named EOS/CHEM-1, launched in 2004) that researches the chemistry and dynamics of Earth's atmosphere, including the very important ozone layer, carries a 2.5 THz methanol gas laser to monitor concentration of the OH molecules, which is a critical component in the ozone cycle [118, 1]. This laser emits 31 mW at 2.5 THz, measures $75 \times 30 \times 10$ cm, weighs 21 kg, and is designed to last for the mission lifetime of 5 years. As can be imagined, a compact solid-state terahertz laser would make an ideal replacement to the gas laser for such a mission. This is but one of the many applications that can benefit from such a device.

1.2 Quantum-cascade lasers (QCLs)

A quantum-cascade laser (QCL) is an electrically-pumped unipolar photonic device in which light emission takes place due to intersubband optical transitions in two-dimensional quantum-wells of a semiconductor heterostructure. This type of a laser, first demonstrated by Jérôme Faist and co-workers in Federico Capasso's group at Bell Labs [54], is often cited as the foremost example of "bandgap engineering". This is because energy-spacing between the subbands, and hence the frequency of light emission, can be artificially engineered by varying the size of the quantum wells. One key characteristic of QCLs is the use of periodic multiple quantum-well (MQW) modules

such that one charge carrier leads to generation of multiple photons as it is transported through the repeated modules. QCLs today have covered a wide wavelength range of 3 μm [142, 132, 42] to 190 μm [91] (and further up to 215 μm [138] with the aid of magnetic field operation), outperforming any other electrically-pumped semiconductor laser source in this wide range of electromagnetic spectrum. These lasers have so far primarily been realized in the conduction band of n-doped InGaAs/InAlAs, GaAs/AlGaAs, and InGaAs/AlAsSb heterostructures. The fact that lasing around $\lambda \sim 10$ μm has been obtained in all of these three material systems attests to the versatility of the QCL concept, which can be tailored to the specifics of the material, and also to the advancement in material-growth technologies today.

Before discussing the operation and development of QCLs, the next section briefly describes the key differences between intersubband and interband lasers to understand the operating principle of QCLs as compared to the semiconductor diode lasers.

1.2.1 Intersubband versus interband semiconductor lasers

Fig. 1-2 shows the difference between interband and intersubband optical transitions in two-dimensional quantum wells of a semiconductor heterostructure schematically. An interband radiative transition involves an electron-hole recombination, and hence an interband laser is inherently bipolar in nature. The frequency of radiation is essentially determined by the bandgap E_g of the quantum well material (of the order of 1 eV or $\lambda \sim 1.24$ μm or $\nu \sim 242$ THz), while small tuning can be obtained by changing the width of quantum wells. The joint density of states for the optical transition is constant, which is typical of two-dimensional density of states for a parabolic $E(k_{\|})$ dispersion. The gain spectrum is typically broad and reflects the thermal distribution of carriers in the conduction and valence bands.

In contrast, intersubband optical transitions are within the subbands of conduction band (or valence band) itself [165]. This has two advantages. First, the radiative frequency can be set by design as a function of the width of the quantum wells. Theoretically, emission for the radiative energy $\hbar\omega$ as high as the heterostructure band-offset ΔE_c (or ΔE_v for valence band), to as low as few meVs (and hence fre-

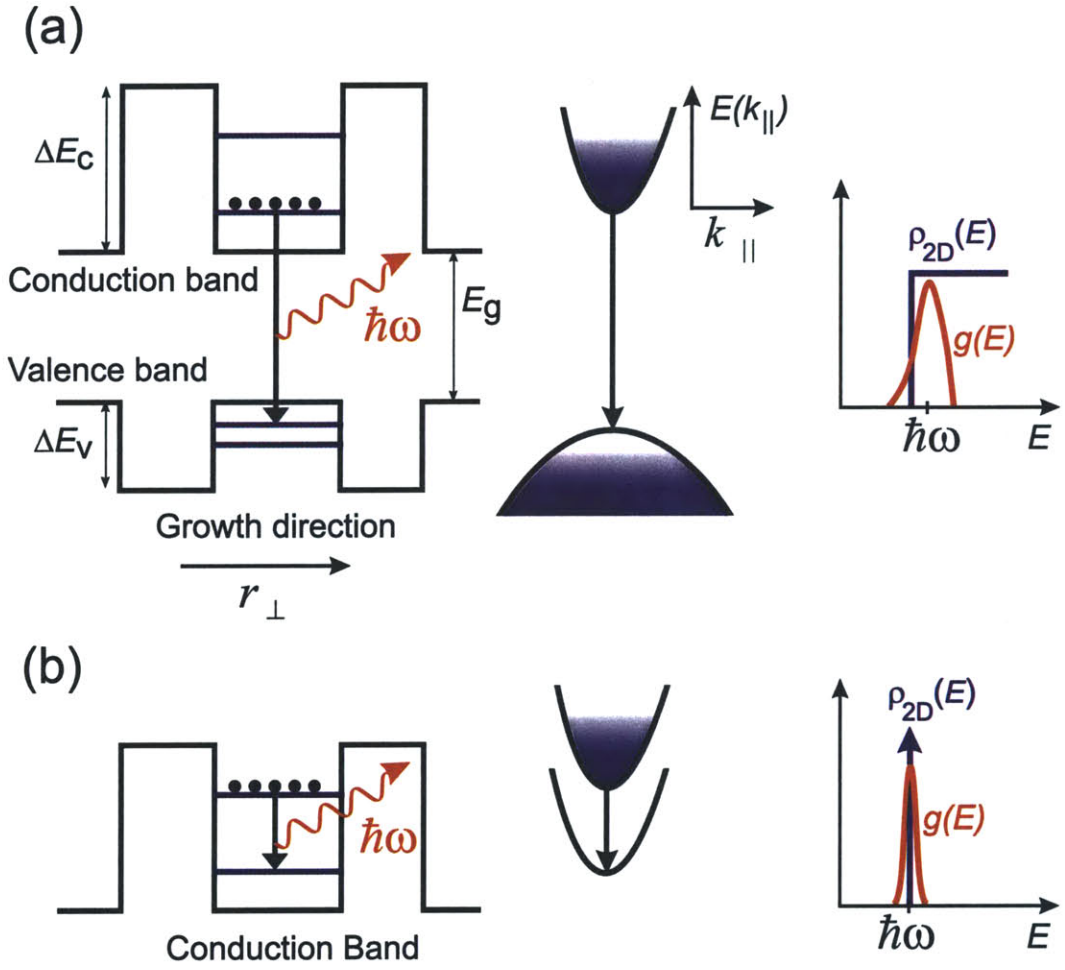


Figure 1-2: Schematic showing (a) interband, and (b) intersubband optical transitions in two- dimensional quantum wells. $\|$ and $\perp$ denote directions parallel and perpendicular to the layers of the quantum wells. $\rho(E)$ and $g(E)$ denote joint density of states, and gain, respectively, as a function of energy of optical transition.

quencies near 1 THz) should be possible as long as enough population inversion could be established between the desired subbands. Second, the energy dispersion of each of the subbands track in $k_{\|}$ space leading to a delta-function like joint density of states at the optical transition energy, $\rho(E) = \delta(E - \hbar\omega)$ where δ is the Dirac delta function, which is similar to that of atomic transitions. This suggests that all inverted carriers should contribute to gain at the same transition energy of $\hbar\omega$, indicating that the gain could potentially be very large (which is not surprising since optical gain $g(E) \propto \rho(E)$). However, this feature is offset by the fact that the non-radiative lifetimes for intersubband transitions are much faster (order of 1 ps) as compared to the non-radiative electron-hole recombination times in an interband diode-laser (order of 1 ns), mostly due to the fast polar longitudinal-optical (LO) phonon intersubband scattering mechanism in semiconductor heterostructures. Hence, small level lifetimes limit the amount of population inversion that could be achieved between the desired subbands. It is for this reason, intersubband lasers are typically made of multiple cascaded modules to obtain enough gain for lasing. The unipolar nature of intersubband transitions makes such a cascading possible.

Based on the aforementioned operating principles there are several characteristics that distinguish interband lasers to intersubband lasers. For an intersubband laser such as a QCL, despite having a delta-function like joint density of states, the gain spectrum has a finite width due to homogeneous (collisional) broadening as well as inhomogeneous broadening (due to fluctuations in material growth) of the subbands. Whereas for an interband laser the gain spectrum is additionally broadened due to a thermal distribution of the carriers. Typical full-width half-maximum (FWHM) spontaneous-emission linewidths observed for QCLs have been $2 - 20$ meV, as compared $\sim$ 50 meV or more for the interband quantum-well diode lasers. As shown schematically in Fig. 1-2, an intersubband laser will have a symmetric Lorentzian-shaped gain spectrum. Consequently, the Kramers-Kronig transformations dictate that contribution of intersubband gain, which determines the imaginary part of electrical susceptibility, to the real part of the refractive index should be asymmetric about the peak gain frequency, and hence zero at the center. The linewidth enhance-

ment factor in an intersubband laser should therefore be much smaller than that of an interband diode-laser (by at least an order of magnitude), which has indeed been verified experimentally in QCLs [10]. A small linewidth enhancement factor should allow an intersubband laser to maintain optical coherence over large device areas [159], which is important for development of high-power surface-emitting lasers, among other applications.

The small non-radiative lifetimes in an intersubband laser have a bearing on its high frequency modulation response. As will be shown in chapter 2, and has been demonstrated experimentally in a QCL [125], the relaxation oscillations that are typically observed in the modulation response for interband diode-lasers are absent in QCLs. The modulation bandwidth for QCLs, however, is limited by the photon lifetime in the cavity (on the order of 10 GHz) similar to that in interband diode-lasers since it is the longest of all other lifetimes that determine the modulation response. In that sense, an intersubband laser will offer no particular advantages over interband-diode lasers despite having non-radiative carrier relaxation mechanisms that are orders of magnitude faster. Multi-mode behavior is often observed in QCLs with Fabry-Pérot cavities and it is stipulated to be an effect of spatial-hole burning on the order of wavelength inside the semiconductor (λ_{semi}), which happens because the diffusion length of the carriers $\sqrt{D\tau}$ is of the order of 100 nm, which is $\ll \lambda_{\text{semi}}$ (where $D = \mu k_{\text{B}} T/e$ is the diffusion coefficient, $\mu = e\tau/m^*$ is the in-plane mobility, and $T \sim 300$ K, $\tau \sim 0.5$ ps, and GaAs effective mass are used for the estimate). Fast non-radiative relaxation mechanisms in intersubband lasers make their gain recovery time of the order of 1 ps, which is much smaller than the typical cavity round-trip time ($\sim$ 25 ps for a 1 mm long cavity in GaAs for example), thus making it hard to see mode-locking in QCLs [64]. Fast gain recovery makes it possible to observe coherent optical phenomena in QCLs. For example, Rabi-splitting of the gain spectrum in a mid-infrared QCL [64], and coherent field oscillations in the time-resolved stimulated emission of a terahertz QCL [90], have recently been reported.

Another distinguishing characteristic for intersubband optical transitions in MQW wells, is the requirement for the electric-field to be polarized perpendicular to the

layers (i.e. in the growth direction). This is because the Bloch wavefunctions for the electrons remain the same in an intersubband transition, and electric-field can couple only with the electric-dipole oscillations between envelope-wavefunctions of the subbands. Whereas, for interband lasers there is no such polarization selection rule since the dipole-matrix element is not limited to a particular direction. This is because the electric-dipole oscillations are caused by electromagnetic coupling of the electron and hole Bloch-wavefunctions in all the three dimensions.

1.2.2 Birth of QCLs: operation in the mid-infrared

The first idea of an intersubband laser was proposed as early as 1960 by Lax [96], even before the interband semiconductor diode laser was developed. The scheme involved optical pumping of Germanium in high-magnetic fields to achieve inversion between Landau subbands of either the conduction or valence bands to achieve a far-infrared "cyclotron resonance maser". The first such a laser was demonstrated by Ivanov *et al.* in 1984 [78]; however, the associated problems due to the requirement of high magnetic fields, requirement of low doping levels, small duty-cycle operation and other such issues has generated only a limited interest in such lasers [15]. On the other hand, with the invent of semiconductor heterostructures and the corresponding advancements in material growth, driven largely in part by the development of semiconductor diode lasers, people had started to work on semiconductor quantum-wells and superlattices in the late 1960s. In 1970, Esaki and Tsu [48] published a paper on the concepts of a superlattice and the possibility of appearance of a negative-differential resistance (NDR) in such a structure. One year later, Kazarinov and Suris [83] suggested, for the first time, the possibility of light amplification in a semiconductor superlattice *independent* of the NDR mechanism, based on a unipolar *photon-assisted* tunneling transition through a potential barrier.

Figure 1-3 shows the proposed idea schematically. An electrically pumped semiconductor superlattice consisting of multiple repeated quantum wells conducts current as the electrons are transported through a one-dimensional periodic potential of the heterostructure. The dominant intersubband scattering mechanism for the electrons

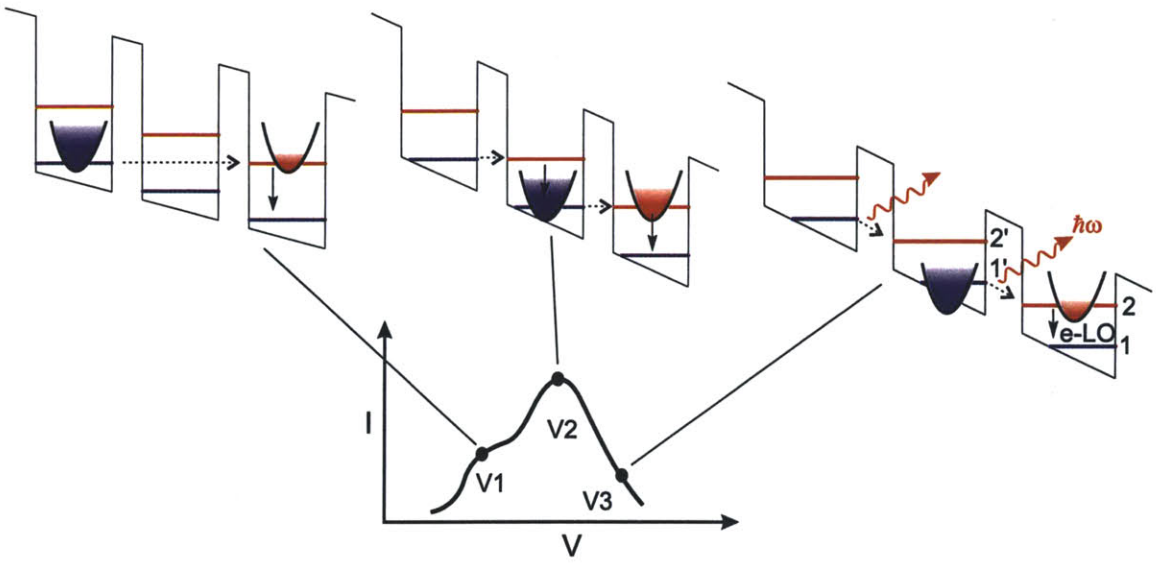


Figure 1-3: Schematic for demonstration of the concept of obtaining light amplification from a semiconductor superlattice based on the proposal of Kazarinov and Suris in 1971 [83]. A hypothetical I-V for such a device is shown along with the conduction band diagrams and location of the subbands in energy space at different bias conditions. Electron transport in the electrically-pumped structure is through the process of sequential resonant-tunneling.

is through the emission of LO-phonons to scatter from the excited state 2 to the ground state 1 within a quantum well. Interwell transport is by coherent resonant-tunneling through the barriers, where the tunneling probability is controlled by the barrier thickness and the static electric field across the structure. This combination of resonant- tunneling and intersubband scattering for electron transport in multiple quantum wells is often described as *sequential resonant-tunneling.* At certain bias conditions (such as $V1$ and $V2$), there is a peak in the current transport as the levels from different wells get aligned in energy space. The dominant current channel happens when the ground state of one well aligns with the excited state in the adjoining well (i.e. bias $V2$). Due to the fast non-radiative relaxation, the ground state is populated more than the excited state at all bias. The interesting situation happens beyond the bias $V2$, when there is enough interaction between the ground state and the excited state of the adjoining well to cause coherent electric-dipole oscillations diagonally across the barrier. Light amplification at $\hbar\omega = E_{1'2}$ can be obtained since a population inversion is maintained between levels $1'$ and 2 due to a short lifetime of

level 2 ($\ll 1$ ps for $E_{21} \gtrsim \hbar\omega_{\mathrm{LO}}$), and a long lifetime of level $1'$ ($\sim 1-10$ ps depending on the diagonality of the radiative transition). Moreover, the energy of the radiative transition $E_{1'2}$ is tunable with bias over a relatively wide range.

The scheme presented above sparked off intense interest to study electrical transport phenomena in semiconductor superlattices in the 1980s. However, the proposed idea has never been realized, and in its original form has some profound limitations for a practical implementation. The original idea proposed undoped superlattices, which meant that electrons were to be injected in the periodic structure from the contact regions at the either ends of the superlattice. This would cause formation of space charge domains across the structure, which prevent the development of a homogeneous static electric field across the structure, ultimately leading to electrical instability in the device. This problem, which is caused due to the unipolar nature of the device and the specific electrical-pumping scheme, can be solved by doping the superlattice within each of its period, such that the positive charges on the ionized donors compensate for the steady-state electron population within each quantum well. However, a yet another bigger problem inherent to the operating principle of such a device exists. The operating bias point for a photon- assisted tunneling transport mechanism at $V3$ is beyond the peak-current bias at $V2$, which is in a negative-differential resistance (NDR) region on the I-V characteristics. A NDR region is inherently unstable for an electrically-pumped superlattice that has multiple repeated modules, since the modules break down into high-field domains in a NDR region [24]. This is because of the lack of a mechanism that could strictly enforce a homogeneous field distribution across the structure due to complex non-linear electron dynamics of charge oscillations in the NDR. Hence, even though a population inversion is guaranteed at $V3$ bias, the superlattice could never be biased at that point. It may be noted that even though a single quantum well structure could be electrically biased in the NDR, a superlattice with multiple quantum wells is almost necessary to achieve sufficient intersubband gain for lasing (for reasons mentioned in section 1.2.1).

While resonant-tunneling through a double-barrier was observed as early as 1974 [36] attesting to the high quality of molecular beam epitaxy (MBE) already achieved

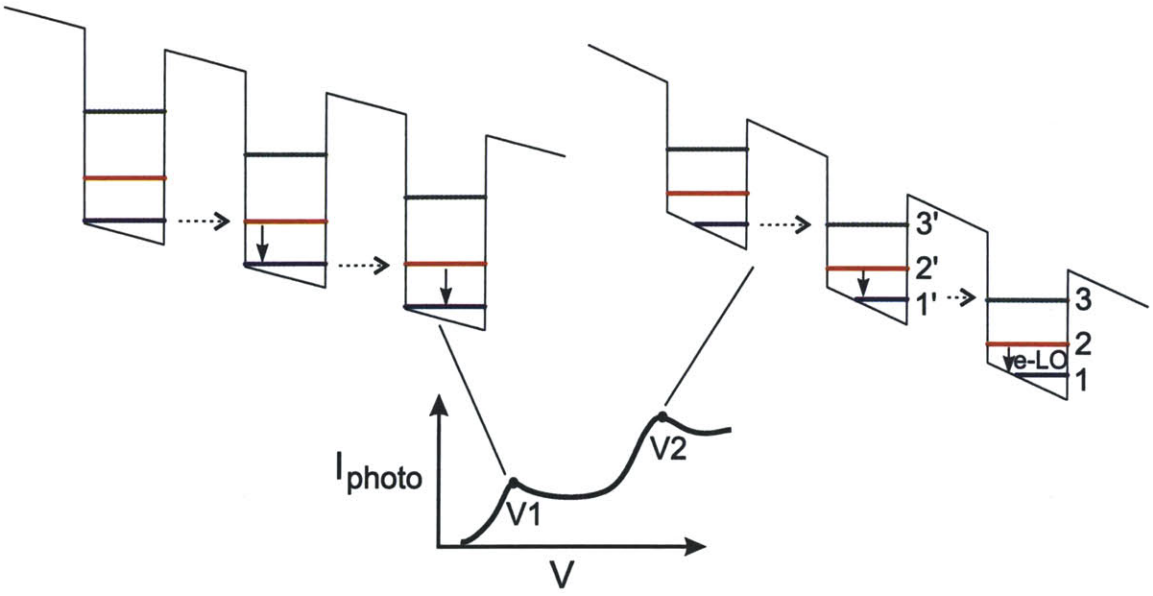


Figure 1-4: First observation of sequential resonant-tunneling in a semiconductor ($In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$) superlattice with 35 periods [33]. Conduction band schematics corresponding to two different bias points in the I-V are shown. The I-V shown here is a qualitative sketch of the measured I-V in Ref. [33]. Note that, at all bias, the ground state 1 in the quantum-wells will be the most populated due to fast electron-LO-phonon scattering from levels $3 \rightarrow 2, 1$ and $2 \rightarrow 1$.

at that time, it took several years before Capasso *et al.* were able to observe sequential resonant-tunneling through a semiconductor superlattice in 1985 [33]. Part of the reason for this was the inherent difficulty in electrically biasing a superlattice beyond the NDR regions. To ensure a rigorously controlled and spatially homogeneous electric field through the structure, the superlattice in Ref. [33] was placed in the i region of a reverse- biased $p^+ - i - n^+$ junction. Electrons, generated as minority carriers through photo-excitation in the $p^+ - i - n^+$ junction, were measured as photocurrent versus the applied reverse bias voltage. This scheme was critical to obtain the results of their experiment since it was able to prevent domain formation in the NDR regions of the I-V characteristics. Conduction band schematic of that structure and a qualitative sketch of the low-temperature I-V, similar to the one actually measured, is shown in Fig. 1-4. The peaks corresponding to the two dominant resonant-tunneling transitions were clearly resolved providing evidence of the high quality of material growth, and paving way for unique device applications based on resonant-tunneling

phenomena in semiconductor superlattices.

Multiple proposals to achieve intersubband lasing were considered in the late 1980s and early 1990s [52]. The structure of Capasso *et al.* based on the schematic in Fig. 1-4 merits further discussion since it bears the closest resemblance to the first superlattice based intersubband laser (i.e. the QCL) that was invented a few years later. In their subsequent paper [32], the authors considered the idea of achieving population inversion between levels 3 and 2 within the same well. At $1' - 3$ resonance (bias point $V2$) in Fig.1-4, the second excited state 3 in a well is selectively populated by the ground state $1'$ of the adjacent well through resonant-tunneling. If a population inversion could be established between levels 3 and 2, stimulated emission is possible through a vertical intrawell radiative transition. Hence, the resonant-tunneling through the barrier no longer needs to be coherent (the difference between coherent and incoherent resonant-tunneling through a barrier will be explained in chapter 2). It may be noted that such a scheme is significantly different compared to the original Kazarinov and Suris proposal outlined in Fig. 1-3, where the radiative transition is an interwell diagonal transition, and the resonant-tunneling through the barrier must be coherent. However, as Capasso *et al.* noted in their paper, it is difficult to obtain population inversion in such a scheme because of fast (sub-ps) electron-LO-phonon scattering times between levels $3 \rightarrow (2,1)$ and $2 \rightarrow 1$.

The electron-LO-phonon scattering matrix element for two dimensional subbands is proportional to $1/|\Delta\mathbf{k}|$, where $|\Delta\mathbf{k}|$ is the momentum exchanged in the scattering process [52]. If the subbands in the scheme of Fig. 1-4 are designed with an energy spacing such that $E_{32} \gg \hbar\omega_{\mathrm{LO}}$, the momentum exchange involved for an electron-LO-phonon scattering event from level $3 \rightarrow 2,1$ is an order of magnitude greater than it is from level $2 \rightarrow 1$ (where $E_{21} \approx \hbar\omega_{\mathrm{LO}}$). This is shown in Fig. 1-5 for parabolic energy dispersion in $In_{0.53}Ga_{0.47}As$ for $E_{32} \approx 300$ meV. Consequently, $\tau_{\mathrm{e-LO}}$ from level $3 \rightarrow 2,1$ ($\sim 2-5$ ps) is an order of magnitude larger than that from level $2 \rightarrow 1$ ($\sim 0.2 - 0.5$ ps), and a population inversion could be achieved between levels 3 and 2 at the $1' - 3$ resonance bias condition.

To obtain level spacings corresponding to a large E_{32}, Faist *et al.* [54] developed a

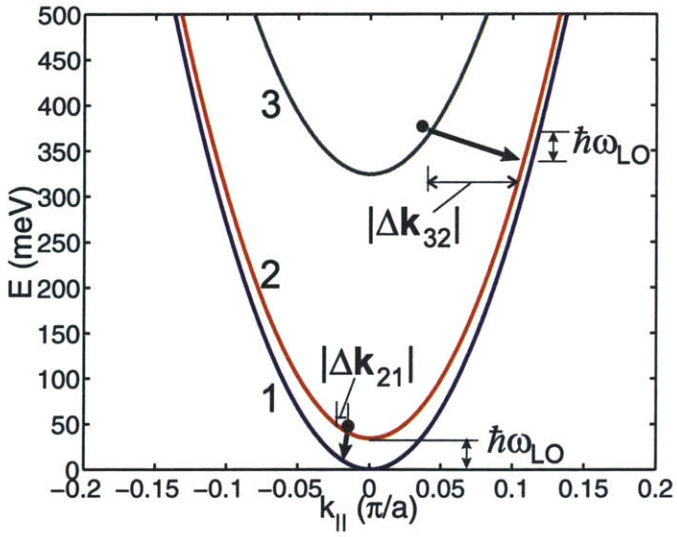


Figure 1-5: Magnitude of the momentum exchanged through intersubband electron-LO- phonon scattering events for level spacings such that $E_{32} \gg \hbar\omega_{\mathrm{LO}}$ and $E_{21} \approx \hbar\omega_{\mathrm{LO}}$. The $E(k_{\|})$ dispersion curves are plotted with $In_{0.53}Ga_{0.47}As$ parameters ($\hbar\omega_{\mathrm{LO}} \approx 34$ meV). A parabolic dispersion is assumed for simplicity (although non-parabolicity is not negligible). $k_{\|}$ is magnitude of the reciprocal lattice vector in the plane of the quantum-well layers.

structure in which levels 3, 2 and 1 are due to adjacent quantum wells that constitute part of a superlattice period, which is then repeated multiple times to form the superlattice. This allows control of level spacing by varying the corresponding well-widths to obtain $E_{32} \gg \hbar\omega_{\mathrm{LO}}$, as shown in Fig. 1-6. The barriers are kept thin to have enough spatial overlap in between the levels. Additionally, a digital-graded alloy, which is essentially a doped multiple quantum-well region, is implemented in the superlattice period. Such a region forms a miniband of levels in the doped region, which helps in extraction of electrons from lower levels 2 and 1 and inject them into the excited state 3 of the next period through resonant-tunneling. This doped region, termed as the *injector*, provides important additional functionality in this design. Firstly, it acts as a Bragg reflector at higher energies to prevent electrons from the excited state 3 to escape into the continuum over the barriers. This is because the injector region forms a minigap just over the injector miniband where no propagating states can exist [52]. Secondly, it limits current flow in the device at fields much lower than the design bias, since the multiple levels in the injector

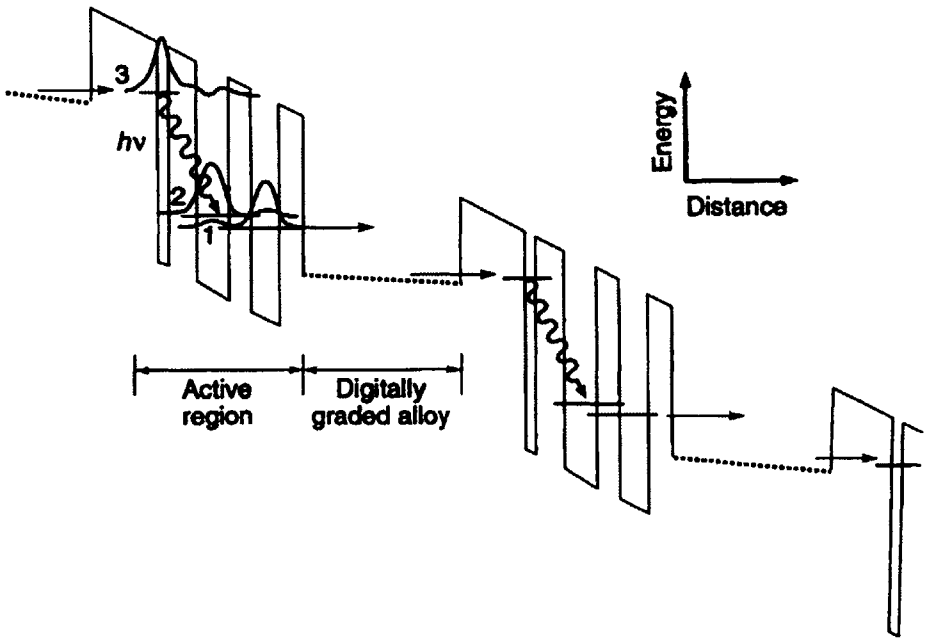


Figure 1-6: Conduction band diagram and magnitude squared wavefunctions for the first intersubband laser in a semiconductor ($In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$) superlattice with 25 periods. The device, which lased at a wavelength of 4.3 μm ($\nu \sim 70$ THz, $\hbar\omega \sim E_{32} \sim 290$ meV), was named as a "quantum cascade laser" by the authors in line with its operating principle. The figure is reproduced from Faist *et al.* [54].

are misaligned with respect to each other at lower bias and hence do not conduct current efficiently. Such a scheme (which is more critical for terahertz QCL designs as will be shown in chapter 3) suppresses any current conduction channels and the related resonant-tunneling peaks prior to the design bias. This prevents occurrence of a NDR region prior to design bias in the device I-V characteristics, and maximizes current flow in the device at design bias. Finally, the addition of the injector region makes the superlattice period longer, which reduces the amount of static electric field needed across the superlattice to obtain the desired alignment of levels at the design bias. This is critical to limit the escape time of the electrons from level 3 into the continuum, since level 3 is located close to continuum in the energy space.

With the aforementioned design, mid-infrared electroluminescence [53], and subsequently lasing [54] was obtained for the first time at a wavelength of 4.3 μm ($\nu \sim 70$ THz, $\hbar\omega \sim 290$ meV). Hence, the first semiconductor superlattice based

intersubband laser, appropriately named as the "quantum cascade laser", was developed and a new field was born. It is worthwhile to note that the radiative efficiency for spontaneous emission, $\eta \sim \tau_{e-LO}/\tau_{spont}$, for such an intersubband laser is very small ($\sim 10^{-5} - 10^{-4}$) due to the fast intersubband non-radiative scattering mechanisms. Consequently, significant population of upper level is needed to observe spontaneous emission in an intersubband device. It is then no surprise that the first QCL was demonstrated within few months of observation of electroluminescence in such a structure.

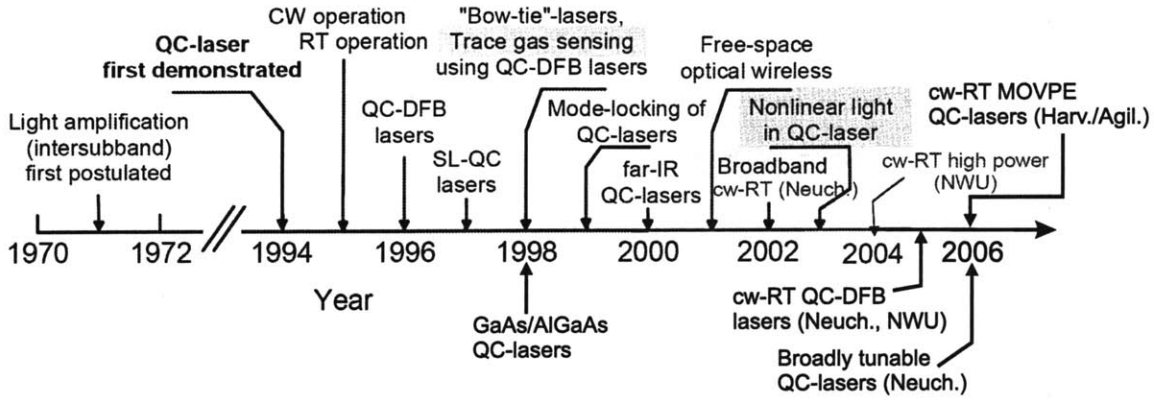


Figure 1-7: Time line of significant developments related to the mid-infrared QCLs. Figure adapted from Ref. [61]. Acronyms used: RT–room- temperature, CW–continuous-wave, DFB–distributed-feedback, SL–superlattice, IR–infrared. Note that "far-IR" in this schematic refers to a range of 20 μm $< \lambda <$ 25 μm. Developments related to terahertz QCLs ($\lambda > 30$ μm) are not included.

Since their inception in 1994, remarkable progress has been made in the development of mid-infrared (MIR) QCLs and the related applications. The reader is directed to a comprehensive review article on the progress of MIR-QCLs until 2001 by Gmachl *et al.* [61]. A time line of some of the significant developments related to MIR-QCLs until 2006 is shown in Fig. 1-7. Today, high-performance MIR-QCLs covering a range of $\lambda \sim 4-10$ μm with continuous-wave (cw) room-temperature operation, and optical power output of the order of 100 mW are available [50, 188]. QCLs that are broadly tunable ($\sim$ 25% of center frequency) by the use of external-cavity gratings have been demonstrated [112]. Until now, most high performance MIR-QCLs

were demonstrated in the InGaAs/InAlAs material system grown by molecular-beam epitaxy (MBE) (either strain-balanced or lattice matched to InP substrate). However, very recently structures grown by metal-organic vapor-phase epitaxy (MOVPE) have been demonstrated with performance comparable to the best results obtained by MBE material with both lattice-matched [43] and strained [44] crystal growth. This has created a large commercial potential for these lasers since MOVPE is a widely established platform for the high-volume production of reliable semiconductor-lasers. Finally, it is worth to mention that MIR-QCLs, which do not use the conventional injector minibands in their periodic structure, have been demonstrated with impressive temperature performance [57]. Within a few years of their introduction, such designs have already demonstrated record-low threshold current densities for any MIR-QCL at room-temperature [58], and may provide an alternative to the current state of the art bound-to-continuum two-phonon resonance designs that have injector minibands [21], similar to the original QCL design shown in Fig. 1-6.

1.2.3 Terahertz QCLs: challenges and first developments

The operating condition of a QCL can be conceptually described by the conduction band schematic in Fig. 1-4 at bias $V2$ (henceforth termed as the *design-bias*). This simple schematic serves as a good model-system to describe the operation of a QCL. In this model, the upper radiative level 3 is populated by resonant-tunneling and/or intersubband scattering from injector level(s) of the previous module. Whereas the lower radiative level 2 is depopulated by a combination of various intersubband scattering mechanisms into level 1 from where the electrons are recycled to complete the electrical transport process in a cascade scheme. In *real-space* the subbands could be designed to have different degrees of overlap, which affects the scattering cross-sections (form-factors) for various intersubband scattering mechanisms. However, in *k-space* the subbands overlap with each other since QCLs have so far been demonstrated in n-type direct band-gap semiconductors only, and all the electrons reside in the $\Gamma-$valley of the conduction band.

Most work related to the development of terahertz QCLs has been done almost

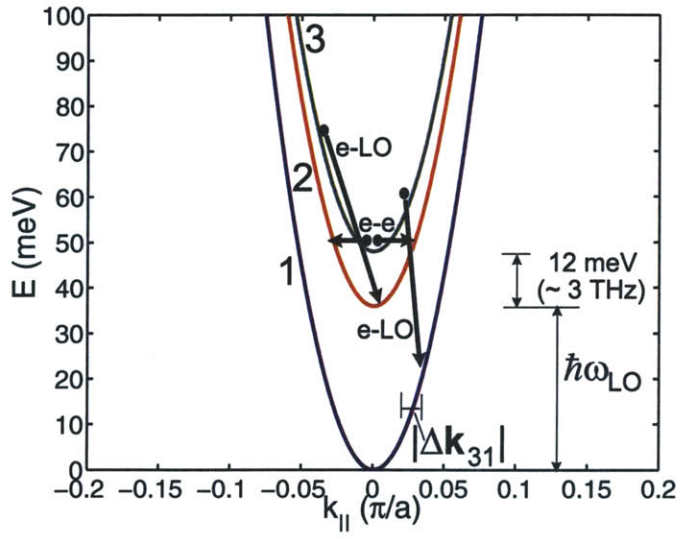


Figure 1-8: Some of the possible intersubband scattering mechanisms for terahertz QCLs. Levels 3 and 2 correspond to the upper and lower radiative levels, respectively, based on the schematic of Fig. 1-4. The $E(k_{\|})$ dispersion curves are plotted with GaAs parameters ($\hbar\omega_{\mathrm{LO}} \approx 36$ meV). A parabolic dispersion is assumed. See Fig. 1-5 for a comparison with MIR-QCLs. Note that E_{21} may be designed to be $< \hbar\omega_{\mathrm{LO}}$.

exclusively in the GaAs/AlGaAs material system. This is because of the high level of maturity attained in this system over years of developments related to interband-diode lasers. The small conduction band discontinuity (ΔE_{c} in Fig. 1-2, ~ 130 meV for $Al_{0.15}Ga_{0.85}As$) for GaAs/AlGaAs works well for terahertz QCL designs due to a relatively small energy separation required between the subbands (of the order of 10 meV). This is in contrast to mid-infrared QCLs where for shorter wavelengths ($\lambda \sim 5$ μm corresponds to $\hbar\omega \sim 250$ meV) the use of higher ΔE_{c} materials such as InGaAs/InAlAs becomes necessary.

Figure 1-8 shows the k-space diagram for the aforementioned three-level model with energy spacings between the subbands typically corresponding to that for terahertz QCLs. This diagram can be used to highlight the main challenges for designing the active region of a terahertz QCL in comparison to a MIR-QCL. Due to the very small energy spacing between radiative levels 3 and 2, various intersubband scattering mechanisms such as electron-electron, electron-impurity and electron- interface-roughness scattering are activated, and hence reduce the amount of population inversion that could be achieved. Moreover, a small E_{32} makes it hard to selectively

populate level 3 by means of resonant-tunneling from the previous module (i.e. the efficiency of injection into the upper level suffers). Additionally, selective depopulation of level 2 also becomes challenging. This is because the momentum exchange $|\Delta\mathbf{k}|$ involved for electron-LO-phonon scattering from level $3 \rightarrow 1$ is almost similar to that for level $2 \rightarrow 1$, unlike in mid-infrared QCLs, and hence this scattering process is also very efficient. The fact that it took seven more years for the invention of a terahertz QCL, after the first mid-infrared QCL was demonstrated, is a testimonial to the scale of difficulty involved. As will be shown below, terahertz QCLs typically got around these challenges by proper design of the overlap of the wavefunctions in real-space to minimize the $3 \rightarrow 1$ scattering form-factors, or by making $E_{21} < \hbar\omega_{\mathrm{LO}}$ to effectively suppress the electron-LO-phonon scattering mechanism, or by a combination of both.

Besides the challenges associated with design of the active region, terahertz waveguide design proved to be equally, if not more, challenging. This is corroborated by the fact that the very first observation of intersubband electroluminescence in a semiconductor superlattice was in the terahertz [68], which happened even before MIR-QCLs were invented. From the year 1997 to year 2000, several groups reported narrow-linewidth terahertz emission from quantum-cascade structures [183, 135, 182, 156, 177, 155]. The fact that intersubband emitters have inherently low radiative efficiencies ($\sim 10^{-5}$) suggests that a large population of carriers is required in the upper level before any spontaneous emission can be observed. This only makes it logical to conclude that high-loss in the waveguides, which were being used for terahertz mode confinement, was a major impediment to making a laser, even though population inversion might have been achieved much earlier.

While conventional dielectric waveguiding worked well for MIR-QCLs, that scheme can not be applied in the terahertz due to the long wavelengths involved (for example λ_{GaAs} at 3 THz is ~ 30 μm), which would require impractically thick dielectric layers. For this reason, the so called "plasmon waveguides", which are compatible with the TM polarization required for intersubband emission (i.e. electric-field perpendicular to the growth layers), were being used for these quantum-cascade structures. Fig. 1-9(a) shows the schematic for such a waveguide. Mode confinement in the MQW

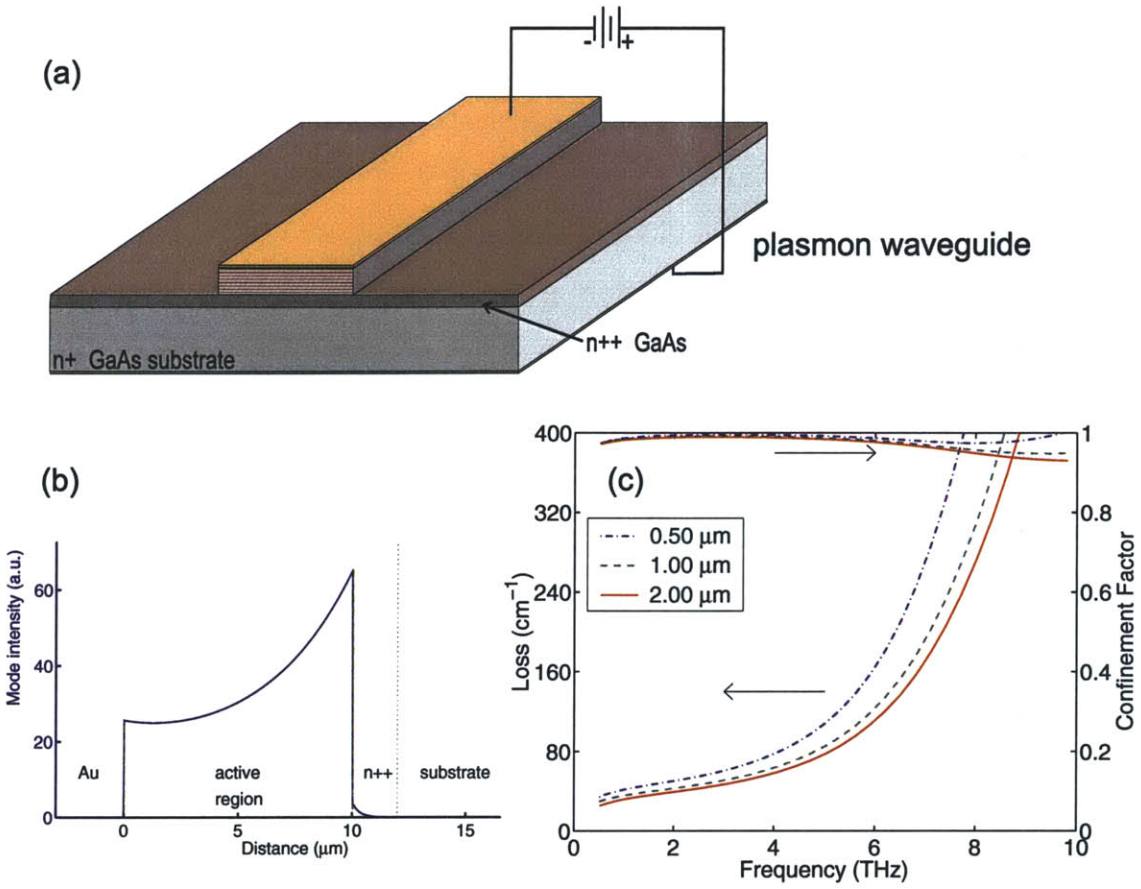


Figure 1-9: (a) Schematic of a plasmon waveguide. A 10 μm thick MQW (loss-less) active region is sandwiched between metal (Au in this figure) on one side and a thin n^{++} GaAs layer on a n^{+} doped substrate on the other. (b) Mode-intensity for the double-plasmon mode inside the waveguide at 5 THz. (c) Propagation loss (α_w) and mode-confinement (Γ) for various n^{++} layer thicknesses at different frequencies, computed within a Drude model. A Drude scattering time of 0.1 ps and doping densities of 5×10^{18} cm^{-3} and 1×10^{18} cm^{-3} are assumed for the n^{++} and n^{+} doped GaAs layers, respectively. This figure is reproduced from Williams [166].

active region is obtained by metal on the top, and a thin n^{++} doped GaAs layer on a n^{+} GaAs substrate at the bottom. At terahertz frequencies, real part of the dielectric constant for both the metal and the highly doped semiconductor is negative. Hence, surface-plasmon propagating modes that are bound to the top metal and the bottom doped semiconductor layer can exist (more details on this in chapter 4). For such a mode, the mode intensity is relatively uniform in the active region because its thickness ($\sim$ 10 μm) is sub- wavelength, as seen from Fig. 1-9(b). However, the mode is heavily attenuated in both the metal and the bottom semiconductor due to plasma- screening. While the plasma frequency in the metal is in the ultra-violet ($>$ 1000 THz), for the highly doped GaAs layer it is $\sim$ 20 THz ($\propto \sqrt{n}$, where n is the doping-density), which makes it very lossy for the terahertz mode due the mode frequency being very close to the plasma-resonance. Consequently, propagation loss for the waveguided terahertz mode is $\sim$ 50 cm^{-1} as shown in Fig. 1-9(c), which is many times higher than the cladding losses in mid-infrared waveguides ($\lesssim$ 10 cm^{-1}). The limits to the lowest loss that could be obtained from this type of waveguide is set by how high the n^{++} GaAs layer could be doped.

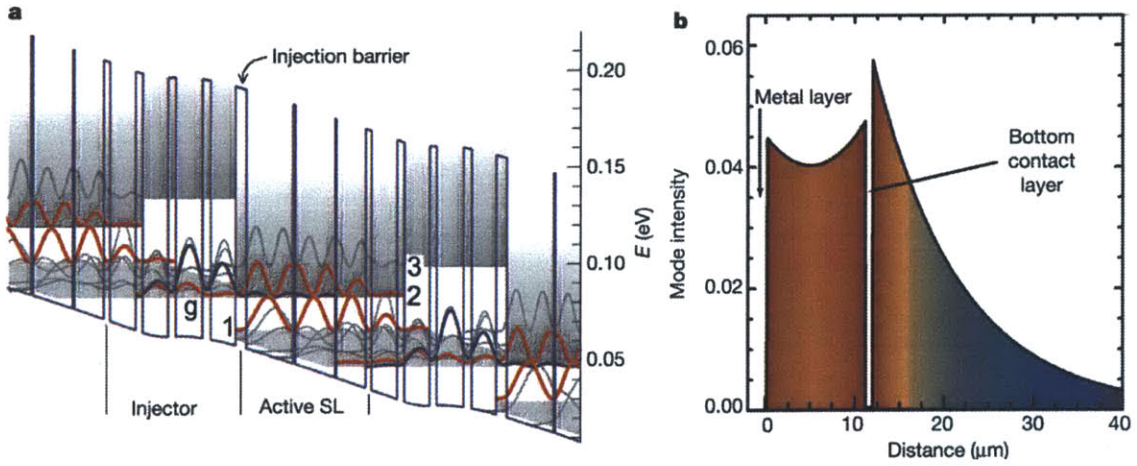


Figure 1-10: (a) Conduction band diagram (for two- modules) at design bias for the first terahertz QCL, which was based on a chirped-superlattice design. (b) Mode shape for the semi-insulating surface plasmon waveguide used for mode-confinement. The figure is reproduced from Köhler *et al.* [89].

After several years of trial by different research groups who had observed terahertz electroluminescence in quantum-cascade structures, the first terahertz QCL was fi-

nally invented in October 2001 by Köhler *et al.* in Alessandro Tredicucci's group at Pisa [89]. Their technique involved the combination of a "chirped-superlattice" based active region first demonstrated in the mid-infrared by Tredicucci *et al.* [154], and a "semi- insulating surface-plasmon" (SISP) waveguide first demonstrated for a different (non-lasing) terahertz quantum-cascade structure by Ulrich *et al.* [155]. This laser operated at 4.4 THz ($\lambda = 67$ μm), and pulsed operation was observed up to a heat-sink temperature of 50 K. The conduction band diagram and waveguided mode-shape for the QCL are shown in Fig. 1-10. Optical transition occurs between levels 3 and 2, which lie at the edge of the upper, and the lower minibands within a QCL period, respectively. This type of design has a high current-carrying capability, since both the minibands provide an efficient conduction mechanism by means of resonant-tunneling and intraminiband scattering. Moreover, a large oscillator strength is obtained because of a vertical optical transition and a greater spread of the wavefunctions in real-space. The lifetime of the lower level is kept ultrashort ($\ll 1$ ps) by intraminiband electron-electron scattering in the lower miniband, because of a large phase space that is available for the electrons to scatter into, and this forms the main mechanism to achieve population inversion. The SISP waveguide is obtained by growing the QCL on a semi-insulating (SI) GaAs substrate (as opposed to a doped one for a plasmon waveguide), and by making the bottom n^{++} GaAs contact layer much thinner (0.8 μm). As can be seen from Fig. 1-10(b), the mode supported in such a waveguide leaks out substantially into the substrate; however, its overlap with the doped GaAs regions is much smaller compared to the scheme of Fig. 1-9, which drastically reduces the waveguide loss α_w (~ 5 cm^{-1}). However, the waveguide may seem unintuitive since the mode confinement factor Γ also reduces considerably ($0.2 - 0.5$). The laser gain threshold is determined by $(\alpha_w + \alpha_m)/\Gamma$, where α_m is the mirror loss. Assuming a negligible α_m/Γ for long-cavity lengths with high-reflectivity coated back-facets (< 5 cm^{-1}), the α_w/Γ factor for the SISP waveguide (< 20 cm^{-1}) stands to be less by a factor of $3 - 5$ over that for the plasmon waveguide of Fig. 1-9 (> 50 cm^{-1}) at terahertz frequencies, which was a determining factor towards the realization of this laser.

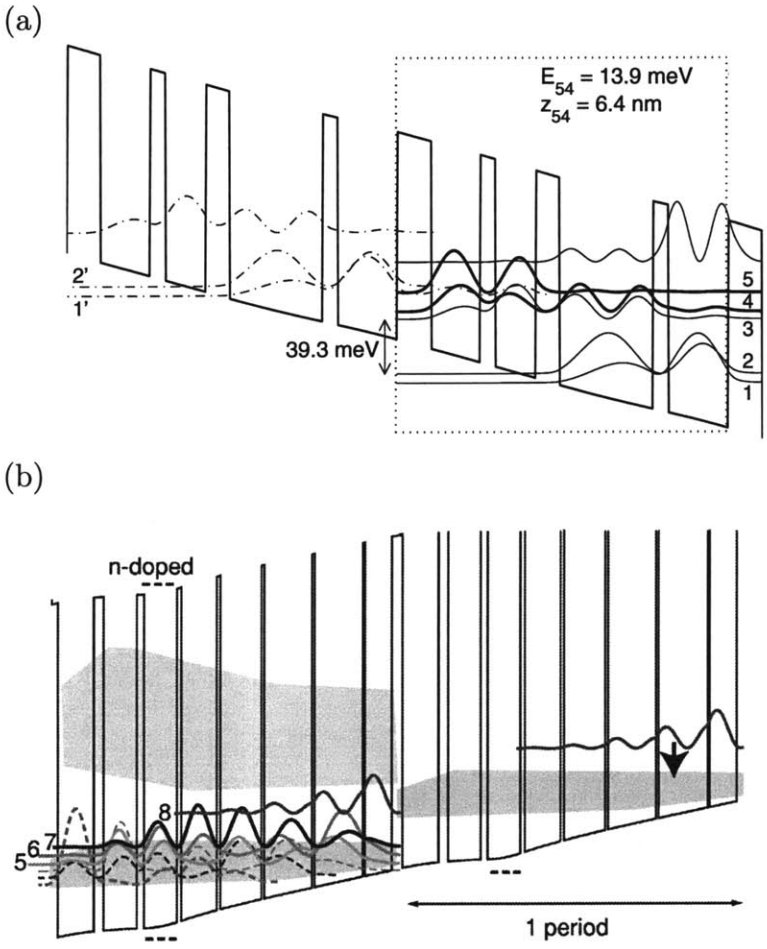


Figure 1-11: (a) Conduction band diagram for the first resonant-phonon terahertz QCL, reproduced from Williams *et al.* [168]. The static electric- field across the modules is from right to left. (b) Conduction band diagram for the first bound-to-continuum terahertz QCL, reproduced from Scalari *et al.* [137]. The electric-field for this diagram is from left to right. In each of the figures, two adjacent QCL modules are shown.

Once the SISP waveguide was proven effective for terahertz QCLs, two other groups published distinctly different working designs in quick succession with the same type of waveguide. This only confirms the importance (and challenge) of obtaining a low-loss terahertz waveguide, which could well have been the last hurdle to achieve terahertz lasing in QCLs for several years. One of those two designs was demonstrated by Williams *et al.* in Qing Hu's group at MIT in November 2002 [168]. This laser operated at 3.4 THz ($\lambda = 87$ μm), and pulsed operation was observed up to a heat-sink temperature of 65 K. The design was based on a "resonant-phonon" depopulation scheme, and is shown in Fig. 1-11. The depopulation scheme employed for this design is significantly different from the miniband based extraction scheme of Fig. 1-10. Lasing transition in this design takes place between levels 5 and 4. The lower level 4 is depopulated by a combination of resonant-tunneling into level 3, and then fast electron-LO-phonon scattering down into the injector levels 1 and 2. From the injector levels, which act as an electron reservoir, electrons are resonant-tunneled back into the upper level 5 of next module. The key feature of this design is fast selective-depopulation of the lower level, while maintaining a large upper-state lifetime. The other significantly different design, which was based on a "bound-to-continuum" transition [51], was demonstrated by Scalari *et al.* in Jérôme Faist's group at Neuchâtel in early 2003 [137]. This laser operated at 3.5 THz ($\lambda = 87$ μm), and pulsed operation was observed up to a heat-sink temperature of 90 K. This design is characterized by an isolated upper laser level created in the minigap formed by a superlattice within the QCL period, while the lower level lies at the edge of the first miniband like in the chirped-superlattice design. The radiative transition is made slightly diagonal, such that the coupling of injector levels to the lower miniband is reduced, thus obtaining high injection efficiency into the upper level.

At this point, the reader is directed to the cited publications to learn more about the chirped-superlattice and bound-to-continuum based designs. The brief introduction in this this section was meant to lay the groundwork for the discussions to follow.

1.3 Overview

The demonstration of the first resonant-phonon terahertz QCL by Williams *et al.* in 2002 [168] was a culmination of several years of investigation by Jürgen Smet [151], Bin Xu [182] and Benjamin S. Williams [166] in Qing Hu's group at MIT. While investigating terahertz electroluminescence from intersubband emitters, Bin Xu [182] developed the so called "metal-metal" waveguides for their ability to provide lower loss waveguiding in the terahertz as compared to plasmon waveguides. These waveguides were later demonstrated, for the first time, to work effectively for terahertz QCLs by Williams *et al.* [170]. These initial developments of both the active region and waveguides for terahertz QCLs form the foundation for the work done in this thesis.

A theoretical model for transport behavior of these QCLs has been developed by Hans Callebaut [28] in parallel during the course of this work. However, due to the complexity of the transport processes involved that are unique to QCL operation in the terahertz (as opposed to mid-infrared), the results from those models cannot yet be taken at face value, although they provide valuable insights. This work combines detailed experimental analyses of multiple types of active region designs (both lasing and non-lasing) with the theoretical insights provided by the models of Hans Callebaut, and also from work done elsewhere.

It is interesting to note that the first resonant-phonon design (Fig. 1-11a) offers far richer physics than what it may appear intuitively, and continues to be developed for yet better performance. Also, the metal-metal waveguides have since been much improved by the use of better fabrication techniques. Presently, some of the record performances for terahertz QCLs have been demonstrated with variations of the first resonant-phonon design in combination of the lower-loss metal-metal waveguides, during the course of this work. These results include the highest pulsed operating temperature of 169 K, the highest cw operating temperature of 117 K [174], the highest optical power output (248 mW in pulsed and 138 mW in cw at 5 K) [175], and the second lowest frequency of operation at 1.59 THz ($\lambda = 188.5\ \mu$m) without the assistance of a magnetic field [91] (the lowest frequency of operation has very recently

been obtained by a different research group, although the exact details are currently unavailable). Several important experiments aimed towards developing these lasers for practical applications have been demonstrated through collaborative work with other research groups. Some of these include frequency and phase-locking of a terahertz QCL to an external far-infrared gas laser for its precise frequency control [22], and demonstration of a heterodyne receiver based on a terahertz QCL as a local-oscillator [59]. Additionally, some applications have been demonstrated in-house, such as real-time imaging over a distance of 25 m with a terahertz QCL as the light source [100].

The key difference between the resonant-phonon depopulation scheme of this design, and the miniband-extraction schemes of chirped-superlattice and bound-to-continuum designs, is that the upper laser level for this design is spatially separated from the injector states. Hence, according to the k-space diagram of Fig. 1-8, while the electron-LO-phonon scattering mechanism from $3 \rightarrow 1$ becomes active in the miniband based designs as carriers become hot (where levels 3 and 2 indicate the laser levels and 1 the injector level), this mechanism is relatively suppressed with the use of the resonant-phonon technique due to real-space delocalization. Additionally, a large energy separation between the lower laser level and the injector levels ($\sim \hbar\omega_{\mathrm{LO}}$) in the resonant-phonon designs provides a thermal barrier against backfilling of the lower laser level by hot electrons from the injector levels. Hence, a better temperature performance is achieved with the resonant-phonon depopulation scheme, despite the fact that the oscillator strength for its radiative transition is about a factor of 2 smaller (= 0.85 for the design in Fig. 1-11) than that for bound-to-continuum and chirped-superlattice based designs (= 1.95 for the designs in Figs. 1-10 and 1-11).

This thesis is organized as follows. Chapter 2 establishes a theoretical formulation for treating intersubband transitions in quantum wells. Simplistic models to estimate of electrical transport and optical gain spectrum in terahertz QCLs are developed. These models can be very useful for analyzing experimental results, especially since accurate computational models for electrical transport calculations in terahertz QCLs are not yet available. The non-radiative scattering lifetimes are taken as parameters;

theoretical investigations for estimating non-radiative lifetimes have been done by Smet [151], and Callebaut [28]. Chapter 3 describes in detail the various different types of terahertz QCL designs that were investigated during the course of this work. Several designs are discussed to address the two primary goals for further development of terahertz QCLs: high temperature operation, and long wavelength operation.

For higher temperature operation, the main challenge for terahertz QCLs stems from the fact that as the carriers become hot, electron-LO-phonon scattering process is activated from the upper laser level to the lower laser level, thereby reducing the population inversion drastically. This is indicated by the $3 \rightarrow 2$ process in Fig. 1-8, which will effectively reduce τ_3 to a value closer to τ_2 once electrons in the upper laser level 3 gain a thermal energy equivalent to $\hbar\omega_{\mathrm{LO}} - E_{32}$. Additionally, with a design having such closely spaced subbands, it becomes challenging to suppress various other parasitic current channels in the structure that arise due to tunneling from the injector levels into some undesired levels (this will be explained in detail in chapter 3). Such parasitic couplings become enhanced at higher temperatures, thereby gradually pinching off the lasing dynamic range with temperature. From a practical perspective, the first big step of obtaining cw operation above liquid-nitrogen temperature was demonstrated during this thesis work in the year 2004 [94]. However, a major goal is to achieve terahertz QCL operating above $\sim$ 250 K, a temperature range which is accessible by thermoelectric coolers. Feasibility arguments and some possible design ideas for achieving that goal are presented in chapter 3.

The second primary goal for terahertz QCLs is to achieve lasing at a frequency as low as 1 THz ($\lambda \sim 300$ μm, $\hbar\omega \sim 4$ meV), to finally bridge the "terahertz gap" that exists between the electronic and optical sources of electromagnetic radiation. Besides a matter of scientific interest, the longer wavelength lasers have some specific applications, as discussed in section 1.1. The primary challenge for QCL operation at such a low frequency will be to maintain selective injection of electrons into the upper laser level, and selective depopulation of the lower laser level, since the collisional broadening of the subbands is of the order of radiative energy ($\sim$ few meVs) itself. Towards this goal, a terahertz QCL operating at 1.59 THz ($\lambda \sim 190$ μm) was recently

demonstrated as part of this work [91]. The design is based on a unique one-well injector scheme [92], which significantly cuts down photon absorption-loss in the active region, and provides improved injection selectivity into the upper laser level. Chapter 3 discusses the one-well injector design in detail, whereby it will become apparent that operation at even lower frequencies is possible with the same design.

Chapter 4 forms the second part of this thesis which is focused on waveguide design for terahertz QCLs. The best temperature performance for terahertz QCLs has been demonstrated in metal-metal waveguides since they provide lower propagation loss in the terahertz, as compared to the initially demonstrated semi-insulating surface-plasmon waveguides. Additionally, they provide almost unity mode confinement, even for very narrow (< 50 μm wide) ridges. Chapter 4 discusses the electromagnetic properties of metal-metal waveguides and discusses various fabrication techniques that were developed to improve their performance significantly since the first-demonstration. The improved thermal properties that are achieved due to sub-wavelength width ridges, and high thermal conductivity of the metal-metal bonding layers, has lead to the demonstration of record high cw temperature of operation in such waveguides [174].

Due to sub-wavelength mode confinement in the vertical dimension, metal-metal waveguides provide very tight mode confinement (or in other words high cavity quality-factors or high reflectivity from the boundaries of the cavity). But as a consequence, the radiation properties and light out-coupling efficiency for these waveguides are poor. The operation of these waveguides is very similar to that of microstrip transmission lines at the microwave frequencies, albeit with one big difference. While in the microwave waveguides, the radiation properties can be improved by incorporating mode-matching structures such as a horn antennas, that is difficult to do in the terahertz because of much shorter wavelengths involved. Towards this goal, surface-emitting distributed feedback (DFB) terahertz QCLs using metal-metal waveguides have been developed during the course of this work [95]. These DFBs have demonstrated robust single mode operation with maximum temperatures within few Kelvins of multi-mode Fabry-Pérot lasers. They additionally provide a single-lobed beam pat-

tern and relatively higher amount of optical power. Implementation of such DFBs in terahertz metal-metal waveguides had proved to be much more challenging than initially expected because of the unique properties of the metal-metal waveguides in the terahertz. Hence, several new techniques had to be developed both for DFB design and fabrication, to obtain successful operation. Extensive use of finite-element simulations was made during the design process. Chapter 4 describes the DFB designs, and the corresponding experimental results in detail.

Chapter 2

Intersubband transitions and transport in multiple quantum wells

This chapter describes some of the relevant theoretical aspects that are key to understanding the operation of a terahertz quantum-cascade laser (QCL). Some of the concepts described here were routinely employed for design of the QCL active regions as presented in chapter 3. The analyses presented here apply well to the electronic intersubband transitions in the conduction band of multiple-quantum-well (MQW) semiconductor heterostructures. All interband transitions are neglected and only unipolar transport is assumed. Device operation is considered only in the steady state. The electronic wavefunctions that are used in the analyses can be computed within a slowly-varying envelope approximation in the MQW growth direction [27]. Within this approximation, the effective mass m^* is the only parameter that describes the energy dispersion of the conduction band, albeit with a different value for the quantum-wells, and the barriers, respectively. Two underlying assumptions are made, which affect the electron energy dispersion in the plane of the MQW layers. First, the conduction band non-parabolicity [121, 147] is neglected. This is a good assumption since the energy levels of the wavefunctions in a terahertz QCL are located within an order of 10 meV from the conduction band edge, which is much

less than the band-gap energy of the semiconductor. Second, coupling of electronic wavefunctions in the plane and perpendicular to the plane of the MQW layers, which happens due to different effective masses in the wells and the barriers and conduction and valence band mixing effects [20, 184], is neglected. This is particularly valid for terahertz QCLs (which may be grown in $GaAs/Al_{0.15}Ga_{0.85}As$, for example) since the effective mass in the barriers is not very different from that in the wells, and the subbands are located close to the bottom of the quantum-wells in energy. Based on these assumptions, electrons can be treated as two-dimensional (2D) free particles in the plane of the MQW layers (i.e. with a parabolic $E(k_{\|})$ dispersion, where $k_{\|}$ is the in-plane momentum), with a mass equal to the conduction band effective mass of the semiconductor. These assumptions, as outlined in a greater detail by Williams [166], afford a simple calculation of the electronic wavefunctions using a 1D Schrödinger equation within the envelope function approximation. In contrast, for mid-infrared QCLs, non-parabolicity must be taken into account due to subbands that are located higher up in energy in the conduction band. This is usually done by making the effective mass energy dependent within the envelope function approximation [52]. Once the envelope wavefunctions and the corresponding energies for each of the subbands are determined, most of the results derived here apply equally well to even mid-infrared QCLs.

Even though the methods mentioned above seem plausible for calculation of stationary state electronic wavefunctions in MQWs, the time evolution of such wavefunctions in a real device, which eventually determines the "steady-state" transport behavior in the device, is highly complex, and more so for terahertz QCLs. This is because of the presence of ultra-fast (order of 1 ps) scattering mechanisms that determine electronic transport, such as electron-electron, electron-impurity, electron-interface-roughness, and electron-LO-phonon scattering. Especially for terahertz QCLs, which have small energy separation between the subbands (of the order of 10 meV), scattering rates are greatly dependent on the in-plane kinetic energy distribution of the electrons in the subbands (which is also of the order of 10 meV). Consequently, it becomes hard to calculate the subband electron populations and the current flowing

through the device in the steady-state, the two key parameters that determine the operating characteristics of a laser. Additionally, the decoherence of electronic wave-functions caused by these scattering mechanisms is extremely challenging to model, which makes it hard to accurately determine the tunnel-coupling between various sub-bands. These tunneling mechanisms, be it resonant-tunneling for electronic transport between two levels in the neighboring modules of the quantum-cascade structure, or the oscillator-strength of a photon-assisted optical tunneling transition between two radiative levels, are extremely important for determining QCL parameters. While a detailed analysis of electronic transport in terahertz QCLs was undertaken by Callebaut [28], the purpose of this chapter is to provide tools for the comparison and evaluation of one QCL design with respect to another, which can provide valuable insights when used in conjunction with experimental data.

2.1 Optical gain due to intersubband radiative transitions

2.1.1 The light-matter interaction

This section provides a review for understanding the light-matter interaction. The *Electric-Dipole Hamiltonian*, which couples two quantum-mechanical electronic states to cause radiative transitions, is derived. The analysis presented here is standard and is partially adapted from Chuang [40], and Williams [166].

In an isotropic homogeneous medium, with spatially invariant permittivity ϵ and permeability μ, the Maxwell's equations could be written in terms of the vector and scalar potentials, $\mathbf{A}(\mathbf{r},t)$ and $\phi(\mathbf{r},t)$, respectively, as

$$\nabla^2\phi + \frac{\partial}{\partial t}(\nabla\cdot\mathbf{A}) = -\frac{\rho}{\epsilon} \tag{2.1}$$

$$\left(\nabla^2\mathbf{A} - \mu\epsilon\frac{\partial^2}{\partial t^2}\mathbf{A}\right) - \nabla\left(\nabla\cdot\mathbf{A} + \mu\epsilon\frac{\partial}{\partial t}\phi\right) = -\mu\mathbf{J} \tag{2.2}$$

where $\rho(\mathbf{r},t)$ and $\mathbf{J}(\mathbf{r},t)$ are charge and current densities, respectively, produced by

all charges other than those associated with the polarization represented in ϵ, and satisfy the continuity equation

$$\nabla \cdot \mathbf{J} + \frac{\partial}{\partial t}\rho = 0 \tag{2.3}$$

The electric and magnetic fields are determined from $\mathbf{A}$ and ϕ by

$$\mathbf{E} = -\nabla\phi - \frac{\partial}{\partial t}\mathbf{A} \tag{2.4}$$

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{2.5}$$

$\mathbf{A}$ and ϕ in the equations above are not uniquely specified. Within gauge transformations ($\mathbf{A} \rightarrow \mathbf{A} + \nabla\xi$ and $\phi \rightarrow \phi - \frac{\partial}{\partial t}\xi$, where $\xi(\mathbf{r}, t)$ is an arbitrary scalar function), the physically measurable quantities such as $\mathbf{E}$ and $\mathbf{B}$ remain the same.

A terahertz QCL conduction band diagram could be represented by an example design shown in Fig. 2-1. Within the scheme of Fig. 1-4, this is the simplest possible design with 3 levels per repeated module, since at least two quantum-wells are needed to independently set the energy separation between the levels 3, 2, and 1, respectively. (In fact, such a design could possibly be made to work successfully as a QCL, but for reasons that will become apparent in chapter 3, it may not make a good laser.) As was alluded to in section 1.2.3, such a prototypical structure is a good model-system to describe the working principles behind a QCL. A terahertz QCL typically consists of multiple (> 100) repeated modules. Hence, unlike in interband diode lasers, it is not desirable to inject carriers from the contact regions. This is because, in steady-state operation, each module will have finite electron density in various subbands that will induce additional non-uniform static electric-field across the superlattice. This will cause distortion to the linear conduction band diagram of Fig. 2-1, and hence to the subbands and their corresponding energies. In general, this problem is avoided by n-doping each of the modules, either partially or uniformly across their entire length. However, a space-charge density ρ is still established since the location of the dopants is not exactly where the electrons might be. In general, $\rho(z)$ is kept small to keep the

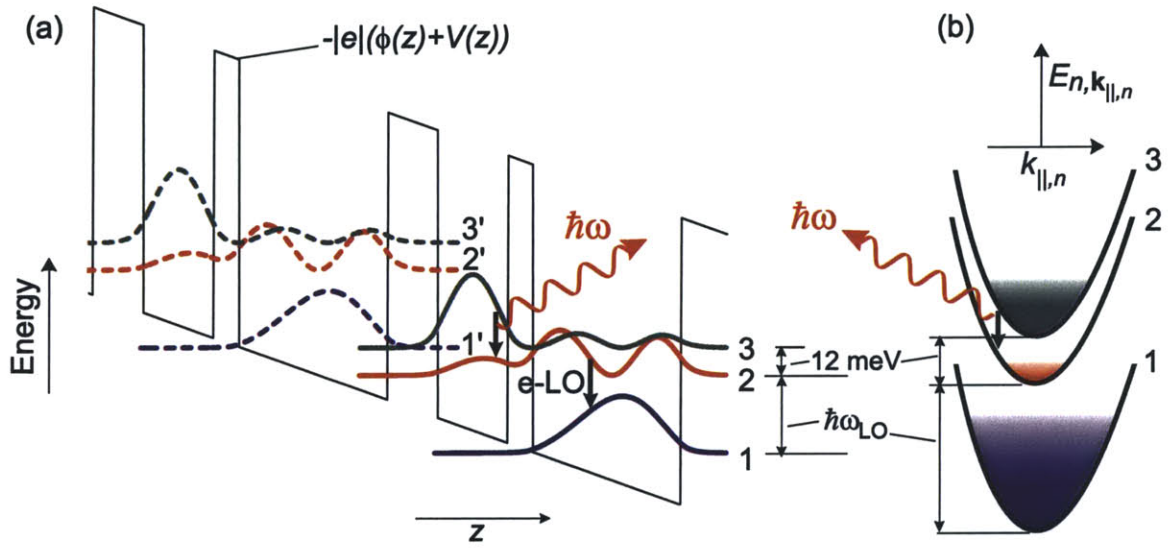


Figure 2-1: (a) Conduction band diagram showing two modules of a simplest possible 3-level terahertz QCL at the $1' - 3$ resonance bias. Square of the envelope wavefunctions ($|\psi_n(z)|^2$, see equation 2.16) for various subbands are plotted, displaced vertically by the subband energy (E_n) in the quantum-wells. Subbands 3 and 2 correspond to the upper and lower radiative levels, and subband 1 acts as an electron-reservoir, also termed as the *injector*. E_{21} may be designed to be $< \hbar\omega_{\mathrm{LO}}$. (b) Schematic of the $E_{n,\mathbf{k}_{\|,n}}$ vs. $k_{\|,n}$ parabolas for the three levels showing a hypothetical distribution of electron populations for a design with population inversion.

conduction band potential profile linear in z, by doping in a location where most of the electrons are expected to reside in the module (for example, in the wide well for Fig. 2-1 since most electrons reside in level 1, which is localized in the wide well).

The Hamiltonian for the light-matter interaction is described more conveniently by a particular choice of a gauge known as the *Coulomb gauge*, in which $\mathbf{A}(\mathbf{r}, t)$ is restricted to satisfy $\nabla \cdot \mathbf{A} = 0$. The Maxwell's equations (2.1) and (2.2) can then be rewritten as

$$\nabla^2 \phi = -\frac{\rho}{\epsilon} \tag{2.6}$$

$$\left(\nabla^2 \mathbf{A} - \mu\epsilon \frac{\partial^2}{\partial t^2}\mathbf{A}\right) - \nabla\left(\mu\epsilon \frac{\partial}{\partial t}\phi\right) = -\mu \mathbf{J} \tag{2.7}$$

Equation (2.6) is known as the *Poisson's equation*, whereas equation (2.7) is the *classical wave equation*. The two equations are de-coupled and can be solved independently. Without the light-matter interaction and in the absence of any other perturbation to the electronic system, ρ, ϕ, and $\mathbf{J}$ are all independent of time if the external electric bias applied across the modules is static. When the light-matter interaction is included, the space-charge density ρ remains independent of time in steady-state operation (as will be explained further in 2.5.1). However, a time dependent current density $\mathbf{J}$ is developed due to the induced electrical polarization in the electronic system ($= \partial \mathcal{P}/\partial t$, where $\mathcal{P}$ is the induced polarization), which acts as a source term to amplify or attenuate the optical field characterized by $\mathbf{A}$, as will be derived in section 2.5.1. Note that the induced $\mathbf{J}$ is a local displacement current density at the frequency of the optical field and may not be measured in the external electrical circuit. If only first order perturbation theory is to be considered, this source term, which is a result of the perturbation (and not a cause of it), need not be considered to derive the light-matter Hamiltonian. Hence, the wave equation to derive the light-matter interaction Hamiltonian can be taken as

$$\nabla^2 \mathrm{A} - \mu\epsilon \frac{\partial^2}{\partial t^2}\mathbf{A} = 0 \tag{2.8}$$

Note that since ϕ is also time independent (for a static external applied bias), the

Lorentz gauge condition, $\nabla \cdot \mathbf{A} + \mu\epsilon\dfrac{\partial}{\partial t}\phi = 0$, is also satisfied. The solution to equation (2.8) is the standard harmonic plane-wave

$$\mathbf{A} = \mathbf{u}\, A_0 \left[e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\right] = \mathbf{u}\, 2A_0 \cos(\mathbf{k}\cdot\mathbf{r} - \omega t) \tag{2.9}$$

for a wave traveling in the direction of $\mathbf{k}$, where $\mathbf{u}$ is a unit vector that specifies the polarization of the plane wave, and is perpendicular to $\mathbf{k}$ (i.e. $\mathbf{u}\cdot\mathbf{k} = 0$) such that $\nabla \cdot \mathbf{A} = 0$ is satisfied. For this to be satisfied, two orthogonal polarizations are possible. Here, $\omega = k/\sqrt{\mu\epsilon}$ is the dispersion relation for the plane-wave, and A_0 is an arbitrary amplitude, taken as a real number for convenience.

Switching from classical to a quantum-mechanical description, the interaction of a free electron with the electromagnetic field can be described by the Hamiltonian [40]

$$\hat{H} = \frac{1}{2m_0}\left(\hat{\mathbf{p}} + |e|\hat{\mathbf{A}}\right)^2 - |e|\hat{\phi} \tag{2.10}$$

where e is the (negative) electronic charge, $\hat{\mathbf{p}}$ $(= -i\hbar\hat{\nabla})$ is the momentum operator, and $\hat{\mathbf{A}}$ and $\hat{\phi}$ are the electromagnetic scalar and vector potential operators, determined from the optical field inside the cavity, and the charge-density distribution inside the quantum-wells, respectively. Note that $\hat{\phi}$ is time-independent for a static external applied bias as was mentioned previously.

For a conduction-band terahertz QCL in a direct band-gap semiconductor (conduction band-edge at the Γ point), the electron wavefunction can be written as the product of a slowly varying envelope wavefunction, and the conduction-band Bloch wavefunction near the bottom-edge of the band, as $\Psi(\mathbf{r},t)u_{c,\mathbf{k}=0}(\mathbf{r})$. Within an effective mass approximation, the free electron mass m_0 in equation (2.10) can then be replaced by the conduction band effective mass m^* ($m^*_{\mathrm{GaAs}} \sim 0.067\, m_0$). The electrostatic potential due to the heterostructure's conduction band edge profile, and the externally applied static electric bias-field across the device can be combined into $\hat{V}(z)$, which can be added as a separate term to the Hamiltonian. The Schrödinger

equation for determining the envelope wavefunction $\Psi(\mathbf{r},t)$ is then written as

$$\hat{H}|\Psi\rangle = i\hbar\frac{\partial}{\partial t}|\Psi\rangle, \quad \text{where,} \quad \hat{H} = \frac{1}{2m^*}\left[\hat{\mathbf{p}} + |e|\hat{\mathbf{A}}\right]^2 - |e|\left[\hat{\phi}(z) + \hat{V}(z)\right] \tag{2.11}$$

For intersubband transitions, the electrons remain close to the bottom of the conduction band, and hence, the Bloch wavefunctions do not change appreciably for such transitions. For this reason, the electromagnetic quantities $\hat{\mathbf{A}}$ and $\hat{\phi}$ couple the envelope wavefunctions only, and can therefore be included in equation (2.11). Expanding the Hamiltonian in equation (2.11) further

$$\hat{H} = \left\{-\frac{\hbar^2}{2m^*}\hat{\nabla}^2 - |e|\left[\hat{\phi}(z) + \hat{V}(z)\right]\right\} + \left\{-\frac{i\hbar|e|}{2m^*}\left[\hat{\nabla}\cdot\hat{\mathbf{A}} + \hat{\mathbf{A}}\cdot\hat{\nabla}\right] + \frac{e^2}{2m^*}\hat{A}^2\right\} \tag{2.12}$$

In the Coulomb gauge ($\nabla\cdot\mathbf{A} = 0$), the chain-rule of differentiation yields $\hat{\nabla}\cdot\hat{\mathbf{A}} = \cancelto{0}{(\hat{\nabla}\cdot\hat{\mathbf{A}})} + \hat{\mathbf{A}}\cdot\hat{\nabla} = \hat{\mathbf{A}}\cdot\hat{\nabla}$. After neglecting the non-linear term with $\hat{A}^2$ as compared to the terms linear in $\hat{\mathbf{A}}$, which is an appropriate assumption for the intensity of the fields involved in a terahertz QCL, the Hamiltonian above becomes

$$\hat{H} = \hat{H}_0 + \hat{H}' \tag{2.13}$$

where

$$\hat{H}_0 = -\frac{\hbar^2}{2m^*}\hat{\nabla}^2 - |e|\left[\hat{\phi}(z) + \hat{V}(z)\right] \tag{2.14}$$

is the unperturbed Hamiltonian, and

$$\hat{H}' = -\frac{i\hbar|e|}{m^*}\hat{\mathbf{A}}\cdot\hat{\nabla} = \frac{|e|}{m^*}\hat{\mathbf{A}}\cdot\hat{\mathbf{p}} \tag{2.15}$$

is the perturbation due to light-matter interaction that induces radiative transitions.

Since the electrons are not confined in the plane of the quantum-wells, the stationary state solutions (i.e. the stationary envelope wavefunctions and the corresponding eigen energies) to the unperturbed Hamiltonian $\hat{H}_0$ could be de-coupled in the directions parallel ($\|$), and perpendicular (z) to the quantum wells, respectively, to

yield

$$\langle \hat{\mathbf{r}} | \Psi \rangle = \Psi_{n,\mathbf{k}_{\parallel,n}}(\mathbf{r}) = \frac{1}{\sqrt{S_{xy}}} e^{i\mathbf{k}_{\parallel,n}\cdot\mathbf{r}_{\parallel}} \psi_n(z) \tag{2.16}$$

$$E_{n,\mathbf{k}_{\parallel,n}} = E_n + E_{\mathbf{k}_{\parallel,n}} = E_n + \frac{\hbar^2 k_{\parallel,n}^2}{2m^*} \tag{2.17}$$

where, n characterizes the subband index, $\mathbf{k}_{\parallel,n}$ is the in-plane wavevector in the n-th subband that determines the kinetic energy inside the subband, and E_n is the energy of the electrons at the bottom of the n-th subband. S_{xy} is the cross-sectional area of the quantum-well layers in real-space, which is included for normalization. These solutions are typically found through an iterative (self-consistent) solution of the Schrödinger equation (2.14) and the Poisson equation (2.6).

Although, a quantum-mechanical description of the electrons has been formalized, the electromagnetic field must also be quantized to fully describe the effect of light-matter interaction on both the quantum-mechanical states of the electrons and that of the electromagnetic modes. However, if only "stimulated emission" into a single cavity mode is considered (which fittingly describes the steady-state operating condition of a laser beyond threshold), the electromagnetic field can be treated classically. The optical optical transition rates due to radiative emission of an inverted electron system into a single cavity mode can be calculated accurately within such a semi-classical treatment. In this case, the vector operator $\hat{\mathbf{A}}$ in the expression for the perturbation $\hat{H}'$ can be replaced by the vector quantity $\mathbf{A}$.

For a harmonic perturbation $\hat{H}' = \hat{V}'^{\pm} e^{\pm i\omega t}$, the rate of transitions (in sec^{-1}) induced between two quantum-mechanical electronic states of $\hat{H}_0$, $|i\rangle$ (initial state) and $|f\rangle$ (final state), can be determined using the *Fermi's Golden Rule* (which was, incidentally, developed by Dirac and not Fermi) [40], as

$$W_{i\rightarrow f}^{\pm} = \frac{2\pi}{\hbar} |V_{fi}'^{\pm}|^2 \delta(E_f - E_i \pm \hbar\omega) \tag{2.18}$$

where the $e^{+i\omega t}$ term causes stimulated emission and the $e^{-i\omega t}$ term causes stimulated absorption. The distinction between the $e^{\pm i\omega t}$ terms is dependent on the convention used and is not absolute. Moreover, since the field is a real quantity, both terms are

always present simultaneously, although only one of the terms satisfies the resonance condition. The δ in the expression is the *Dirac delta* function (with units of inverse of energy in this case), which is a consequence of energy conservation in the transition events. The optical matrix element $V'^{\pm}_{fi}$ is given by

$$V'^{\pm}_{fi} = \langle f|\hat{V}'^{\pm}|i\rangle \tag{2.19}$$

For the "classical" electromagnetic field specified by equation (2.9), the optical matrix element due to the harmonic perturbation of equation (2.15) (with $\mathbf{A}$ instead of $\hat{\mathbf{A}}$), when the perturbation is rewritten in the form $\hat{H}' = \hat{V}'^{\pm}e^{\pm i\omega t}$, becomes

$$V'^{\pm}_{(f,\mathbf{k}_{\parallel,f})(i,\mathbf{k}_{\parallel,i})} = \frac{|e|A_0}{m^*}\mathbf{u}\cdot\langle f,\mathbf{k}_{\parallel,f}|e^{\mp i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{p}}|i,\mathbf{k}_{\parallel,i}\rangle \tag{2.20}$$

where, $\mathbf{k}$ is the wavevector of the electromagnetic field, and $\mathbf{k}_{\parallel,n}$ specifies the in-plane wavevector for an electron in the n-th subband. The optical transition rate between two electronic states is

$$W^{\pm}_{(i,\mathbf{k}_{\parallel,i})\rightarrow(f,\mathbf{k}_{\parallel,f})} = \frac{2\pi}{\hbar}|V'^{\pm}_{(f,\mathbf{k}_{\parallel,f})(i,\mathbf{k}_{\parallel,i})}|^2\delta(E_{f,\mathbf{k}_{\parallel,f}} - E_{i,\mathbf{k}_{\parallel,i}} \pm \hbar\omega) \tag{2.21}$$

where, $|n,\mathbf{k}_{\parallel,n}\rangle$ specifies the n-th "stationary" state of the unperturbed Hamiltonian $\hat{H}_0$, and $E_{n,\mathbf{k}_{\parallel,n}}$ its energy, as determined from equations (2.16) and (2.17).

While the optical matrix element in equation (2.20) is proportional to the momentum matrix element between the two eigen states of $\hat{H}_0$, it is instructive to derive an expression that involves the dipole matrix element instead. This can be done within the so-called *electric-dipole approximation*, in which $e^{i\mathbf{k}\cdot\mathbf{r}}$ can be approximated as ~ 1, as will be argued next. The electromagnetic wavevector inside the semiconductor is of the order of 0.01 μm^{-1} ($\lambda_{\text{GaAs}} = 28$ μm at 3 THz, $k_{\text{GaAs}} = 1/\lambda_{\text{GaAs}} = 0.036$ μm^{-1}), whereas the maximum z extent of $\psi_n(z)$ in the quantum wells is typically of the order of 0.01 μm. Hence, for calculation of the integral along the z direction, $(k_z z)_{\max} \sim 10^{-4}, e^{ik_z z} \approx 1$ in equation (2.20). For the integral along the in-plane direction, $r_{\parallel\max}$ can be approximately taken as the in-plane

diffusion length of the electrons, which is typically of the order of 0.1 μm (as determined from the mobility of the semiconductor in the plane of the quantum wells). Hence, $e^{i\mathbf{k}_{\parallel}\cdot\mathbf{r}_{\parallel}} \approx 1$ also holds (where $\mathbf{k}_{\parallel}$ is the in-plane component of the electromagnetic wavevector $\mathbf{k}$, not to be confused by the electronic wavevectors $\mathbf{k}_{\parallel,n}$). It may be noted that a similar approximation cannot be made for the electronic in-plane wavefunction, $e^{i\mathbf{k}_{\parallel,n}\cdot\mathbf{r}_{\parallel}}$ (equation 2.16), since $k_{\parallel,n}$ is typically much larger. For example, for a kinetic energy of $\sim$ 5 meV for an electron in a given subband, which corresponds to an electronic temperature of $\sim$ 60 K, the in-plane electronic wavevector $k_{\parallel,n}$ is $\sim 100\ \mu\text{m}^{-1}$.

Using the commutation relation $[\hat{\mathbf{p}}, \hat{\mathbf{r}}] = -i\hbar$, it can be shown that $\hat{\mathbf{p}} = \dfrac{im^*}{\hbar}\left[\hat{H}_0, \hat{\mathbf{r}}\right]$. Using this, and the fact that $e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1$, the momentum matrix element in equation (2.20) can be simplified as

$$\begin{aligned}
\langle f, \mathbf{k}_{\parallel,f}|\hat{\mathbf{p}}|i, \mathbf{k}_{\parallel,i}\rangle &= \frac{im^*}{\hbar}\langle f, \mathbf{k}_{\parallel,f}|\left[\hat{H}_0, \hat{\mathbf{r}}\right]|i, \mathbf{k}_{\parallel,i}\rangle \\
&= \frac{im^*}{\hbar}\langle f, \mathbf{k}_{\parallel,f}|\hat{H}_0\hat{\mathbf{r}} - \hat{\mathbf{r}}\hat{H}_0|i, \mathbf{k}_{\parallel,i}\rangle \\
&= \frac{im^*}{\hbar}\left(E_{f,\mathbf{k}_{\parallel,f}} - E_{i,\mathbf{k}_{\parallel,i}}\right)\langle f, \mathbf{k}_{\parallel,f}|\hat{\mathbf{r}}|i, \mathbf{k}_{\parallel,i}\rangle \qquad (2.22)
\end{aligned}$$

Equation (2.22) expresses the momentum matrix element between two states in terms of their dipole matrix element. The optical matrix element in equation (2.20), within the electric-dipole approximation, can thus be rewritten as

$$\begin{aligned}
V'^{\pm}_{(f,\mathbf{k}_{\parallel,f})(i,\mathbf{k}_{\parallel,i})} &= |e|\left(\mp i\omega A_0\right)\mathbf{u}\cdot\langle f, \mathbf{k}_{\parallel,f}|\hat{\mathbf{r}}|i, \mathbf{k}_{\parallel,i}\rangle \\
&= \mp i|e|E_0\mathbf{u}\cdot\langle f, \mathbf{k}_{\parallel,f}|\hat{\mathbf{r}}|i, \mathbf{k}_{\parallel,i}\rangle \qquad (2.23)
\end{aligned}$$

where, $\left(E_{f,\mathbf{k}_{\parallel,f}} - E_{i,\mathbf{k}_{\parallel,i}}\right)$ in equation (2.22) can be replaced by $\mp\hbar\omega$ if the energy conservation term in the optical transition rate is a δ function, and $E_0 = \omega A_0$ is the amplitude of the time-harmonic component of the electric field, which is determined from $\mathbf{E} = -\dfrac{\partial}{\partial t}\mathbf{A} = \mathbf{u}\left(\mp i\omega A_0\right)e^{\pm i\omega t}$ (equations 2.4 and 2.9), in the electric-dipole approximation $e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1$. The optical matrix element above can be identified with the classical expression for the potential energy of an electric dipole with a dipole-

moment $q\mathbf{r}$ as $-q\mathbf{E}\cdot\mathbf{r}$, where $\mathbf{r}$ is the displacement vector from point charge $-q$ to $+q$. Since the perturbation Hamiltonian is given by $\hat{H}' = \hat{V}'^{\pm}e^{\pm i\omega t}$, equation (2.15) can be rewritten in the more intuitive form (within the electric-dipole approximation) as

$$\hat{H}' = |e|\hat{\mathbf{E}}\cdot\hat{\mathbf{r}} \tag{2.24}$$

The dipole matrix element in equation (2.23) can be further simplified as

$$\begin{aligned}\langle f,\mathbf{k}_{\parallel,f}|\hat{\mathbf{r}}|i,\mathbf{k}_{\parallel,i}\rangle &= \langle f|\hat{\mathbf{z}}|i\rangle\langle\mathbf{k}_{\parallel,f}|\mathbf{k}_{\parallel,i}\rangle + \langle f|i\rangle\langle\mathbf{k}_{\parallel,f}|\hat{\mathbf{r}}_{\parallel}|\mathbf{k}_{\parallel,i}\rangle \\ &= \mathbf{z}\, z_{fi}\,\delta^{\mathrm{kr}}(\mathbf{k}_{\parallel,f},\mathbf{k}_{\parallel,i})\end{aligned} \tag{2.25}$$

where, $\mathbf{z}$ is a unit vector in the z direction, and z_{fi} is the dipole-matrix element between the z component of the envelope wavefunctions (explicitly written in equation 2.28). The second term in the equation vanishes for $f \neq i$, which has to be true for a finite energy of optical transition. The dipole-matrix element in equation (2.25) is z directed, which indicates that only the z polarization of the time-harmonic electromagnetic field will couple the electronic states. This is known as the *intersubband polarization selection rule*, which has been experimentally verified [106, 144]. Additionally, the result also indicates that the in-plane wavevector for the electrons needs to be conserved in an optical transition, which is a consequence of momentum conservation (as indicated by the unitless *Kronecker delta* function). This is because the field wavevector is much smaller than the in-plane electronic wavevectors, and hence the field cannot impart much momentum to the electron in an optical transition.

Combining equations (2.25), (2.23), and (2.21), the following expression is obtained for the optical transition rate (in sec^{-1}), in which either $+$ is used for emission, or $-$ for absorption, between two electronic states of the unperturbed Hamiltonian $\hat{H}_0$ (as formulated in equations 2.16 and (2.17))

$$W^{\pm}_{(i,\mathbf{k}_{\parallel,i})\to(f,\mathbf{k}_{\parallel,f})} = |\mathbf{u}\cdot\mathbf{z}|^2\,\frac{2\pi}{\hbar}e^2E_0^2|z_{fi}|^2\delta(E_f - E_i \pm \hbar\omega)\,\delta^{\mathrm{kr}}(\mathbf{k}_{\parallel,f},\mathbf{k}_{\parallel,i}) \tag{2.26}$$

where, the time-harmonic electric field is given by equations (2.4) and (2.9) as

$$\mathbf{E} = \mathbf{u}\, E_0 \left[-ie^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + ie^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\right] = \mathbf{u}\, 2E_0 \sin(\mathbf{k}\cdot\mathbf{r} - \omega t) \tag{2.27}$$

with $\mathbf{u}$ being the unit vector that determines the field polarization, $\mathbf{z}$ is the unit vector in the z direction, and $e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1$ is assumed for calculation of the optical transition rate (the electric dipole approximation). It may be noted that $\delta(E_{f,\mathbf{k}_{\parallel},f} - E_{i,\mathbf{k}_{\parallel},i} \pm \hbar\omega)$ in the original expression can be replaced by $\delta(E_f - E_i \pm \hbar\omega)$, since the subbands have the same $E(k_{\parallel})$ dispersion, and since the in-plane momentum is conserved in the optical transition. The (scalar) dipole matrix element (in meter) is given by

$$z_{fi} = \langle f|\hat{z}|i\rangle = \int_{-\infty}^{+\infty} dz\ \psi_f^*(z)\, z\, \psi_i(z) \tag{2.28}$$

2.1.2 Expression for the optical gain

The net optical transition rate per unit active region volume V_{ac} (in sec^{-1} meter^{-3}) between the initial subband i and the final subband f (assuming that subband i has higher energy, i.e. $E_i > E_f$), accounting for both the stimulated emission and absorption processes related to a single cavity mode, can be written by summing the rates over all the in-plane electronic states as

$$\begin{aligned} R_{i\to f} = R^{\text{em}}_{i\to f} - R^{\text{abs}}_{f\to i} &= \frac{2}{V_{\text{ac}}} \sum_{\mathbf{k}_{\parallel},i} \sum_{\mathbf{k}_{\parallel},f} \Big\{ W^{+}_{(i,\mathbf{k}_{\parallel},i)\to(f,\mathbf{k}_{\parallel},f)} f(E_{i,\mathbf{k}_{\parallel},i}) \left[1 - f(E_{f,\mathbf{k}_{\parallel},f})\right] - \\ &\qquad W^{-}_{(f,\mathbf{k}_{\parallel},f)\to(i,\mathbf{k}_{\parallel},i)} f(E_{f,\mathbf{k}_{\parallel},f}) \left[1 - f(E_{i,\mathbf{k}_{\parallel},i})\right] \Big\} \\ &= |\mathbf{u}\cdot\mathbf{z}|^2 \frac{2\pi}{\hbar} e^2 E_0^2 |z_{fi}|^2 \delta(E_f - E_i + \hbar\omega) \\ &\quad \times \frac{2}{V_{\text{ac}}} \sum_{\mathbf{k}_{\parallel},i} \sum_{\mathbf{k}_{\parallel},f} \left\{ \delta^{\text{kr}}(\mathbf{k}_{\parallel},f, \mathbf{k}_{\parallel},i) \left[f(E_{i,\mathbf{k}_{\parallel},i}) - f(E_{f,\mathbf{k}_{\parallel},f}) \right] \right\} \\ &= |\mathbf{u}\cdot\mathbf{z}|^2 \frac{2\pi}{\hbar} e^2 E_0^2 |z_{fi}|^2 \delta(E_f - E_i + \hbar\omega)(N_i - N_f) \end{aligned} \tag{2.29}$$

where, the optical transition rate expression derived in equation (2.26) has been used for transitions between the in-plane electronic states of the subbands. $f(E_{n,\mathbf{k}_{\parallel},n})$ is

the Fermi function that determines the occupation probability of a particular electronic state in the n-th subband, and a factor of 2 is included to account for electron spin degeneracy in the occupation statistics. The double summation over the in-plane electronic wavevectors $\mathbf{k}_{\|,n}$, which is equivalent to a single summation due to momentum conservation, simply yields the difference of electron population densities N_i and N_f between the two subbands computed over V_{ac} (in meter^{-3}). This is because the statement of energy conservation, $\delta(E_f - E_i + \hbar\omega)$, is independent of the in-plane wavevectors and hence can be taken out of the summation. In other words, all the electrons in the subbands contribute to the optical transition at same ω, equivalent to saying that the *joint density of states* for the optical transition is delta-like in ω. In contrast, for interband transitions in quantum-wells, the energy conservation depends on the in-plane wavevectors. Therefore, the joint density of states and consequently the net transition rate, becomes dependent on ω.

If a population inversion is established in the system (i.e. if $N_i > N_f$), the net optical transition rate $R_{i\rightarrow f}$ from subband $i \rightarrow f$ is positive, leading to a net stimulated emission into the cavity mode being considered. Correspondingly, power would be added to the field as it propagates through the system. Each optical transition adds an energy equivalent of $\hbar\omega$ to the field (the basis of this assumption is in a quantum mechanical description of the field which will be explained in section 2.2). The optical gain coefficient in the semiconductor material g_{mat} (in meter^{-1}), which is calculated by assuming the cavity volume to be the same as that of the active region, can be written as

$$g_{\text{mat}} = \frac{\hbar\omega\,[\text{Joules}] \times R_{i\rightarrow f}\,[\text{sec}^{-1}\ \text{meter}^{-3}]}{\dfrac{\epsilon_0 n_{\text{r}} c (2E_0)^2}{2}\,[\text{Watts meter}^{-2}]} \tag{2.30}$$

where, the quantity in the denominator is the electromagnetic power flow density in the semiconductor for a propagating *plane-wave* with an electric field of the form $\mathbf{E} = \mathbf{u}\,2E_0 \sin(\mathbf{k}\cdot\mathbf{r} - \omega t)$, with n_{r} being the refractive index of the semiconductor.

Whereas a thermal distribution of the electron populations within the subbands does not contribute to linewidth broadening for intersubband gain, the gain spectrum still has a finite linewidth. This is due to homogeneous (collisional) broadening

(caused by "lifetime broadening", and "pure dephasing" due to various scattering mechanisms), as well as inhomogeneous broadening (caused by fluctuations in material growth from one QCL module to another) of the subbands. The broadened linewidth can be approximated by a *normalized* Lorentzian function, which could be incorporated in the expression for the net rate $R_{i\rightarrow f}$ by replacing the Dirac delta (which, in this case, has units of Joules^{-1}) as

$$\delta(E_f - E_i + \hbar\omega) \rightarrow \frac{\Delta E/(2\pi)}{(E_f - E_i + \hbar\omega)^2 + (\Delta E/2)^2} \tag{2.31}$$

where, ΔE is the full-width at half-maximum (FWHM) for the Lorentzian lineshape (in Joules), and is a phenomenological parameter. The basis for this substitution is purely quantum mechanical, and is related to the various relaxation/dephasing mechanisms that can cause damping of the coherent oscillations of the electric-dipole oscillators. It is interesting to note that such broadening mechanisms merely redistribute the transitions strengths over a wider frequency space, while keeping their sum constant. The peak transition strength at resonance is, therefore, lower for a larger linewidth. Combining equations (2.31), (2.30), and (2.29), the following expression is obtained for the optical gain coefficient (in meter^{-1}) due to intersubband radiative transitions between subbands i and f, where $E_i > E_f$, and optical transitions due to only a single electromagnetic mode at ω are considered

$$g_{\text{mat}}(\omega) = \frac{\pi e^2 \omega}{\epsilon_0 n_{\text{r}} c} |z_{fi}|^2 \Delta N_{\text{3D}} \frac{\Delta E/(2\pi)}{(E_f - E_i + \hbar\omega)^2 + (\Delta E/2)^2} \tag{2.32}$$

and the peak gain that is obtained at the resonance condition $\omega = \omega_{if} \triangleq \dfrac{E_i - E_f}{\hbar}$ is

$$\begin{aligned} g_{\text{mat}}^{\text{peak}} = g_{\text{mat}}(\omega_{if}) &= \frac{2e^2}{\hbar\epsilon_0 n_{\text{r}} c} |z_{fi}|^2 \Delta N_{\text{3D}} \frac{\nu_{if}}{\Delta\nu} \\ &= \frac{e^2}{2\pi m^* \epsilon_0 n_{\text{r}} c} \Delta N_{\text{3D}} \frac{f_{fi}}{\Delta\nu} \\ &\overset{\text{GaAs}}{\approx} 70 \, \frac{[\Delta N_{\text{3D}}/(10^{15}\ \text{cm}^{-3})]}{[\Delta\nu/(1\ \text{THz})]} \, f_{fi} \quad (\text{in cm}^{-1}) \end{aligned} \tag{2.33}$$

In these expressions, $\nu_{if} = \omega_{if}/(2\pi)$ (in Hz), $\Delta N_{3D} = N_i - N_f$ is the 3D population inversion density between the subbands computed over V_{ac}, and ΔE has been replaced with $\Delta\nu = \Delta E/h$, which is the FWHM for the linewidth in frequency. The last expression is to aid a quick estimation of gain in GaAs, where $n_r \sim 3.6$ and $m^*_{GaAs} \sim 0.067\, m_0$. Also, the term f_{fi} is defined as the *oscillator strength* of the optical transition, and is written as

$$f_{fi} = \frac{2m^*(E_i - E_f)}{\hbar^2}|z_{fi}|^2 \tag{2.34}$$

It is instructive to write the expression for the peak gain in terms of the oscillator strength, which is a unitless quantity, and is directly related to the strength of the optical transition as is evident from the gain expression. In equation (2.33), there is no explicit frequency dependence and g^{peak}_{mat} only depends on three parameters: ΔN_{3D}, $\Delta\nu$, and f_{fi}.

In classical electromagnetics, the oscillator strength of a single Hertzian dipole is *defined* to be 1. For a system of oscillators that can interact with light at many different resonant frequencies, the oscillator strength is used as a statistical weight indicating the relative "number" of oscillators bound at those frequencies. For example, the relative contribution to the electric susceptibility by different optical-phonon modes in an ionic crystal, due to their coupling with light, is weighted by their oscillator strengths [47].

For the quantum-mechanical system being considered, f_{ij}, as defined in equation (2.34), is the relative strength of the optical absorption from the state $i \rightarrow j$ (conversely f_{ji} determines the relative strength of optical emission from state $i \rightarrow j$). In other words, if this system was irradiated with a spectrally uniform power density over all its transitions, f_{ij} percent of the total absorbed power would be due to transitions from state $i \rightarrow j$ (assuming that the ensemble of the electrons is initialized in the state i, which is taken to be the ground state). This holds regardless of the linewidths of the individual transitions, which is one of the reasons why different optical transitions can be characterized by their oscillator strengths only. For this case,

an oscillator strength *sum-rule* can be written as

$$\sum_j f_{ij} = 1 \tag{2.35}$$

where, the summation is over the complete set of eigenstates for a given Hamiltonian. For the Hamiltonian of the form $\hat{H}_0$ in equation (2.14), the expression for the oscillator strength in equation (2.34) can be shown to satisfy the sum-rule, by using simple commutator relations, and also the completeness relation $\sum_j |j\rangle\langle j| = \hat{I}$, where $\hat{I}$ is the unity operator. It may be noted that a particular value of the oscillator strength f_{ij} can be > 1 within this definition. This can happen if the state i is not the ground state, since in that case, the terms corresponding to the states below i in the summation will be negative. For an ensemble of electrons initialized in state i, the optical transitions corresponding to a negative oscillator strength would emit optical power rather than absorb power, which makes the power absorbed by the $i \rightarrow j$ transitions greater than the net power absorbed by the system.

As a side note, it must be mentioned that sometimes in literature the free electron mass m_0 instead of the effective mass m^* is used in the expression for the oscillator strength (equation 2.34). In that case, the sum-rule summation in equation (2.35) will give m_0/m^*, which can be much different than 1.

2.2 Spontaneous and stimulated emission lifetimes

The light-matter interaction was treated semi-classically in section 2.1, that is, while matter (the electronic system) was treated quantum-mechanically, the interacting optical field was described classically. This is accurate as long as changes to the optical field are not considered as a result of interaction. Hence, the optical transition rates derived in equations (2.26) and (2.29) are correctly described by the semi-classical picture. A quantum mechanical description of the optical field was implicitly implied when the gain was estimated from the net optical transition rate in section 2.1.2 because of two reasons. First, each optical transition was assumed to add an energy

equivalent of $\hbar\omega$ to the field (since $\hbar\omega$ appears in equation 2.30), a fact that cannot be described semi-classically. Second, all optical transitions (for the cavity mode that was being considered) were assumed to contribute coherently to the field. Both these assumptions have their basis in a quantum mechanical description of the optical field. Within these assumptions, the "semi-classical" derivation for gain due to stimulated emission and absorption related to a single cavity mode in section 2.1.2 was accurate. In this section, those assumptions will be justified by using a quantized description of the field. This also allows an estimation of the rates of spontaneous emission, which cannot be obtained from the semi-classical model.

In the Lorentz gauge, the full Hamiltonian for a quantized electromagnetic field can be written as a summation of *plane-wave* Hamiltonians as [66]

$$\hat{H}_{\text{emag}} = \sum_{\sigma}\sum_{\mathbf{k}} \mathbf{u}_{\sigma}\, \hbar\omega_{\mathbf{k}} (\hat{a}^{\dagger}_{\mathbf{k},\sigma}\hat{a}_{\mathbf{k},\sigma} + 1) \tag{2.36}$$

where, $\mathbf{k}$ is the wavevector for a plane-wave, $\mathbf{u}_{\sigma}$ is a unit vector and σ $(= 1, 2)$ symbolizes one of the two orthogonal polarizations that are required for a full description. $\hat{a}^{\dagger}$ and $\hat{a}$ are the creation and the annihilation operators, respectively. The summation includes cavity modes with all possible frequencies $\omega_{\mathbf{k}}$, each two fold degenerate for the polarizations symbolized by σ. In this description, the harmonic component of the electric field operator $\hat{\mathbf{E}}$, which does not commute with $\hat{H}_{\text{emag}}$, can be written as

$$\begin{aligned} \hat{\mathbf{E}}(\mathbf{r}) &= \sum_{\sigma}\sum_{\mathbf{k}} \mathbf{u}_{\sigma} \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2V_{\text{cav}}\epsilon}} (e^{i\mathbf{k}\cdot\mathbf{r}}\hat{a}_{\mathbf{k},\sigma} + e^{-i\mathbf{k}\cdot\mathbf{r}}\hat{a}^{\dagger}_{\mathbf{k},\sigma}) \\ &\approx \sum_{\sigma}\sum_{\mathbf{k}} \mathbf{u}_{\sigma} \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2V_{\text{cav}}\epsilon}} (\hat{a}_{\mathbf{k},\sigma} + \hat{a}^{\dagger}_{\mathbf{k},\sigma}) \end{aligned} \tag{2.37}$$

where, V_{cav} is the volume of the electromagnetic cavity, and ϵ is the dielectric constant of the medium. Also, $e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1$ is assumed within the electric-dipole approximation, as described in section 2.1.1.

A quantum mechanical state of the field for a cavity mode $(\mathbf{k}, \sigma)$, which has n (integer) number of photons, can be written as $|n^{\mathbf{k},\sigma}\rangle$. Whereas, for the electronic

system $|m, \mathbf{k}_{\parallel,m}\rangle$ specifies a particular state in the m-th subband with an in-plane electronic wavevector $\mathbf{k}_{\parallel,m}$. The combined state for the coupled system can be written as $|n_a^{\mathbf{k},\sigma} : a, \mathbf{k}_{\parallel,a}\rangle$, where a is the index that characterizes the coupled state. For the perturbation Hamiltonian of equation (2.24), the optical matrix element between states $|n_a^{\mathbf{k},\sigma} : a, \mathbf{k}_{\parallel,a}\rangle$ and $|n_b^{\mathbf{k},\sigma} : b, \mathbf{k}_{\parallel,b}\rangle$ can be written as a product of two parts since the operators $\hat{\mathbf{E}}$ and $\hat{\mathbf{r}}$ in the perturbation Hamiltonian act on different basis. In this representation, the optical matrix element can be expressed as

$$\begin{aligned} V_{(n_b^{\mathbf{k},\sigma}:b,\mathbf{k}_{\parallel,b})(n_a^{\mathbf{k},\sigma}:a,\mathbf{k}_{\parallel,a})} &= |e| \langle n_b^{\mathbf{k},\sigma} : b, \mathbf{k}_{\parallel,b}|\hat{\mathbf{E}} \cdot \hat{\mathbf{r}}|n_a^{\mathbf{k},\sigma} : a, \mathbf{k}_{\parallel,a}\rangle \\ &= \mathbf{u}_\sigma \cdot \mathbf{z}\, |e| \langle n_b^{\mathbf{k},\sigma}|\hat{E}|n_a^{\mathbf{k},\sigma}\rangle \langle b, \mathbf{k}_{\parallel,b}|\hat{r}|a, \mathbf{k}_{\parallel,a}\rangle \\ &= \mathbf{u}_\sigma \cdot \mathbf{z}\, |e| \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2V_{\text{cav}}\epsilon}} \left[\sqrt{n_a^{\mathbf{k},\sigma}}\, \delta^{\text{kr}}\left(n_b^{\mathbf{k},\sigma}, n_a^{\mathbf{k},\sigma} - 1\right) + \right. \\ &\quad \left. \sqrt{n_a^{\mathbf{k},\sigma} + 1}\, \delta^{\text{kr}}\left(n_b^{\mathbf{k},\sigma}, n_a^{\mathbf{k},\sigma} + 1\right)\right] z_{ba}\, \delta^{\text{kr}}(\mathbf{k}_{\parallel,b}, \mathbf{k}_{\parallel,a}) \end{aligned} \tag{2.38}$$

where, z_{ba} is the dipole-matrix element between subbands b and a, as specified by equation (2.28). As can be noted from the Kronecker delta terms, the electric-field operator only couples those field states, which differ in their photon number by unity.

While the electric field operator causes a state transition in the optical field by creating or annihilating a photon, the position operator couples two electronic states together causing charge dipole oscillation, thus allowing transfer of energy (a photon) from one sub-system to another. The rate at which this process happens is determined from the Fermi's golden rule, as given by equation (2.18). The optical transition rate (in sec^{-1}) from state $|n_a^{\mathbf{k},\sigma} : a, \mathbf{k}_{\parallel,a}\rangle$ to $|n_b^{\mathbf{k},\sigma} : b, \mathbf{k}_{\parallel,b}\rangle$ becomes

$$\begin{aligned} &W^{\text{abs/em}}_{(n_a^{\mathbf{k},\sigma}:a,\mathbf{k}_{\parallel,a})\to(n_b^{\mathbf{k},\sigma}:b,\mathbf{k}_{\parallel,b})} = \\ &\quad |\mathbf{u}_\sigma \cdot \mathbf{z}|^2 \frac{2\pi}{\hbar} e^2 \frac{\hbar\omega_{\mathbf{k}}}{2V_{\text{cav}}\epsilon} |z_{ba}|^2 \delta^{\text{kr}}(\mathbf{k}_{\parallel,b}, \mathbf{k}_{\parallel,a}) \\ &\quad \times \begin{cases} \delta(E_b - E_a - \hbar\omega_{\mathbf{k}})\ \left(n_a^{\mathbf{k},\sigma}\right)\ \delta^{\text{kr}}\left(n_b^{\mathbf{k},\sigma}, n_a^{\mathbf{k},\sigma} - 1\right) & \ldots \text{absorption} \\ \delta(E_b - E_a + \hbar\omega_{\mathbf{k}})\ \left(n_a^{\mathbf{k},\sigma} + 1\right)\ \delta^{\text{kr}}\left(n_b^{\mathbf{k},\sigma}, n_a^{\mathbf{k},\sigma} + 1\right) & \ldots \text{emission} \end{cases} \end{aligned} \tag{2.39}$$

For this transition, the creation operator $\hat{a}^\dagger_{\mathbf{k},\sigma}$, which is associated with the term $e^{+i\omega_{\mathbf{k}}t}$

(in the convention being used), causes a photon to be created in the mode $(\mathbf{k}, \sigma)$, and hence causes stimulated emission into that mode. The converse is true for the (stimulated) absorption process. The electronic state populations determine which of these two is the more dominant process. The emission process is slightly different than absorption, as can be seen from equation (2.39). Even for no photons in the mode being considered ($n_a^{\mathbf{k},\sigma} = 0$), a finite emission rate exists, which is the spontaneous emission process. Even though a single cavity mode may be the dominant mode in a laser, the spontaneous emission process happens into all the available cavity modes $\left(W^{\text{sp}} = \sum_{\mathbf{k},\sigma} W^{\text{em}}_{n_a^{\mathbf{k},\sigma}=0}\right)$ corresponding to a given frequency $\omega_{\mathbf{k}} = (E_a - E_b)/\hbar$, adding noise to the coherent laser field in the cavity. It may be noted that the Fermi's golden rule is a consequence of the first order time-dependent perturbation theory and cannot describe multi-photon processes. The non-linear perturbation term in the original Hamiltonian (equation 2.12), will lead to two-photon processes at high field intensities, and are not accounted for in this analysis.

For stimulated emission/absorption, the optical transition rate derived in equation (2.39) with the quantized electromagnetic field is the same as the semi-classical result of equation (2.26). This can be noted from the fact that the energy density in the lasing mode $(\mathbf{k}, \sigma)$ besides the zero-point term is given by

$$E_{\text{dens},(\mathbf{k},\sigma)}\,[\text{Joules meter}^{-3}] = \frac{n_a^{\mathbf{k},\sigma}\,\hbar\omega_{\mathbf{k}}}{V_{\text{cav}}} = \frac{1}{2}\epsilon\, 4\frac{n_a^{\mathbf{k},\sigma}\,\hbar\omega_{\mathbf{k}}}{2V_{\text{cav}}\epsilon} \tag{2.40}$$

which makes $\dfrac{n_a^{\mathbf{k},\sigma}\hbar\omega_{\mathbf{k}}}{2V_{\text{cav}}\epsilon}$ in equation (2.39) analogous to E_0^2 in the semi-classical expression of equation (2.26) (where $2E_0$ is the amplitude of the time harmonic electric field), thus proving the validity of the argument.

Calculation of the net rate of stimulated transitions from subband $a \rightarrow b$ into an electromagnetic mode $(\mathbf{k}, \sigma)$ follows the exact same steps as in section 2.1.2, equa-

tion (2.29). The net rate (in sec^{-1}) becomes

$$W^{\text{st em}}_{a\to b,(\mathbf{k},\sigma)}\left(=-W^{\text{st abs}}_{b\to a,(\mathbf{k},\sigma)}\right)=|\mathbf{u}_\sigma\cdot\mathbf{z}|^2\frac{2\pi}{\hbar}e^2\frac{\hbar\omega_\mathbf{k}}{2V_{\text{cav}}\epsilon}|z_{ba}|^2\delta(E_b-E_a+\hbar\omega_\mathbf{k})(n_a-n_b)n_a^{\mathbf{k},\sigma}$$
$$\equiv\frac{n_a-n_b}{\tau^{\text{st}}_{a\to b,(\mathbf{k},\sigma)}(\hbar\omega_\mathbf{k})}\equiv\frac{(n_a-n_b)n_a^{\mathbf{k},\sigma}}{\tau^{\text{sp}}_{a\to b,(\mathbf{k},\sigma)}(\hbar\omega_\mathbf{k})} \tag{2.41}$$

Here, n_a and n_b are the electron population *numbers* in the respective subbands. For a mode that is not entirely localized within the active region, a mode confinement factor Γ, which specifies the fraction of the mode in the cavity that spatially overlaps with the active region, should be multiplied into the expression above. A stimulated emission lifetime $\tau^{\text{st}}_{a\to b,(\mathbf{k},\sigma)}(\hbar\omega_\mathbf{k})$ can also be defined, which is proportional to $V_{\text{cav}}/n_a^{\mathbf{k},\sigma}$, where $n_a^{\mathbf{k},\sigma}$ is the *number* of photons in the mode $(\mathbf{k},\sigma)$, and V_{cav} is the volume of the cavity. Correspondingly, $\tau^{\text{sp}}_{a\to b,(\mathbf{k},\sigma)}(\hbar\omega_\mathbf{k})$ is the spontaneous emission lifetime into the mode $(\mathbf{k},\sigma)$. Both these lifetimes are a function of the frequency of the mode $\omega_\mathbf{k}$. To account for broadened linewidth in a real device, the Dirac delta function can be approximated by a normalized Lorentzian function, as in equation (2.31).

In a laser, the photon number $n_a^{\mathbf{k},\sigma}$ builds up in only a single cavity mode (assuming single mode operation). Hence, stimulated emission, which happens in only the lasing mode, is unaffected by the cavity modal properties. However, calculation of the net rate of spontaneous emission requires summation over all the available cavity modes, since spontaneous emission is independent of the level population. The net rate of spontaneous emission (in sec^{-1}) from subband $a\to b$ is thus dependent on the cavity modal properties. Assuming $E_a>E_b$, the net rate can be written as

$$\begin{aligned}
W^{\text{sp em}}_{a\to b} &= 2\sum_{\mathbf{k}_{\|,a}}\sum_{\mathbf{k}_{\|,b}}\sum_{\mathbf{k},\sigma}\left\{W^{\text{em}}_{(n_a^{\mathbf{k},\sigma}:a,\mathbf{k}_{\|,a})\to(n_b^{\mathbf{k},\sigma}:b,\mathbf{k}_{\|,b}),n_a^{\mathbf{k},\sigma}=0}f(E_{a,\mathbf{k}_{\|,a}})\left[1-f(E_{b,\mathbf{k}_{\|,b}})\right]\right\}\\
&= \frac{2\pi}{\hbar}e^2\frac{1}{2\epsilon}|z_{ba}|^2\times 2\sum_{\mathbf{k}_{\|,a}}\sum_{\mathbf{k}_{\|,b}}\left\{\delta^{\text{kr}}(\mathbf{k}_{\|,b},\mathbf{k}_{\|,a})f(E_{a,\mathbf{k}_{\|,a}})\left[1-f(E_{b,\mathbf{k}_{\|,b}})\right]\right\}\\
&\qquad\times\frac{1}{V_{\text{cav}}}\sum_{\mathbf{k},\sigma}\left\{|\mathbf{u}_\sigma\cdot\mathbf{z}|^2\hbar\omega_\mathbf{k}\delta(E_b-E_a+\hbar\omega_\mathbf{k})\right\}\\
&\approx \frac{2\pi}{\hbar}e^2\frac{\hbar\omega_{ab}}{2\epsilon}|z_{ba}|^2n_a\rho_{\text{cav}}(\hbar\omega_{ab})\equiv\frac{n_a}{\tau^{\text{sp}}_{a\to b}}
\end{aligned} \tag{2.42}$$

where, $\omega_{ab} \triangleq \dfrac{E_a - E_b}{\hbar}$, and $\rho_{\text{cav}}(E) = (1/V_{\text{cav}}) \sum_{\mathbf{k},\sigma} \{|\mathbf{u}_\sigma \cdot \mathbf{z}|^2 \, \hbar\omega_{\mathbf{k}} \delta(\hbar\omega_{\mathbf{k}} - E)\}$ is the density of electromagnetic modes in the cavity (in Joules^{-1} meter^{-3}), which can couple to the intersubband transitions. The spontaneous emission lifetime into all available cavity modes is defined as $\tau^{\text{sp}}_{a\to b}$. The population distribution of the photons generated due to these transitions will have a $\delta(\hbar\omega_{\mathbf{k}} - \hbar\omega_{ab})$ dependence in energy space, which can be replaced by a Lorentzian for a broadened linewidth. For a non-unity mode-confinement, the aforementioned factor Γ should also be included in equation (2.42). The final expression above is derived by neglecting "state-blocking" in the lower subband, by assuming $f(E_{b,\mathbf{k}_{\|,b}}) \ll 1$. This is a good approximation for terahertz QCLs, which typically are estimated to have a lower radiative state population of less than 1×10^{10} cm^{-2} [29]. The 2D electronic density of states in the subbands, $g_{2\text{D}} = m^*/(\pi\hbar^2)$, is approximately 2.8×10^{10} cm^{-2} meV^{-1} in GaAs. The in-plane energy distribution of electrons can extend by much more than 10 meV ($T_{\text{e}} \sim 115$ K) above the bottom-edge of the subband even at liquid-helium operating temperatures [31, 29]. Hence, the quasi-Fermi level in the subbands for a thermalized electronic distribution is much below the bottom edge of the subband. As is evident from Fig. 2-2, the electron distribution in the subbands becomes Maxwell-Boltzmann like, and the occupation probability of an electron $f(E_{n,\mathbf{k}_{\|,n}}) \ll 1$.

The density of modes for a cavity, whose dimension is much greater than the wavelength in the material $\lambda_{\text{mat}}(= \lambda_{\mathbf{k}}/n_{\text{r}})$ in all the three-dimensions, can be calculated as

$$\begin{aligned}
\rho^{\text{3D}}_{\text{cav}}(\hbar\omega_{ab}) &= \frac{1}{V_{\text{cav}}} \sum_{\mathbf{k},\sigma} \left[|\mathbf{u}_\sigma \cdot \mathbf{z}|^2 \, \delta(\hbar\omega_{\mathbf{k}} - \hbar\omega_{ab}) \right] \\
&= \frac{1}{(2\pi)^3} \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} d\phi \, d\theta \, dk \; k^2 \sin^3\theta \, \delta\left(\hbar k \frac{c}{n_{\text{r}}} - \hbar\omega_{ab} \right) \\
&= \frac{(\hbar\omega_{ab})^2 n_{\text{r}}^3}{3\pi^2 \hbar^3 c^3}
\end{aligned} \tag{2.43}$$

In this calculation, $|\mathbf{u}_\sigma \cdot \mathbf{z}|^2 = \sin^2\theta$, is due to the intersubband polarization selection rule (equation 2.25), which results in a value that is a factor of 3 smaller than a similar calculation for a randomly polarized electric-field. If a Lorentzian is used instead of the Dirac delta to account for a finite linewidth, the result of the calculation will not

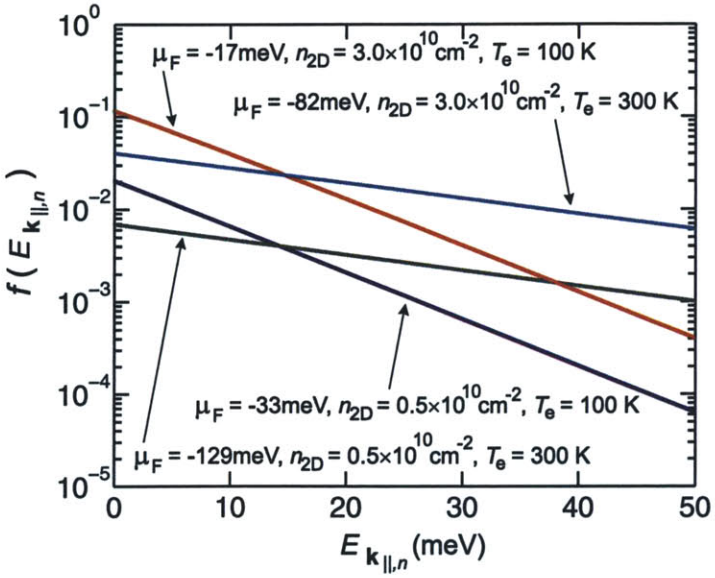


Figure 2-2: Fermi distribution of electrons in the 2D subbands, for GaAs parameters. Subband populations and electronic temperatures, representative of typical terahertz QCLs [31, 29], are chosen. The fermi energy μ_F is measured with reference to the bottom-edge of the subband.

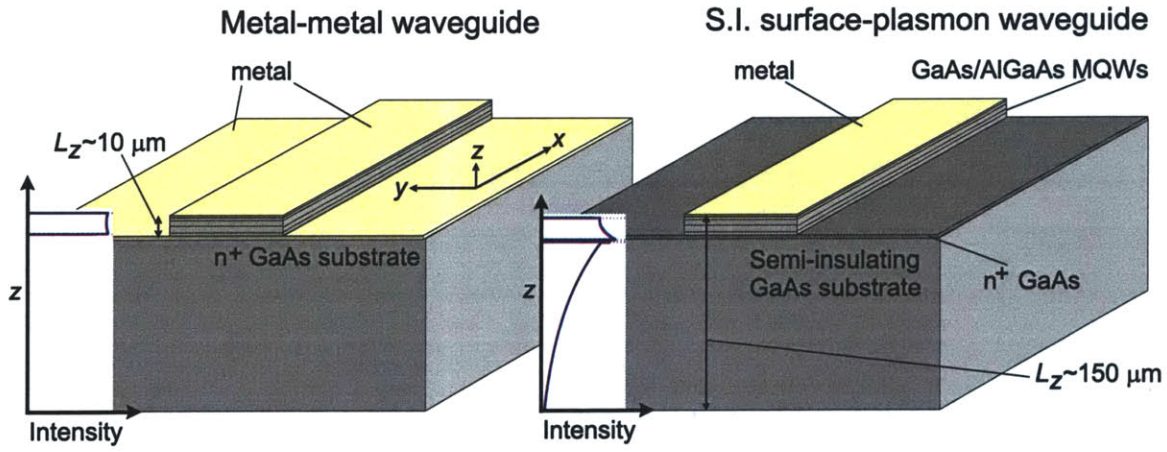


Figure 2-3: Schematic and mode-shape plots at $\nu \sim 3$ THz ($\lambda_{\mathrm{GaAs}} \sim 28$ μm) for the two types of waveguides that are used for terahertz QCLs (details in section 3.6.1).

change appreciably, as long as the linewidth $\Delta\omega \ll \omega_{ab}$.

At terahertz frequencies ($\nu = 1 - 10$ THz, $\lambda = 300 - 30$ μm), the wavelength inside the semiconductor ($= \lambda/n_{\rm r}$) is typically greater than twice the thickness of the active region used. Hence, the modes that have any significant overlap with the active region, have a negligible value of the wavevector along the height of the waveguide (in their plane-wave decomposition). Figure 2-3 shows the schematic and the typical mode-plots for the two types of waveguides that are used for terahertz QCLs. In a metal-metal waveguide, all the supported modes have $k_z \approx 0$ due to its sub-wavelength vertical dimension. For a semi-insulating surface-plasmon waveguide, the "active" modes also have $k_z \approx 0$ (all the rest of the modes are mostly localized in the substrate rather than the active region, which makes the corresponding mode-confinement factor $\Gamma \approx 0$). Due to this reason, the density of available modes in a terahertz QCL cavity can be determined from a 2D calculation as

$$\begin{aligned}
\rho_{\rm cav}^{\rm THz}(\hbar\omega_{ab}) = \frac{1}{L_z}\rho_{\rm cav}^{\rm 2D}(\hbar\omega_{ab}) &= \frac{1}{L_z}\frac{1}{A_{\rm cav}}\sum_{\mathbf{k},\sigma}^{k_z=0}\left[|\mathbf{u}_\sigma\cdot\mathbf{z}|^2\,\delta(\hbar\omega_{\mathbf{k}}-\hbar\omega_{ab})\right] \\
&= \frac{1}{L_z}\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{\infty} d\theta\, dk\; k\,\delta\left(\hbar k\frac{c}{n_{\rm r}}-\hbar\omega_{ab}\right) \\
&= \frac{1}{L_z}\frac{(\hbar\omega_{ab})n_{\rm r}^2}{2\pi\hbar^2c^2} = \frac{3}{4}\frac{(\lambda_{ab}/n_{\rm r})}{L_z}\rho_{\rm cav}^{\rm 3D}(\hbar\omega_{ab}) \qquad (2.44)
\end{aligned}$$

where, L_z is the cavity dimension in the z direction, $A_{\rm cav}$ is the area of the cavity in the xy plane, $\lambda_{ab} = (2\pi c/\omega_{ab})$, and $|\mathbf{u}_\sigma\cdot\mathbf{z}|^2 = 1$ since for $k_z = 0$ the electric-field is polarized along the z direction, which automatically satisfies the polarization selection rule. For metal-metal waveguides, the cavity thickness L_z is smaller than the wavelength inside the material $\lambda_{ab}/n_{\rm r}$, which enhances the density of modes, and hence the spontaneous emission rate. This is the so-called *Purcell* effect, which was first analyzed in 1946 by Purcell [129], who predicted an enhancement in the rate of nuclear magnetic moment transitions for a nuclear-magnetic medium placed in a resonant electrical circuit. The alteration of spontaneous emission rate has since been studied extensively in optical microcavities, in which at least one of the cavity's dimension is on the order of a wavelength. Optical confinement rearranges the usual

free-space spectral mode density and the density of modes at some wavelengths is increased, whereas at others, it is decreased depending on cavity resonances. For an off-resonant cavity, the spontaneous emission can conversely be inhibited [85], instead of being enhanced. For a planar microcavity surrounded by two metallic-mirrors in the z direction the spontaneous emission rate is always enhanced for sub-wavelength dimension in the z direction [26], as predicted by equation (2.44) for $L_z < (3/4)\lambda_{ab}/n_{\rm r}$.

It is instructive to write expressions for the material gain $g_{\rm mat}$ (in meter^{-1}), which is defined as the increase in photon number in the lasing mode per unit length, in terms of the stimulated and the spontaneous emission lifetimes, respectively. Using equations (2.40), (2.41), and (2.42), the corresponding expressions become

$$g_{\rm mat}(\hbar\omega) = \frac{(n_a - n_b)}{\tau_{\rm st}(\hbar\omega) n_{\rm ph}} \frac{n_{\rm r}}{c} = \frac{(n_a - n_b)}{\tau'_{\rm sp}(\hbar\omega)} \frac{n_{\rm r}}{c} \tag{2.45}$$

$$= \frac{(n_a - n_b)}{\tau_{\rm sp}} \frac{n_{\rm r}}{c} \left[\frac{|\mathbf{u}_\sigma \cdot \mathbf{z}|^2 \delta(\hbar\omega - \hbar\omega_{ab})}{V_{\rm cav} \rho_{\rm cav}(\hbar\omega_{ab})} \right] \tag{2.46}$$

where, $\tau_{\rm st}(\hbar\omega) \equiv \tau^{\rm st}_{a\to b,(\mathbf{k},\sigma)}(\hbar\omega_{\mathbf{k}})$ is the lifetime due to stimulated emission into the lasing mode, $\tau'_{\rm sp}(\hbar\omega) \equiv \tau^{\rm sp}_{a\to b,(\mathbf{k},\sigma)}(\hbar\omega_{\mathbf{k}})$ is the lifetime due to spontaneous emission into the lasing mode, $\tau_{\rm sp} \equiv \tau^{\rm sp}_{a\to b}$ is the lifetime due to spontaneous emission into all available modes, and $n_{\rm ph} \equiv n_a^{\mathbf{k},\sigma}$ is the number of photons in the lasing mode. While equation (2.45) is exact, equation (2.46) holds true within the assumption of negligible state-blocking, as was done in equation (2.42). These expressions, which are another form for equation (2.32), are written in terms of *numbers* rather than *densities* to avoid any confusion between the volumes of the active region and the cavity. Here, n_a and n_b are the respective electron numbers in subbands a and b, $n_{\rm ph}$ is the photon number in the lasing cavity mode, $n_{\rm r}$ is the material refractive index, and c is the speed of light in vacuum. To account for a finite linewidth, the Dirac delta function $\delta(\hbar\omega - \hbar\omega_{ab})$ (which also appears in the expressions for $\tau_{\rm st}$ and $\tau'_{\rm sp}$) should be replaced by a normalized Lorentzian, as in equation (2.31).

It may be mentioned that the *microcavity effect*, which reduces the spontaneous emission lifetime $\tau_{\rm sp}$ by increasing the density of available modes $\rho_{\rm cav}(\hbar\omega_{ab})$, does not affect the gain in the active region. Gain, like loss, is a material property, which is

due to a modification in (predominantly) the imaginary component of the electrical susceptibility $\chi(\omega) = \chi'(\omega) + i\chi''(\omega)$ of a medium, as will be shown in section 2.5.1. For an electrical field $\mathcal{E}e^{i\omega t}$ propagating in such a medium, a polarization current density $-\omega\epsilon_0\chi''\mathcal{E}e^{i\omega t}$ is induced which is 180° out of phase to the original field and hence adds energy to the field. The lasing action is a linear process as opposed to the non-linear processes of optical parametric oscillators (OPO), second-harmonic generation (SHG), and Raman effect. As such, the gain coefficient is independent of the field strength. Consequently, the modal properties of the cavity do not come into play in this picture, and thus the microcavity effect should not change the material gain of the active region. To verify this experimentally, terahertz QCLs in metal-metal waveguide microcavities with different thicknesses (L_z) have been experimentally tested, the results of which will be presented in chapter 3 (section 3.3.2).

Using the expressions developed in equations (2.41) and (2.42), the values for the spontaneous and stimulated emission lifetimes in GaAs can be calculated as

$$\begin{aligned}
\tau'_{\rm sp}(\hbar\omega_{ab}) &= 1.7\,\frac{[\Delta\nu/(1\,\text{THz})]\,[V_{\rm cav}/(10^6\,\mu\text{m}^3)]}{f} \quad (\text{in ms}) \\
\tau_{\rm st}(\hbar\omega_{ab})\left(=\frac{\tau'_{\rm sp}(\hbar\omega_{ab})}{n_{\rm ph}}\right) &= 6.0\,\frac{[\Delta\nu/(1\,\text{THz})]\,[\nu_{ab}/(3\,\text{THz})]}{f\,[E^2_{\rm amp}/(1\,\text{kV/cm})^2]} \quad (\text{in ps}) \\
\tau_{\rm sp} &= 4.0\,\frac{[L_z/(10\,\mu\text{m})]}{f\,[\nu_{ab}/(3\,\text{THz})]} \quad (\text{in }\mu\text{s})
\end{aligned} \tag{2.47}$$

where, $\tau_{\rm st}$ and $\tau'_{\rm sp}$ are evaluated at the resonance frequency $\omega = \omega_{ab}$, and an electric field of the form $\mathbf{E} = \mathbf{z}\,E_{\rm amp}\sin(\mathbf{k}\cdot\mathbf{r} - \omega_{ab}t)$ is assumed inside the cavity, which corresponds to the number of photons $n_{\rm ph} = \dfrac{V_{\rm cav}\epsilon_0 n_{\rm r}^2 E^2_{\rm amp}}{2\hbar\omega_{ab}}$. For evaluation of $\tau_{\rm sp}$, the expression in equation (2.44) is used to calculate the density of modes. f is the oscillator strength for $a \rightarrow b$ optical transition, $V_{\rm cav}$ is the modal volume in the cavity, ν_{ab} is the lasing frequency, and $\Delta\nu$ is the linewidth of the optical transition.

Finally, it must be mentioned that expressions derived in this section need some modification before they could be used for a real system. The material gain $g_{\rm mat}$ is usually defined for a case when the entire electromagnetic cavity is filled with the active gain medium. For a cavity where only a fraction (= Γ) of the mode overlaps

with the active region, the stimulated radiative lifetimes derived in this section become longer by a factor of $1/\Gamma$, and consequently the actual gain is reduced by the factor Γ. Also, the expressions derived here are for *plane-wave* propagation inside the material, therefore, the gain is for the direction of the propagation. In a waveguide, the propagating mode along the *waveguided* direction could be a superposition of plane-waves. In that case, the expression for gain in the waveguided direction should be modified, as will be shown in chapter 4.

2.3 Rate equations analysis of a 3-level QCL

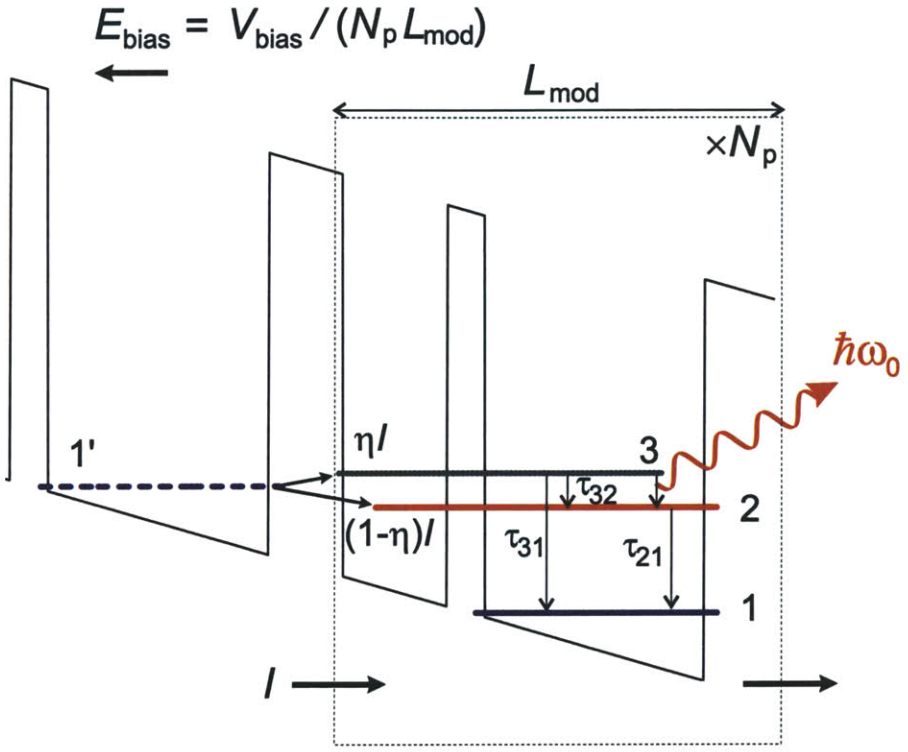


Figure 2-4: Schematic to describe rate equations analysis of a 3-level QCL.

Although, electron transport calculation in a terahertz QCL is a complex problem [28], simple rate equations can be written to solve for a coupled electronic and electromagnetic system once certain parameters are available. Figure 2-4 shows the schematic of a representative QCL, which was initially presented in Fig. 2-1. Dynamic electron and photon populations can be determined as a function of various

non-radiative and radiative lifetimes, and the source term I, which is the current flowing through the quantum-cascade structure. Here, 3 is the upper radiative level, 2 is the lower radiative level, and 1 is a reservoir (injector) level. η is the fraction of the current injected into the upper level 3 and is a function of current I itself, which in turn is a function of the static electric-field bias across the device E_{bias}. In general, I and η are much harder to determine computationally in an accurate manner, but once they are taken as parameters, the rate equations describe the dynamics of the coupled electron-photon system well. The rate equations are written as

$$\frac{dn_3}{dt} = \frac{\eta I N_\text{p}}{|e|} - \frac{n_3}{\tau_{32}} - \frac{n_3}{\tau_{31}} - \frac{n_3}{\tau_\text{sp}}\frac{\Gamma}{N_\text{p}} - (n_3 - n_2)\frac{\Gamma}{N_\text{p}}\int \frac{n_\text{ph}(\hbar\omega)}{\tau'_\text{sp}(\hbar\omega)}\hbar d\omega \tag{2.48}$$

$$\frac{dn_2}{dt} = \frac{(1-\eta) I N_\text{p}}{|e|} - \frac{n_2}{\tau_2} + \frac{n_3}{\tau_{32}} + \frac{n_3}{\tau_\text{sp}}\frac{\Gamma}{N_\text{p}} + (n_3 - n_2)\frac{\Gamma}{N_\text{p}}\int \frac{n_\text{ph}(\hbar\omega)}{\tau'_\text{sp}(\hbar\omega)}\hbar d\omega \tag{2.49}$$

$$\frac{dn_\text{ph}(\hbar\omega)}{dt} = \frac{n_3}{\tau_\text{sp}}\frac{\Gamma}{N_\text{p}}\mathcal{L}(\hbar\omega) + (n_3 - n_2)\frac{\Gamma}{N_\text{p}}\frac{n_\text{ph}(\hbar\omega)}{\tau'_\text{sp}(\hbar\omega)} - \frac{n_\text{ph}(\hbar\omega)}{\tau_\text{ph}} \tag{2.50}$$

The symbols used in these expressions have the following meaning:

- I is the *electronic* current (in the direction of electron transport) flowing through the device.
- N_p is the number of repeated modules in the QCL.
- n_3 and n_2 are the electron population numbers in subbands 3 and 2, respectively, *summed* over all QCL modules.
- Γ is the fraction of the optical mode that overlaps with the *entire* active region.
- $\tau'_\text{sp}(\hbar\omega)N_\text{p}/\Gamma$ is the spontaneous emission lifetime into a single cavity mode at frequency ω, and $\tau_\text{sp}N_\text{p}/\Gamma$ is the spontaneous emission lifetime into all available cavity modes (as determined from equations 2.41, 2.42, 2.45 and 2.46).
- τ_{32}, τ_{31}, and τ_2 are the lifetimes due to non-radiative transitions $3 \rightarrow 2$, $3 \rightarrow 1$, and $2 \rightarrow 1$, respectively.

- $\tau_{\rm ph}$ is the photon lifetime in the cavity, which can be expressed as

$$\frac{1}{\tau_{\rm ph}} = \frac{c}{n_{\rm r}}(\alpha_{\rm w} + \alpha_{\rm m}) \tag{2.51}$$

where, $\alpha_{\rm w}$ and $\alpha_{\rm m}$ are the waveguide and the mirror losses (in meter^{-1}), respectively, c is the speed of light in vacuum, and $n_{\rm r}$ is the refractive index of the active region. Typically $\alpha_{\rm w} = \alpha_{\rm mat}\Gamma$, where $\alpha_{\rm mat}$ is the loss in the semiconductor active region material, however, additional loss terms have to be added if significant fraction of the mode propagates in a lossy inactive material. The "mirror" loss is due to out-coupling of photons from the cavity, and is discussed in greater detail in chapter 4. $\alpha_{\rm w}$ is typically a slowly varying function of frequency ω. However, $\alpha_{\rm m}$ can depend critically on both ω (such as in a cavity with frequency dependent feedback), or the direction of the mode propagation.

- $n_{\rm ph}(\hbar\omega)$ is the number of photons per unit energy (in Joules^{-1}) generated into all available cavity modes, due to both spontaneous and stimulated optical transitions. $\int n_{\rm ph}(\hbar\omega)\hbar d\omega$ gives the total number of photons in the cavity, and $\int \hbar\omega n_{\rm ph}(\hbar\omega)\alpha_{\rm m}(c/n_{\rm r})\hbar d\omega$ is the total amount of optical power emitted.

-

$$\mathcal{L}(\hbar\omega) = \frac{\hbar\Delta\omega/(2\pi)}{(\hbar\omega - \hbar\omega_{32})^2 + (\hbar\Delta\omega/2)^2} \tag{2.52}$$

is the normalized Lorentzian lineshape function (in Joules^{-1}), such that $\hbar\Delta\omega$ is the FWHM linewidth, and $\int \mathcal{L}(\hbar\omega)\hbar d\omega = 1$.

It may be noted that all the lifetimes depend on the subband wavefunctions, which in turn depend on $E_{\rm bias}$, and hence on I. However, the bias does not change significantly in the range of laser operation, therefore, the lifetimes are taken as constants. Also, the lifetimes $\tau'_{\rm sp}$, $\tau_{\rm sp}$, and $\tau_{\rm ph}$ are strictly derived only for plane-wave propagation modes. For modes that can be represented as a superposition of plane-waves, their expressions need to be slightly modified as shown in chapter 4.

2.3.1 Operation below lasing threshold

Below lasing threshold, any photons that are emitted are distributed amongst all cavity modes and $n_{\rm ph}(\hbar\omega)$ has a broad linewidth (as will be derived below). As was estimated for GaAs in equation (2.47), the lifetimes due to radiative transitions, $\left[\frac{\Gamma}{N_{\rm p}}\int\frac{n_{\rm ph}(\hbar\omega)}{\tau'_{\rm sp}(\hbar\omega)}\hbar d\omega\right]^{-1}$ and $\frac{\tau_{\rm sp}N_{\rm p}}{\Gamma}$, are typically many orders of magnitude greater than the non-radiative lifetimes. In the steady-state ($d/dt \rightarrow 0$), equations (2.48) and (2.49) yield for subband populations

$$\begin{aligned} n_3 &= I\frac{\eta_I N_{\rm p}}{|e|}\frac{\tau_{31}\tau_{32}}{(\tau_{31}+\tau_{32})} \\ n_2 &= I\frac{N_{\rm p}}{|e|}\frac{\tau_2(\tau_{31}+(1-\eta_I)\tau_{32})}{(\tau_{31}+\tau_{32})} \end{aligned} \tag{2.53}$$

Below threshold, the current flowing through the structure I is not enough to develop sufficient population inversion $(n_3 - n_2)$ for the gain to counter the losses for any cavity mode. Thus, a population of coherent photons cannot be developed in any cavity mode. However, depending on the electron population of the upper level 3, significant number of spontaneously emitted photons could build up inside the cavity. Their distribution in frequency is determined by the population difference $(n_3 - n_2)$ (which determines whether there is gain or loss due to optical transitions between subbands 3 and 2), and the photon lifetime in the cavity $\tau_{\rm ph}$ (which may itself be frequency dependent). In steady state, $n_{\rm ph}(\hbar\omega)$ could be determined from equation (2.50) as

$$n_{\rm ph}(\hbar\omega) = \frac{\frac{n_3}{\tau_{\rm sp}}\frac{\Gamma}{N_{\rm p}}\mathcal{L}(\hbar\omega)}{\frac{1}{\tau_{\rm ph}} - \frac{(n_3-n_2)}{\tau'_{\rm sp}(\hbar\omega)}\frac{\Gamma}{N_{\rm p}}} \tag{2.54}$$

$$\equiv \frac{A\,n_3\,\mathcal{L}(\hbar\omega)}{1 - B\,(n_3-n_2)\,\mathcal{L}(\hbar\omega)}$$

$$= A\,n_3\,\frac{\Delta\omega}{\Delta\omega'}\,\mathcal{L}'(\hbar\omega) \tag{2.55}$$

where, using equations (2.42) and (2.41), A and B are written as

$$\begin{aligned} A &= \frac{\Gamma}{N_p}\frac{2\pi}{\hbar}e^2\frac{\hbar\omega_{32}}{2\epsilon}|z_{23}|^2\rho_{cav}(\hbar\omega) \\ B &= \frac{\Gamma}{N_p}\frac{2\pi}{\hbar}e^2\frac{\hbar\omega}{2V_{cav}\epsilon}|z_{23}|^2 \end{aligned} \tag{2.56}$$

The Lorentzian in the denominator of equation (2.54) is due to the fact that $\tau'_{sp}(\hbar\omega) \propto 1/\mathcal{L}(\hbar\omega)$. Assuming τ_{ph}, A, and B to be slowly varying functions of frequency (as compared to the Lorentzian $\mathcal{L}(\hbar\omega)$), equation (2.54) can be simplified further and written in terms of a modified Lorentzian $\mathcal{L}'(\hbar\omega)$, which has a linewidth $\Delta\omega' = \Delta\omega\sqrt{1 - \frac{2B(n_3 - n_2)}{\pi\hbar\Delta\omega}}$ where $\Delta\omega$ is the linewidth of $\mathcal{L}(\hbar\omega)$, which corresponds to that of the radiative transition $3 \rightarrow 2$. The center frequency of $\mathcal{L}'(\hbar\omega)$ remains the same as ω_{32}.

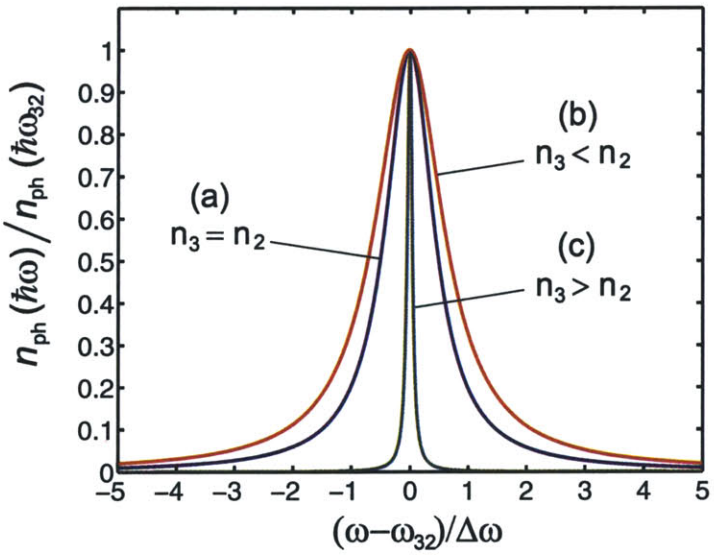


Figure 2-5: $n_{ph}(\hbar\omega)$ distribution in frequency for three hypothetical cases. (a) No population inversion, for which the linewidth $\Delta\omega'$ remains the same as $\Delta\omega$. (b),(c) With loss and gain such that $\frac{2B|n_3 - n_2|}{\pi\hbar\Delta\omega} = 0.99$.

The frequency distribution of the emitted optical power is approximately the same as that of $n_{ph}(\hbar\omega)$, which is shown for three hypothetical cases in Fig. 2-5. The linewidth of the emission spectrum $\Delta\omega'$ is narrowed as population inversion increases, and eventually $\Delta\omega' \rightarrow 0$, which characterizes onset of lasing (section 2.3.2).

The total intensity of the spontaneous emission increases by an amount $\Delta\omega/\Delta\omega'$ due to the gain in the active region. It may be noted that in case of $n_2 > n_3$, there is loss due to intersubband absorption rather than gain, and the linewidth $\Delta\omega'$ conversely becomes broader than $\Delta\omega$. This is because photons with frequencies close to the center frequency ω_{32} encounter greater intersubband loss as compared to the photons at frequencies on the wings of the Lorentzian $\mathcal{L}(\hbar\omega)$. The broadening is typically small even for a relatively large $(n_2 - n_3)$ as seen from Fig. 2-5. Moreover, it is less likely for this broadening to be measured experimentally, since the spontaneous emission power is bound to be low when population in level 3 is small.

2.3.2 Operation near lasing threshold

As the lasing threshold is approached, the linewidth for $n_{\text{ph}}(\hbar\omega)$ becomes infinitesimally small and the photons at the peak frequency are all due to stimulated transitions into a single cavity mode (the *lasing* mode). Hence, $n_{\text{ph}}(\hbar\omega)$ can be approximated as $n_{\text{ph}}^{\text{st}}\delta(\hbar\omega - \hbar\omega_0)$ where, $\delta(\hbar\omega)$ is the Dirac-delta function (in Joules^{-1}), and $n_{\text{ph}}^{\text{st}}$ is the total *number* of photons in the lasing mode (at some particular frequency ω_0 which may not necessarily be ω_{32}). However, $n_{\text{ph}}^{\text{st}}$ may not necessarily be a large number and the lifetime due to stimulated transitions, which now becomes $\dfrac{\tau'_{\text{sp}}(\hbar\omega_0)N_{\text{p}}}{n_{\text{ph}}^{\text{st}}\Gamma}$, can still be many orders of magnitude greater than the non-radiative lifetimes (for example, for the typical parameters given in equation 2.47, a value of $n_{\text{ph}}^{\text{st}} = 1.7 \times 10^8$ is needed for the stimulated lifetime to become as small as 10 ps). In the steady-state, equations (2.48) and (2.49) yield for subband populations

$$\begin{aligned} n_3 \equiv n_3^{\text{th}} &= I_{\text{th}}\frac{\eta_{I\text{th}}N_{\text{p}}}{|e|}\frac{\tau_{31}\tau_{32}}{(\tau_{31}+\tau_{32})} \\ n_2 \equiv n_2^{\text{th}} &= I_{\text{th}}\frac{N_{\text{p}}}{|e|}\frac{\tau_2(\tau_{31}+(1-\eta_{I\text{th}})\tau_{32})}{(\tau_{31}+\tau_{32})} \end{aligned} \tag{2.57}$$

These expressions are similar to those in equation (2.53). I_{th} identifies the current as the threshold current, which establishes enough population inversion to satisfy the threshold condition specified by equation (2.50) for a solution of the form $n_{\text{ph}}^{\text{sp}}(\hbar\omega) =$

$n_{\rm ph}^{\rm st}\delta(\hbar\omega - \hbar\omega_0)$. In this case, the optical gain g (in meter^{-1}) in the cavity becomes equal to the optical losses (as characterized by $\tau_{\rm ph}$, which is typically in the range of $1 - 10$ ps for terahertz QCLs), and can be written as

$$\begin{aligned} g(\hbar\omega_0) &= (n_3^{\rm th} - n_2^{\rm th})\frac{\Gamma}{N_{\rm p}}\frac{1}{\tau'_{\rm sp}(\hbar\omega_0)}\frac{n_{\rm r}}{c} \\ &= \frac{I_{\rm th}}{|e|}\frac{(\eta_{I{\rm th}}\tau_{31}\tau_{32} - \tau_2\tau_{31} - (1-\eta_{I{\rm th}})\tau_2\tau_{32})}{(\tau_{31}+\tau_{32})}\frac{\Gamma}{\tau'_{\rm sp}(\hbar\omega_0)}\frac{n_{\rm r}}{c} \\ &= \frac{1}{\tau_{\rm ph}}\frac{n_{\rm r}}{c} = (\alpha_{\rm w} + \alpha_{\rm m}) \end{aligned} \tag{2.58}$$

Note that even though the linewidth for $n_{\rm ph}(\hbar\omega)$ distribution becomes infinitesimally small, the photons generated purely due to spontaneous emission, $\frac{n_3}{\tau_{\rm sp}}\frac{\Gamma}{N_{\rm p}}\mathcal{L}(\hbar\omega)$, are still present in the cavity. Although, the intensity of such photons is many orders of magnitude lower compared to the photon intensity at the lasing frequency. The expressions above allow determination of $I_{\rm th}$ as a function of cavity losses. $\tau'_{\rm sp}(\hbar\omega_0)$ in these expressions is equivalent to $\tau^{\rm sp}_{a\to b,(\mathbf{k},\sigma)}(\hbar\omega_{\mathbf{k}})$ of equation (2.41), which can be used to evaluate it for subbands 3 and 2, respectively, after replacing the Dirac delta with the Lorentzian $\mathcal{L}(\hbar\omega_0)$ of equation (2.52).

2.3.3 Operation above lasing threshold

Above lasing threshold, the photon number in the lasing mode $n_{\rm ph}^{\rm st}$ builds up significantly, and correspondingly the lifetime due to stimulated transitions $\frac{\tau'_{\rm sp}(\hbar\omega_0)N_{\rm p}}{n_{\rm ph}^{\rm st}\Gamma}$ cannot be neglected in comparison to non-radiative lifetimes. The solutions to equations (2.48), (2.49), and (2.50) in steady-state result in

$$\begin{aligned} n_3 &= n_3^{\rm th} + (I - I_{\rm th})\frac{N_{\rm p}}{|e|}\frac{\tau_2\tau_{31}}{(\tau_2+\tau_{31})} \\ n_2 &= n_2^{\rm th} + (I - I_{\rm th})\frac{N_{\rm p}}{|e|}\frac{\tau_2\tau_{31}}{(\tau_2+\tau_{31})} \\ n_{\rm ph}^{\rm st} &= \frac{N_{\rm p}}{|e|}\tau_{\rm ph}\left[(I - I_{\rm th})\chi + I(\eta_I - \eta_{I{\rm th}})\right] \end{aligned} \tag{2.59}$$

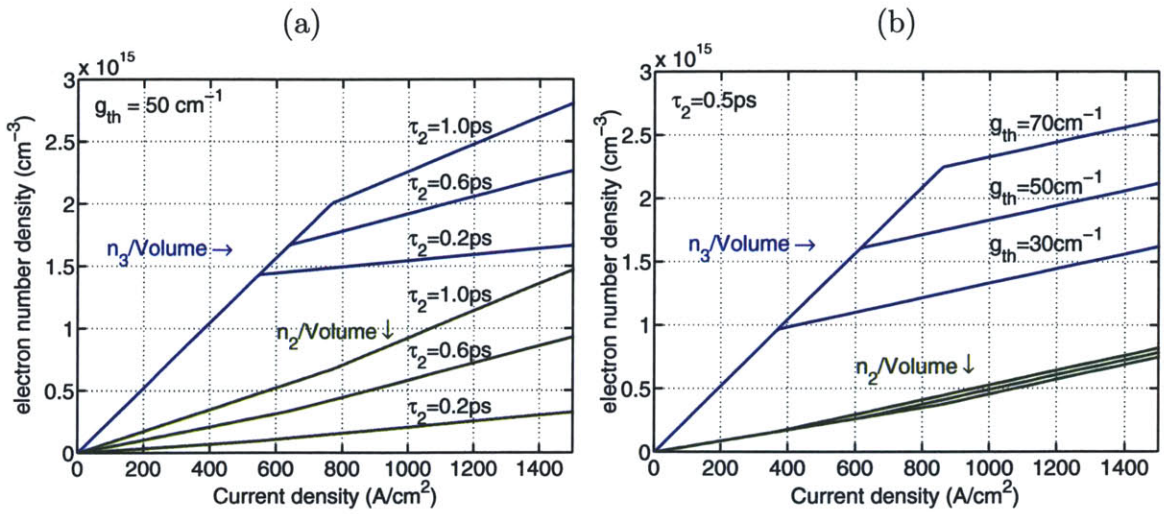


Figure 2-6: Population densities of the radiative levels n_3 and n_2 as a function of current per unit cross-sectional area. The following values are used for the plots: $\eta_{I\mathrm{th}} = \eta_I = 1$, $\tau_{32} = 3$ ps, $\tau_{31} = 7$ ps, $f_{23} = 0.8$, and $\Delta\nu = 1.5$ THz, where the threshold gain $g_{\mathrm{th}} \equiv (\alpha_{\mathrm{w}} + \alpha_{\mathrm{m}})$ is related to the population inversion $\Delta n_{\mathrm{th}} \equiv (n_3^{\mathrm{th}} - n_2^{\mathrm{th}})$ by equation (2.33). $L_{\mathrm{mod}} = 50$ nm is assumed to calculate 3D densities. The discontinuity in the curves corresponds to the threshold current-density.

where,

$$\chi = \frac{(\eta_{I\mathrm{th}}\tau_{31}\tau_{32} - \tau_2\tau_{31} - (1-\eta_{I\mathrm{th}})\tau_2\tau_{32})}{(\tau_2 + \tau_{31})\tau_{32}} \tag{2.60}$$

Above threshold, $n_3 - n_2$ does not change from its value at threshold. However, the populations of the individual subbands increase linearly with current as shown in Fig. 2-6 for some typical parameter values corresponding to those for terahertz QCLs. This increase is due to inefficient extraction of electrons from the lower subband (due to a non-zero lower subband lifetime τ_2), which does not allow all the extra current being pumped into the modules beyond threshold to be converted into photons, thus reducing the optical efficiency of the laser. Note that n_2-I curves are piecewise linear with two different slopes. The slope above I_{th} is higher due to stimulated emission transitions from 3 to 2. The power out-coupled from the laser becomes

$$\begin{aligned} P_{\mathrm{out}} &= \hbar\omega_0\, n_{\mathrm{ph}}^{\mathrm{st}}\, \alpha_{\mathrm{m}}\, \frac{c}{n_{\mathrm{r}}} \\ &= \frac{N_{\mathrm{p}}\hbar\omega_0}{|e|}\frac{\alpha_{\mathrm{m}}}{(\alpha_{\mathrm{w}} + \alpha_{\mathrm{m}})}\left[(I - I_{\mathrm{th}})\chi + I(\eta_I - \eta_{I\mathrm{th}})\right] \end{aligned} \tag{2.61}$$

Correspondingly, the *slope efficiency* dP_{out}/dI becomes

$$\frac{dP_{\text{out}}}{dI} = \frac{N_{\text{p}}\hbar\omega_0}{|e|}\frac{\alpha_{\text{m}}}{(\alpha_{\text{w}}+\alpha_{\text{m}})}\chi' \tag{2.62}$$

where,

$$\chi' = \chi + (\eta_I - \eta_{I\text{th}}) + I\frac{d\eta_I}{dI} \tag{2.63}$$

The slope efficiency can show discontinuities in the P-I (also known as L-I) characteristics of the lasers, which have multiple injector levels, since the *injection efficiency* η can change significantly at some particular bias resulting in a large $d\eta_I/dI$. The quantity χ' ($\approx \chi$ at most bias conditions) is known as the *internal quantum efficiency*, and a rough estimate of its value can be obtained from the I-V characteristics of a laser. χ' is lowered from its ideal value of unity due to a non-unity injection efficiency ($\eta < 1$), and a non-zero lower level lifetime ($\tau_2 > 0$). $\chi' N_{\text{p}}\alpha_{\text{m}}/(\alpha_{\text{w}}+\alpha_{\text{m}})$ determines the *differential quantum efficiency* in number of photons generated per electron, which for the best terahertz QCLs is reported to be $\sim$ 10% of the total value of N_{p} [175]. This quantity is largely dependent on the type of waveguide used, due to a big difference in the *out-coupling efficiency* $\alpha_{\text{m}}/(\alpha_{\text{w}}+\alpha_{\text{m}})$ of metal-metal and semi-insulating surface-plasmon waveguides.

Terahertz QCLs are typically made with a large number of periods ($N_{\text{p}} > 100$). For the metal-metal waveguides (section 3.6.1), a large N_{p} reduces the fractional overlap of the desired waveguide mode with the lossy cladding layers in comparison to that of the active region, thereby reducing α_{w}, and hence the threshold gain requirement (equation 2.58). For the semi-insulating surface-plasmon waveguides (section 3.6.1), a large N_{p} increases the mode confinement factor Γ, which also effectively reduces the threshold gain requirement (equation 2.58). Moreover, for a large N_{p}, the slope efficiency is higher because it is proportional to N_{p} and also because α_{w} is lowered (equation 2.62). However, increasing N_{p} comes at the cost of a higher voltage that is required to bias the N_{p} modules, which causes increased electrical power dissipation in the active region affecting continuous-wave performance of the laser due to lattice heating. The *wall-plug efficiency*, which is the fraction of electrical power that

is converted into optical power, is a combination of various factors that include the slope efficiency, the value of the threshold current $I_{\rm th}$, the range of the current above threshold, and also the *voltage-efficiency* in the design ($\hbar\omega_0/(E_3 - E_1)$ in Fig. 2-4). The best 5 K wall-plug efficiencies measured so far for terahertz QCLs are $\sim$ 0.5% in cw operation [12, 175] (see also section 3.6.2, Fig. 3-50), and $\sim$ 4% in pulsed operation [13].

All the experimental values mentioned in the preceding paragraphs were for terahertz QCLs operating in continuous-wave mode at liquid helium temperatures. These results have been obtained from QCLs developed during the course of this work, the experimental results of which will be discussed in greater detail in chapters 3 and 4.

2.3.4 Small-signal modulation response above threshold

The small-signal current modulation response of a QCL can be characterized by the time-harmonic steady-state response. This is done by assuming solutions of the form

$$\begin{aligned} I(t) &= I + \mathcal{R}\mathrm{e}(\tilde{I}e^{i\Omega t}) \\ n_3(t) &= n_3 + \mathcal{R}\mathrm{e}(\tilde{n}_3 e^{i\Omega t}) \\ n_2(t) &= n_2 + \mathcal{R}\mathrm{e}(\tilde{n}_2 e^{i\Omega t}) \\ n_{\rm ph}^{\rm st}(t) &= n_{\rm ph}^{\rm st} + \mathcal{R}\mathrm{e}(\tilde{n}_{\rm ph}^{\rm st} e^{i\Omega t}) \end{aligned} \tag{2.64}$$

Assuming amplitudes of the time varying components to be much smaller than the corresponding zero-frequency values, the cross terms for the small-signal components could be neglected in the rate equations, which then become a system of linear algebraic equations. In steady state operation above threshold, $\tilde{n}_{\rm ph}^{\rm st}$ becomes

$$\left.\frac{\tilde{n}_{\rm ph}^{\rm st}}{\tilde{I}}\right|_{\tau_2 \to 0} = \frac{\eta_I N_{\rm p}}{|e|}\tau_{\rm ph}\left(\frac{\Omega_{\rm R}^2}{\Omega_{\rm R}^2 - \Omega^2 + i\Omega\gamma}\right) \tag{2.65}$$

where Ω_R^2 and γ are given by

$$\begin{aligned}
\Omega_R^2 &= \left(\frac{\eta_I I}{\eta_{I\mathrm{th}} I_{\mathrm{th}}} - 1\right) \frac{1}{\tau_{\mathrm{ph}} \tau_3} \\
\gamma &= \left(\frac{\eta_I I}{\eta_{I\mathrm{th}} I_{\mathrm{th}}}\right) \frac{1}{\tau_3} \\
\tau_3 &\equiv \frac{\tau_{31} \tau_{32}}{(\tau_{31} + \tau_{32})}
\end{aligned} \tag{2.66}$$

Equation (2.65) describes a two-pole response of the optical power (since optical power $\propto n_{\mathrm{ph}}^{\mathrm{st}}(t)$) for a harmonic current modulation in a QCL, which is similar to that for interband diode lasers. The parameter $(2\Omega_R)/\gamma$ characterizes the oscillatory behavior of the response, and can be expressed as

$$\frac{2\Omega_R}{\gamma} = \sqrt{\left(\frac{\eta_{I\mathrm{th}} I_{\mathrm{th}}}{\eta_I I}\right)\left(1 - \frac{\eta_{I\mathrm{th}} I_{\mathrm{th}}}{\eta_I I}\right) \frac{4\tau_3}{\tau_{\mathrm{ph}}}} \tag{2.67}$$

The frequency response is overdamped if $2\Omega_R/\gamma < 1$. The maximum value of the expression $2\Omega_R/\gamma$ is obtained for $\eta_I I = 2\eta_{I\mathrm{th}} I_{\mathrm{th}}$ when $(2\Omega_R/\gamma)_{\max} = \sqrt{\tau_3/\tau_{\mathrm{ph}}}$. In either a terahertz or a mid-infrared QCL, typically τ_3 is of the order of 1 ps, whereas $\tau_{\mathrm{ph}} \sim 1 - 10$ ps, for which $2\Omega_R/\gamma < 1$ is always satisfied. Hence, unlike in interband diode lasers, the fast non-radiative scattering times in a QCL do not allow occurrence of relaxation oscillations in its modulation response. The 3-dB frequency, defined as the frequency where the square of the magnitude of the frequency response goes to half of the zero frequency value, becomes

$$\begin{aligned}
\Omega_{3\mathrm{dB}} &= \sqrt{\sqrt{\left(\frac{\gamma^2}{2} - \Omega_R^2\right)^2 + \Omega_R^4} - \left(\frac{\gamma^2}{2} - \Omega_R^2\right)} \\
&\underset{(2\Omega_R/\gamma) \ll 1}{\approx} \frac{\Omega_R^2}{\gamma} \\
&= \left(1 - \frac{\eta_{I\mathrm{th}} I_{\mathrm{th}}}{\eta_I I}\right) \frac{1}{\tau_{\mathrm{ph}}}
\end{aligned} \tag{2.68}$$

Hence, for QCLs the maximum 3-dB modulation frequency, $f_{3\mathrm{dB}}$ should be limited by the photon lifetime inside the cavity limiting it to the range of $10 - 100$ GHz. The

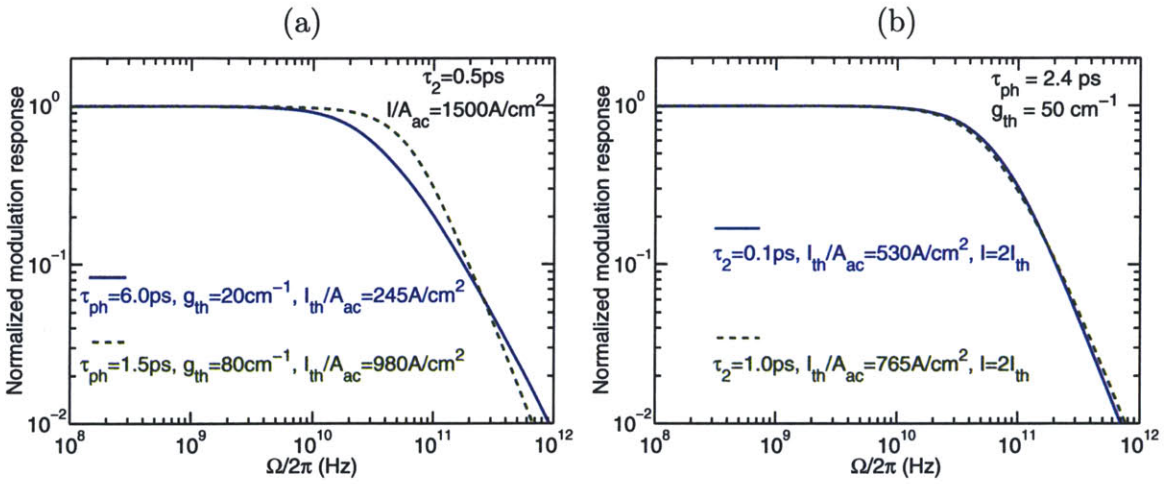


Figure 2-7: Normalized small signal modulation response $\frac{\tilde{n}_{\rm ph}^{\rm st}(\Omega)/\tilde{I}(\Omega)}{\tilde{n}_{\rm ph}^{\rm st}(0)/\tilde{I}(0)}$ plotted for the following parameters: $\eta_{I\rm th} = \eta_I = 1$, $L_{\rm mod} = 50$ nm, $\tau_{32} = 3$ ps, $\tau_{31} = 7$ ps, $f_{23} = 0.8$, and $\Delta\nu = 1.5$ THz, where the threshold gain $g_{\rm th} \equiv (n_{\rm r}/c\tau_{\rm ph})$ is related to the population inversion $\Delta n_{\rm th} \equiv (n_3^{\rm th} - n_2^{\rm th})$ by equation (2.33). $A_{\rm ac}$ is the cross-sectional area of the active region, which can be chosen arbitrarily once $L_{\rm mod}$ is fixed. $\Omega_{\rm 3dB}$ is defined as the frequency where the normalized modulation response becomes $1/\sqrt{2}$.

normalized small-signal modulation response for some typical parameters corresponding to those for terahertz QCLs is plotted in Fig. 2-7. The conclusions derived above agree with a similar calculation done for mid-infrared QCLs in Ref. [131], and puts QCLs on the same footing as interband diode lasers with regard to the maximum frequency of their modulation response.

The results derived here for the maximum modulation response are contrary to some of the claims in literature, which have suggested that even faster modulation response should be possible in QCLs [34, 125], possibly reaching even THz modulation bandwidths [120, 52]. In Refs. [52, 34, 125], the following standard expression, as derived in Yariv [187] for a vanishing lower state lifetime, is used

$$\left.\frac{\tilde{n}_{\rm ph}^{\rm st}}{\tilde{I}}\right|_{\tau_2 \to 0} \propto \left(\frac{\Omega_{\rm R}^2}{\Omega_{\rm R}^2 - \Omega^2 + i\Omega\gamma}\right) \tag{2.69}$$

where Ω_R^2 and γ are expressed in terms of τ_{st} instead of I as

$$\begin{aligned} \Omega_R^2 &= \frac{1}{\tau_{ph}\tau_{st}} \\ \gamma &= \frac{1}{\tau_3} + \frac{1}{\tau_{st}} \end{aligned} \tag{2.70}$$

The characteristic frequencies for the modulation response in equation (2.69) become

$$\Omega_{\pm} = i\frac{1}{2}\left(\frac{1}{\tau_3} + \frac{1}{\tau_{st}}\right) \pm \sqrt{\frac{1}{\tau_{ph}\tau_{st}} - \frac{1}{4}\left(\frac{1}{\tau_3} + \frac{1}{\tau_{st}}\right)^2} \tag{2.71}$$

The expressions in equation (2.70) correspond exactly to the ones derived in equation (2.66) (for $\tau_2 \rightarrow 0$). Typically τ_{st} is much smaller than τ_3, and since τ_3 is of the same order as τ_{ph}, the characteristic frequencies in the expression above will have no real component implying absence of relaxation oscillations. In the limit $\tau_{st} \ll (\tau_{ph}, \tau_3)$

$$\Omega_{\pm}|_{\tau_{st} \ll (\tau_{ph}, \tau_3)} = \begin{cases} i\dfrac{1}{\tau_{ph}} \\ i\dfrac{1}{\tau_{st}} \end{cases} \tag{2.72}$$

which is same as the result derived in equation (2.68) for $I \gg I_{th}$, implying that τ_{ph} limits the 3dB bandwidth. However, in Refs. [52, 34, 125], this limit is incorrectly calculated to be i/τ_3, which allowed for a conclusion that modulation bandwidth could be increased further from $1/\tau_{ph}$ by reducing the upper state lifetime. For the sake of a physical argument, an upper limit for the fastest modulation response should be obtainable by estimating the limiting mechanism for the photon density modulation. Although τ_3 may be made smaller than τ_{ph}, it seems plausible that the limiting decay rate for the photon density would still be the photon lifetime τ_{ph}, which should thus determine the modulation bandwidth.

2.4 Resonant-tunneling through a potential barrier

2.4.1 Density matrix formalism

The Fermi's golden rule (equation 2.18) for determining the probability of state transitions gives no information about phase relation between the initial and the final states. This is insufficient to describe "coherent" phenomena, in which phase of the interacting states plays an important role. The density matrix formalism is an elegant way to include the phase information while determining the time evolution of quantum mechanical states. Most importantly, compared to the wave function approach, it allows a phenomenological description of a statistical ensemble of many states, such as in a multi-electron system.

Suppose that there is a state (Hilbert) space, with a denumerable orthonormal basis $|\phi_n\rangle$, where n is an integer. If the system is in state $|\psi(t)\rangle$ at time t, the following is the expansion in this basis

$$|\psi(t)\rangle = \sum_n a_n(t)|\phi_n\rangle \tag{2.73}$$

For an observable (self-adjoint operator) $\hat{O}$, the matrix elements of $\hat{O}$ in this basis can be written as

$$O_{mn} = \langle\phi_m|\hat{O}|\phi_n\rangle \tag{2.74}$$

In this representation, the operator $\hat{O}$ can also be written as a Hermitian matrix $\bar{\bar{O}}$ with the above mentioned elements. The average (expectation) value of $\hat{O}$ at time t, for the system in state $|\psi(t)\rangle$ becomes

$$\langle\hat{O}\rangle(t) = \langle\psi(t)|\hat{O}|\psi(t)\rangle = \sum_n \sum_m a_m^*(t)a_n(t)O_{mn} \tag{2.75}$$

Consider a self-adjoint (Hermitian) operator, $\hat{\rho}(t)$, which is defined as

$$\hat{\rho}(t) \equiv |\psi(t)\rangle\langle\psi(t)| \tag{2.76}$$

This is called the *density operator*. The elements of the *density matrix* $\bar{\bar{\rho}}(t)$ become

$$\rho_{mn}(t) = \langle \phi_m | \hat{\rho}(t) | \phi_n \rangle = a_m(t) a_n^*(t) \tag{2.77}$$

The expectation value of the observable $\hat{O}$ can then be written in terms of $\bar{\bar{\rho}}(t)$ as

$$\begin{aligned} \langle \hat{O} \rangle(t) &= \sum_n \sum_m \rho_{nm}(t) O_{mn} \\ &= \sum_n \left[\bar{\bar{\rho}}(t) \bar{\bar{O}} \right]_{nn} \\ &= \mathrm{Tr} \left[\bar{\bar{\rho}}(t) \bar{\bar{O}} \right] \end{aligned} \tag{2.78}$$

The time evolution of a state is given by the Schrödinger equation

$$\hat{H}(t) | \psi(t) \rangle = i\hbar \frac{d}{dt} | \psi(t) \rangle \tag{2.79}$$

Correspondingly, time evolution of the density operator (and the matrix) becomes

$$\begin{aligned} \frac{d}{dt} \hat{\rho}(t) &= -\frac{i}{\hbar} \left[\hat{H}(t), \hat{\rho}(t) \right] \\ \frac{d}{dt} \bar{\bar{\rho}}(t) &= -\frac{i}{\hbar} \left[\bar{\bar{H}}(t), \bar{\bar{\rho}}(t) \right] \end{aligned} \tag{2.80}$$

The density operator, unlike the state vector, has no phase ambiguity and it remains the same under a phase transformation $|\psi(t)\rangle \rightarrow e^{i\theta} |\psi(t)\rangle$. Moreover, the expectation values of observables are quadratic in $|\psi(t)\rangle$, but only linear in $\hat{\rho}(t)$, where $\hat{\rho}^2(t) = \hat{\rho}(t)$. It may be noted that the diagonal terms of the density matrix $\rho_{nn}(t) = |a_n(t)|^2$ are real and positive, whereas the off-diagonal terms can be complex. The real advantage of the density matrix formalism is in the fact that the matrix element $\rho_{mn}(t)$ can be defined as an ensemble average [186, 119]

$$\rho_{mn}(t) \equiv \overline{a_m(t) a_n^*(t)} \tag{2.81}$$

for a statistical mixture of state vectors, such as in a multi-electron system, where each

electron is described by its "independent" state vector within a statistical distribution. The off-diagonal elements in this case are called *coherences* since their behavior with time depends on how well the individual state vectors maintain their phase. The diagonal elements determine the probability of finding a random state-vector in the ensemble in a particular basis state, and therefore, the *population* of that basis state.

2.4.2 Current transport in the QCL with 3-levels per module

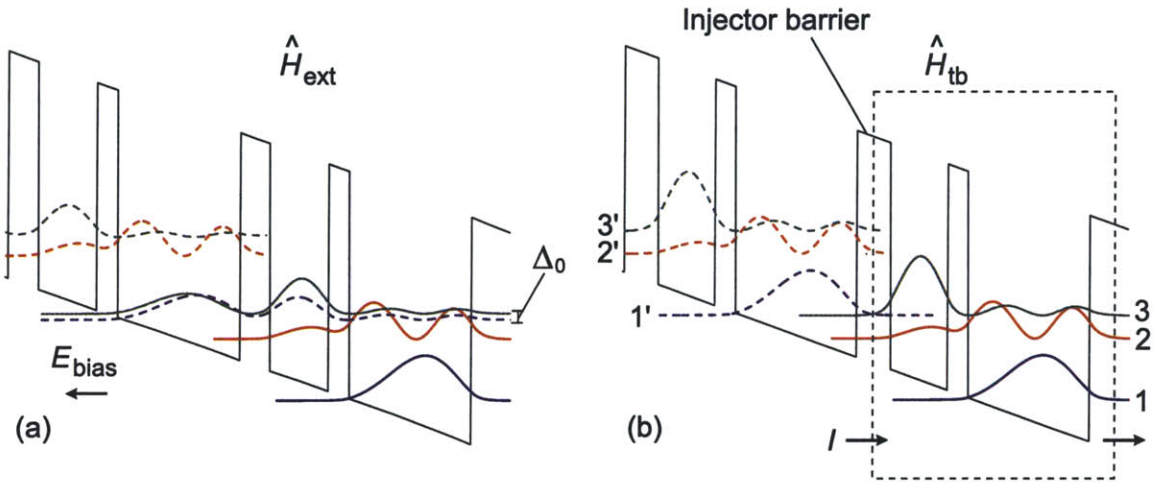


Figure 2-8: Plot of the magnitude squared envelope wavefunctions for the 3-level QCL of Fig. 2-1 with two different sets of basis functions. Plot (a) is for the "extended" scheme, where energy splitting due to the injector anticrossing Δ_0 is visible, and the wavefunctions are calculated for a potential profile $\hat{H}_{ext}$ as it appears in the figure. Plot (b) is for the "tight-binding" scheme, where $\hat{H}_{tb}$ is formed by making the potential barriers at the boundaries of a module infinitely thick to confine the wavefunctions within that single module.

The density matrix formalism can provide an elegant way to include the collisional broadening, as well as lifetime broadening of the states of a Hamiltonian in the presence of a perturbation, such as the light-matter interaction, for a many-body system [186]. Analogously, it can be used to estimate current transport in a multiple quantum-well structure due to resonant-tunneling through a potential barrier [166, 29]. For the best performing QCL designs, both in the mid-infrared and in the terahertz, the *injector barrier* is intentionally set as the bottleneck for current transport so that a high injection efficiency can be maintained and the current

transport can be limited at an applied field below the design bias. In this section, the current flow due to resonant-tunneling through the injector barrier in the 3-levels per module QCL design of Figs. 2-1 and 2-4 will be calculated using density matrices. This will allow determination of current I, which was used as an independent parameter in the rate-equation description of such a QCL in section 2.3.

Figure 2-8(a) shows the conduction band diagram for the aforementioned design at *design bias*, when the levels $1'$ and 3 are lined up in energy for maximum tunneling current through the injector barrier. The wavefunctions are calculated with the full Hamiltonian $\hat{H}_{\text{ext}}$. At $1'-3$ resonance, both subbands are extended on either side of the injector barrier with an energy splitting Δ_0, which known as the *anticrossing gap*. The wavefunctions are calculated again in a tight-binding scheme with a modified Hamiltonian $\hat{H}_{\text{tb}}$, as shown in Fig. 2-8(b). $\hat{H}_{\text{tb}}$ is modified from the original Hamiltonian $\hat{H}_{\text{ext}}$ to keep the wavefunctions confined within a module. Whereas the wavefunctions for $\hat{H}_{\text{ext}}$ depend sensitively on the external static field E_{bias} near the design bias, the wavefunctions for $\hat{H}_{\text{tb}}$ do not change significantly, which keeps the form-factors for various scattering rate calculations in between those wavefunctions approximately independent of bias.

The tunneling mechanism is critically dependent on the energy broadening of the levels since it is a resonance phenomenon. This broadening is caused by both lifetime broadening and collisional broadening (dephasing), and is not easy to incorporate in the scheme of Fig. 2-8(a) due to the sensitive dependence of the extended wavefunctions on E_{bias}. However, with the tight-binding wavefunctions as the basis states, density-matrix formalism that allows a phenomenological treatment of level broadening, is easier to implement. As will be shown next, Δ_0 plays a critical role in determining the current flow at resonance due to the effect of level broadening [29].

The density matrix for the 3-level problem at hand is a 3×3 matrix. To keep the analysis simple, the current flow close to design bias will be assumed entirely due to $1'-3$ tunneling. This is analogous to assuming a unity injection efficiency ($\eta = 1$). Within this approximation, the time evolution of the density matrix with the basis

states of Fig. 2-8(b) could be written as in equation (2.80)

$$\frac{d}{dt}\bar{\bar{\rho}}_{(3\times3)} = -\frac{i}{\hbar}\left[\bar{\bar{H}}_{\text{ext}}, \bar{\bar{\rho}}_{(3\times3)}\right] - \frac{i}{\hbar}\left[\bar{\bar{H}}', \bar{\bar{\rho}}_{(3\times3)}\right] \tag{2.82}$$

where, $\bar{\bar{H}}'$ includes radiative and non-radiative perturbation terms. The expression above can be expanded as

$$\begin{aligned}\frac{d}{dt}\begin{pmatrix}\rho_{1'1'} & \rho_{1'2} & \rho_{1'3}\\ \rho_{21'} & \rho_{22} & \rho_{23}\\ \rho_{31'} & \rho_{32} & \rho_{33}\end{pmatrix} &= -\frac{i}{\hbar}\left[\begin{pmatrix}E_{1'} & 0 & \frac{-\Delta_0}{2}\\ 0 & E_2 & 0\\ \frac{-\Delta_0}{2} & 0 & E_3\end{pmatrix}, \begin{pmatrix}\rho_{1'1'} & \rho_{1'2} & \rho_{1'3}\\ \rho_{21'} & \rho_{22} & \rho_{23}\\ \rho_{31'} & \rho_{32} & \rho_{33}\end{pmatrix}\right]\\ &+ \begin{pmatrix}\frac{\rho_{33}}{\tau_{31}} + \frac{\rho_{22}}{\tau_2} & -- & \frac{-\rho_{1'3}}{\tau_{\|}}\\ -- & \frac{\rho_{33}}{\tau_{32}} + \frac{\rho_{33}-\rho_{22}}{\tau_{\text{st}}} - \frac{\rho_{22}}{\tau_2} & --\\ \frac{-\rho_{31'}}{\tau_{\|}} & -- & -\frac{\rho_{33}}{\tau_3} - \frac{\rho_{33}-\rho_{22}}{\tau_{\text{st}}}\end{pmatrix}\end{aligned} \tag{2.83}$$

where, $\langle n|\hat{H}_{\text{ext}}|n\rangle \approx \langle n|\hat{H}_{\text{tb}}|n\rangle = E_n$, $\langle 1'|\hat{H}_{\text{ext}}|3\rangle = \langle 3|\hat{H}_{\text{ext}}|1'\rangle \approx -\Delta_0/2$, $\tau_{\text{st}} = \tau'_{\text{sp}}/n^{\text{st}}_{\text{ph}}$ ($\gg (\tau_3, \tau_2)$ near and below lasing threshold), and $\tau_3 \equiv \tau_{31}\tau_{32}/(\tau_{31}+\tau_{32})$. The interaction with the optical field is included through τ_{st}, and the photon rate equation, which determines the condition of lasing threshold, is assumed to be solved separately. As described above, the diagonal terms of the density matrix are equivalent to the population of the basis states, and the off-diagonal terms describe the coherences. For this system, the coherences corresponding to level 2 (indicated by $--$) do not affect the population dynamics and hence need not be included [29]. $\tau_{\|}$ is the dephasing time, which determines relaxation (damping) of the coherent interaction between level $1'$ and level 3. $\tau_{\|}$ includes terms due to both lifetime broadening and collisional broadening as [29]

$$\frac{1}{\tau_{\|}} = \frac{1}{2\tau_1} + \frac{1}{2\tau_3} + \frac{1}{T_2^*} \approx \frac{1}{2\tau_3} + \frac{1}{T_2^*} \tag{2.84}$$

The time T_2^* is the *pure-dephasing* time and is due to the destruction of phase coherences between states that results from various scattering mechanisms (collisions) such as interface-roughness (which reduces the specularity of tunneling through the barriers), impurity scattering, and electron-electron scattering. Pure-dephasing is

exclusive of the level lifetime τ_3, which characterizes lifetime broadening due to the uncertainty relation between the energy of a level and its lifetime. Lifetime broadening can be explained by the time-dependent perturbation theory of which the Fermi's golden rule is a special case. The factor of 2 in the lifetime contribution to dephasing is because population relaxation applies to the square of the amplitude of wavefunctions whereas dephasing applies to the amplitudes directly [166].

The finite off-diagonal terms in $\bar{\bar{H}}_{\text{ext}}$ corresponding to levels $1'$ and 3 are indicative of coherent interaction in between those basis states. In other words, the phase coherence of those levels is linked to their populations. Hence, the dephasing time $\tau_{||}$ is included as a relaxation term for their corresponding off-diagonal elements in equation (2.83).

Equation (2.83) is similar in many ways to the rate equations analyzed in sections 2.3. Additionally, it describes the resonant-tunneling process also, thus allowing a calculation of the current flowing in the structure rather than it being taken as a source parameter. The rate equations accurately describe the population dynamics of level 2 since it does not have coherent interaction with the other two levels close to the design bias (i. e. the bias for $1' - 3$ alignment). Hence, the results derived from rate equations, which specify the population of level 2 ($= \rho_{22}$) in terms of that of level 3 ($= \rho_{33}$), could be used to convert equation (2.83) to an equation with a 2×2 density matrix. The simplified equation becomes

$$\frac{d}{dt}\bar{\bar{\rho}}_{1',3} = -\frac{i}{\hbar}\left[\begin{pmatrix} E_{1'} & \frac{-\Delta_0}{2} \\ \frac{-\Delta_0}{2} & E_3 \end{pmatrix}, \bar{\bar{\rho}}_{1',3}\right] + \begin{pmatrix} \frac{\rho_{33}}{\tau_{31}} + \frac{\rho_{22}}{\tau_2} & \frac{-\rho_{1'3}}{\tau_{||}} \\ \frac{-\rho_{31'}}{\tau_{||}} & -\frac{\rho_{33}}{\tau_{31}} - \frac{\rho_{22}}{\tau_2} \end{pmatrix} \tag{2.85}$$

where, $\bar{\bar{\rho}}_{1',3} \equiv \begin{pmatrix} \rho_{1'1'} & \rho_{1'3} \\ \rho_{31'} & \rho_{33} \end{pmatrix}$. Note that $\frac{\rho_{33}}{\tau_3} + \frac{\rho_{33}-\rho_{22}}{\tau_{\text{st}}}$ from equation (2.83) is replaced by $\frac{\rho_{33}}{\tau_{31}} + \frac{\rho_{22}}{\tau_2}$ since these terms are equivalent (both below and above threshold $\frac{\rho_{22}}{\tau_2} = \frac{\rho_{33}}{\tau_{32}} + \frac{\rho_{33}-\rho_{22}}{\tau_{\text{st}}}$), as a statement of current continuity. In the steady-state ($d/dt \rightarrow 0$), the following solutions can be used for ρ_{22}, as deduced from the results of section 2.3

for $\eta = 1$

$$\rho_{22} = \begin{cases} \rho_{33}\dfrac{\tau_2}{\tau_{32}} & \ldots (I < I_{\text{th}}) \\ \rho_{33} - \Delta n_{\text{th}} & \ldots (I \geq I_{\text{th}}) \end{cases} \tag{2.86}$$

where, $\Delta n_{\text{th}} = \rho_{33} - \rho_{22}$ is the population inversion, which is pinned to a constant value above threshold, as derived in equation (2.58). Equation (2.85) can be solved analytically in the steady-state using the expressions for ρ_{22} from equation (2.86). Solving for the current I (with the constraint $(\rho_{11} + \rho_{22} + \rho_{33}) = n_{\text{tot}}$) the following expressions are obtained

$$I \equiv \frac{|e|}{N_{\text{p}}}\left(\frac{\rho_{33}}{\tau_{31}} + \frac{\rho_{22}}{\tau_2}\right) \tag{2.87}$$

$$= \begin{cases} \dfrac{n_{\text{tot}}|e|}{N_{\text{p}}}\left[\dfrac{\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}(\tau_{31}+\tau_{32})}{\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\tau_{31}(2\tau_{32}+\tau_2) + 2\left(\frac{E_{1'3}}{\hbar}\right)^2 \tau_{\parallel}^2(\tau_{31}+\tau_{32}) + 2(\tau_{31}+\tau_{32})}\right] & \ldots (I < I_{\text{th}}) \\ \dfrac{n_{\text{tot}}|e|}{N_{\text{p}}}\left[\dfrac{\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\left((\tau_{31}+\tau_2) - \frac{\Delta n_{\text{th}}}{n_{\text{tot}}}(2\tau_{31}-\tau_2)\right)}{3\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\tau_{31}\tau_2 + 2\left(\frac{E_{1'3}}{\hbar}\right)^2 \tau_{\parallel}^2(\tau_{31}+\tau_2) + 2(\tau_{31}+\tau_2)}\right] & \ldots (I \geq I_{\text{th}}) \end{cases}$$

The $I - E_{1'3}$ expressions derived above have a Lorentzian lineshape, and the maximum current flows through the structure at resonance ($E_{1'3} \equiv E_{1'} - E_3 = 0$).

Figure 2-9 shows calculated $I - E_{1'3}$ curves for typical parameters corresponding to terahertz QCLs. A wealth of information can be obtained from these plots; however, further discussions about these results are deferred to chapter 3, where results from various terahertz QCL active regions will be analyzed. It may be noted that the current I calculated in these plots is due to $1' \rightarrow 3$ tunneling only, where $1'$ is the injector level and 3 is the upper radiative level. In reality, various *parasitic* current channels also conduct current due to resonant interaction of the injector level with levels other than the upper radiative level. In that case, the threshold current I_{th} gets artificially enhanced from the calculated values in Fig. 2-9 and thus does not accurately correspond to the $1' \rightarrow 3$ current required to meet the gain threshold g_{th}. These issues will be discussed in detail in chapter 3 (section 3.2).

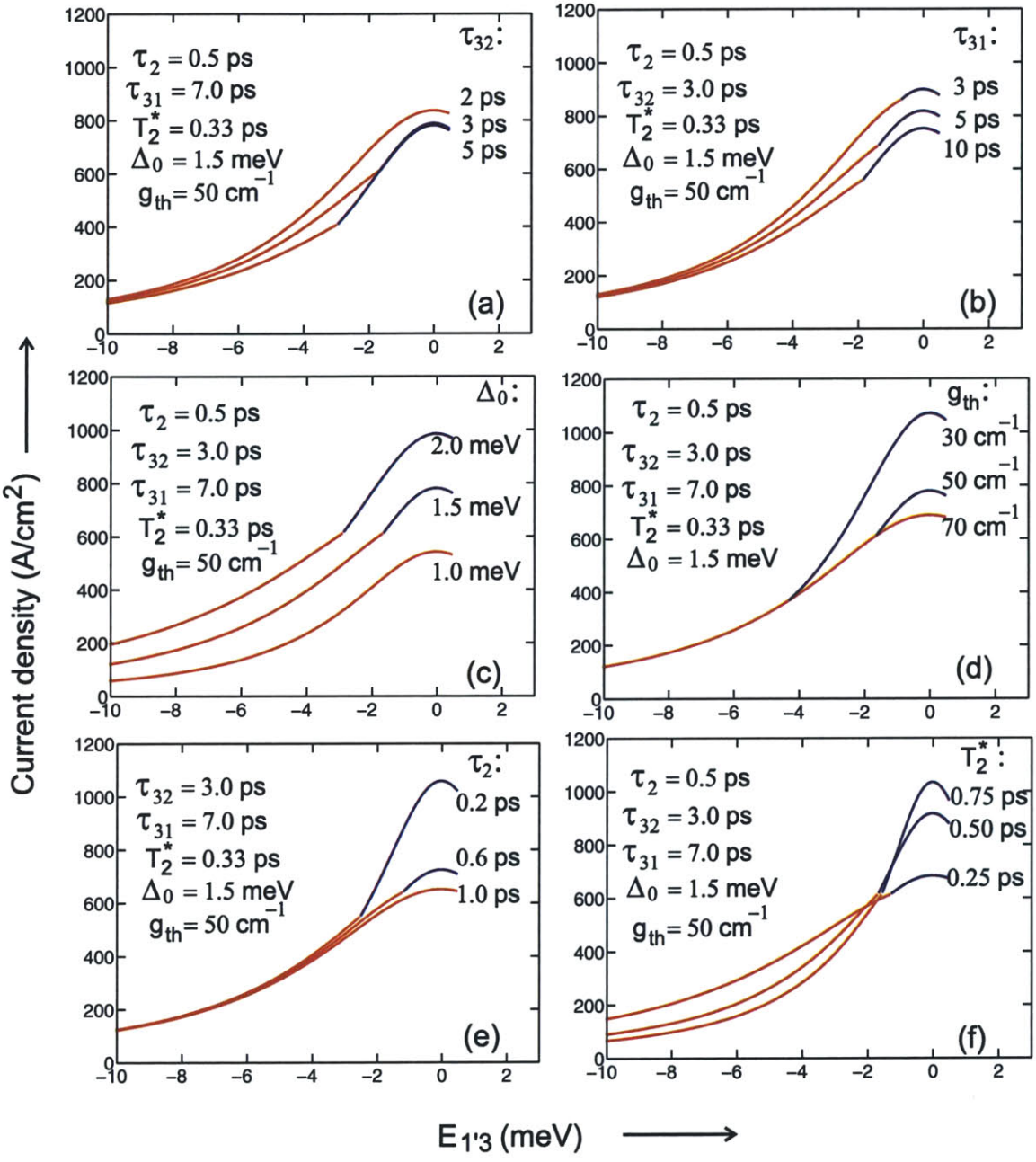


Figure 2-9: Steady state current density (I/A_{ac}) versus bias plots for the 3-level QCL, calculated using density-matrices and rate-equations for a range of parameters, typically corresponding to those for terahertz QCLs. A doping density $n_{\text{tot}}/(N_{\text{p}}L_{\text{mod}}A_{\text{ac}}) = 5 \times 10^{15}$ cm^{-3} is used, where A_{ac} is the cross-sectional area of the active region and is arbitrarily chosen. The expression for peak gain in GaAs, as derived in equation (2.33), is used to relate population inversion Δn_{th} to the threshold gain g_{th}, with an oscillator strength $f_{23} = 0.8$, Lorentzian linewidth $\Delta\nu = 1.5$ THz, and a mode-confinement factor Γ of unity. The red region of the curves corresponds to $I < I_{\text{th}}$, while blue corresponds to $I > I_{\text{th}}$.

It is instructive to rewrite equation (2.87) for the case of ultra-short lifetime of the lower radiative state ($\tau_2 \rightarrow 0$). The expressions then become

$$I|_{(\tau_2 \rightarrow 0)} = \begin{cases} \dfrac{n_{\text{tot}}|e|}{2N_{\text{p}}} \left(\dfrac{1}{\tau_3}\right) \left[\dfrac{\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\|} \tau_3}{\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\|} \tau_3 + \left(\frac{E_{1'3}}{\hbar}\right)^2 \tau_{\|}^2 + 1} \right] & \ldots (I < I_{\text{th}}) \quad \text{(a)} \\ \dfrac{n_{\text{tot}}|e|}{2N_{\text{p}}} \left(\dfrac{1}{\tau_3}\right) \left[\dfrac{\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\|} \tau_3 \left(1 - \frac{2\Delta n_{\text{th}}}{n_{\text{tot}}}\right)}{\left(\frac{E_{1'3}}{\hbar}\right)^2 \tau_{\|}^2 + 1} \right] & \ldots (I \geq I_{\text{th}}) \quad \text{(b)} \end{cases} \tag{2.88}$$

Equation (2.88)(a) is the same as the one derived by Kazarinov and Suris [83] for their superlattice scheme shown in Fig. 1-3, since for $\tau_2 = 0$ ($\Rightarrow \rho_{22} = 0$), the superlattice with 3-levels per module essentially becomes equivalent to a superlattice with 2-levels per module. It was later used by Sirtori *et al.* [148] to describe the resonant-tunneling phenomena in mid-infrared quantum-cascade lasers. More recently, a time-domain Monte-Carlo simulation of resonant-tunneling transport through a potential barrier was demonstrated by Callebaut *et al.* [29], which was likened to the *optical Rabi oscillations* caused by the electromagnetic coupling between the two electronic states. While $\Omega = (\Delta_0/\hbar)$ is the frequency of oscillation of the electron wavepacket across the barrier (analogous to the optical *Rabi frequency*), $\tau_{\|}^{-1}$ corresponds to the damping rate of the oscillation. Similar to the lineshape of optical emission, the $I - E_{1'3}$ curves are Lorentzians with the FWHM linewidths given by

$$\Delta E_{1'3,\text{FWHM}}|_{\tau_2 \rightarrow 0} = \begin{cases} \dfrac{2\hbar\sqrt{\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\|} \tau_3 + 1}}{\tau_{\|}} & \ldots (I < I_{\text{th}}) \\ \dfrac{2\hbar}{\tau_{\|}} & \ldots (I \geq I_{\text{th}}) \end{cases} \tag{2.89}$$

Typically, several current channels are active at any given bias in a QCL, hence the shape of the experimental I-V curves cannot be used to estimate $\tau_{\|}$, and hence the pure dephasing time T_2^*. As will be shown in section 2.5, the spontaneous emission linewidth is also determined from T_2^* and can provide a better estimate since it can be directly measured in an experiment.

Discontinuity in differential resistance at the lasing threshold

A discussion about the discontinuity in the slope of the $I - E_{1'3}$ curves at the lasing threshold is in order. The discontinuity is a manifest of the fact that the population inversion gets pinned to a constant value above threshold, and the upper state lifetime starts to decrease due to stimulated emission. The change in slope of the $I - E_{1'3}$ curve at the discontinuity can be experimentally determined from the differential resistance-voltage (dV/dI-V) characteristics of a device, and is shown below to be a direct probe of the population inversion $\Delta n_{\rm th}$ in the laser. The voltage applied across the device $V_{\rm bias}$ is related to $E_{1'3}$ by

$$V_{\rm bias} = \frac{E_{1'3} N_{\rm p}}{f_{\rm V} |e|} \tag{2.90}$$

where, $f_{\rm V}$ is the fraction of the applied voltage that appears as the energy difference between levels $1'$ and 3. $f_{\rm V}$ is typically a non-linear function of the applied voltage V and is a characteristic of a particular design. For the main result to be derived in this section, the particular form of $f_{\rm V}$ is inconsequential. Using equation (2.90) the differential resistance becomes

$$\mathcal{R} = \frac{dV_{\rm bias}}{dI} = \frac{1}{f_{\rm V}} \frac{N_{\rm p}}{|e|} \frac{dE_{1'3}}{dI} \tag{2.91}$$

To determine the value of $\mathcal{R}$ just below and above threshold, an expression for $E_{1'3,\rm th}$ can be derived, for which the population inversion condition $n_3 - n_2 = \Delta n_{\rm th}$ is satisfied. From the steady state solution of the density matrix in equation (2.85), the following expression is obtained

$$E_{1'3,\rm th} = -\frac{\hbar}{\tau_{\|}} \sqrt{\frac{\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\|} \tau_3 \left[n_{\rm tot} - 2\Delta n_{\rm th} - (n_{\rm tot} + \Delta n_{\rm th})\tau_2/\tau_{32}\right] - 2\Delta n_{\rm th}}{2\Delta n_{\rm th}}} \tag{2.92}$$

Substituting $E_{1'3} = E_{1'3,\rm th}$ in $dE_{1'3}/dI$ obtained from the $I - E_{1'3}$ expressions in

equation (2.87), the following expression is obtained for $\mathcal{R}$ just below threshold

$$\mathcal{R}|_{(I\lesssim I_{\rm th})} = \frac{n_{\rm tot}}{(-E_{1'3,\rm th})f_{\rm V}} \left(\frac{N_{\rm p}\tau_3(1-\tau_2/\tau_{32})\Delta_0}{|e|2\Delta n_{\rm th}} \right)^2 \tag{2.93}$$

and

$$\mathcal{R}|_{(I\gtrsim I_{\rm th})} = \mathcal{R}|_{(I\lesssim I_{\rm th})} \left[1 - \frac{2\Delta n_{\rm th}}{n_{\rm tot}} \left(\frac{1-\frac{\tau_2}{2\tau_{31}}}{1+\frac{\tau_2}{\tau_{31}}} \right) \right] \tag{2.94}$$

for $\mathcal{R}$ just above threshold. Thus, the fractional change in the differential resistance $\mathcal{R}$ at threshold becomes

$$\begin{aligned} \frac{\Delta\mathcal{R}_{\rm th}}{\mathcal{R}_{\rm th}} &\equiv \frac{\mathcal{R}|_{(I\lesssim I_{\rm th})} - \mathcal{R}|_{(I\gtrsim I_{\rm th})}}{\mathcal{R}|_{(I\lesssim I_{\rm th})}} \\ &= \frac{2\Delta n_{\rm th}}{n_{\rm tot}} \left(\frac{1-\frac{\tau_2}{2\tau_{31}}}{1+\frac{\tau_2}{\tau_{31}}} \right) \end{aligned} \tag{2.95}$$

The value of $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ is plotted for some typical parameters in Fig. 2-10. It may be noted that $n_{\rm tot} = n_1 + n_2 + n_3$ is assumed for these calculations where $n_{\rm tot}$ is the doping in the active region. However, for a QCL design with more than 3 levels per module, $n_{\rm tot}$ in the above expressions must be replaced by $n_i + n_u + n_l$ (as will be shown in section 3.3.4), where i, u, l specify the injector level, the upper radiative level, and the lower radiative level, respectively. In that case, the observed value of $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ will be enhanced as compared to the values calculated in Fig. 2-10 when other parameters are kept the same, since $n_i + n_u + n_l$ becomes smaller than the total doping (number of electrons) available per module.

For the sake of completeness of discussion, it may be mentioned that it is theoretically possible for the slope of the $I - E_{1'3}$ to decrease beyond threshold, making $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ negative. This can happen for $2\tau_{31} < \tau_2 < \tau_{32}$, which is neither realizable in a realistic design, and nor desired (since $3 \rightarrow 1$ is essentially a leakage channel for current flow). For this case, the current I in a non-lasing structure is greater as compared to that in a lasing structure above threshold. This happens because more electrons get piled up in the lower radiative level in a lasing scenario (due to onset of $3 \rightarrow 2$ stimulated transitions), which makes less population available in the

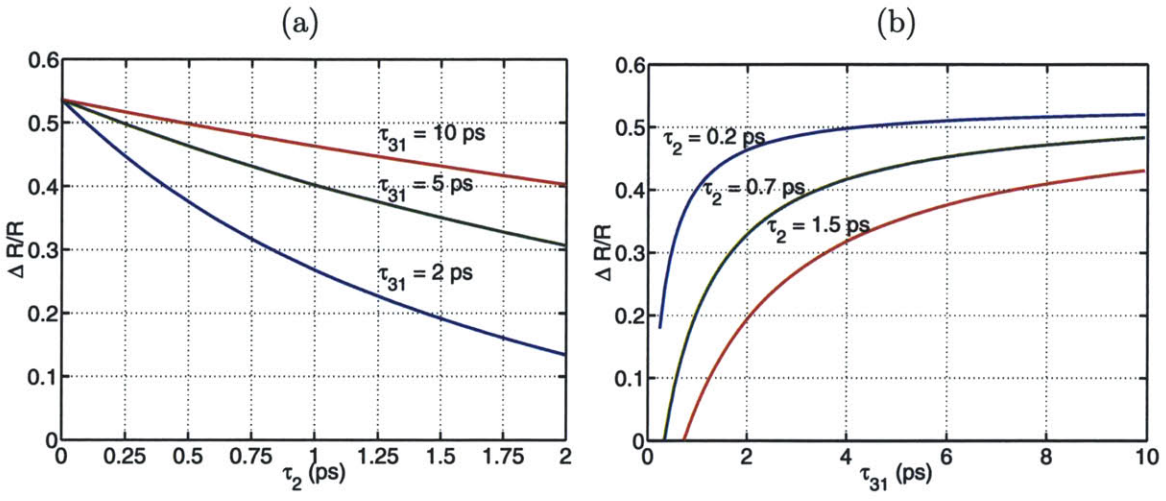


Figure 2-10: Plots for the discontinuity in differential resistance at threshold ($\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$) according to equation (2.95). The following parameters are used: $g_{\rm th} = 50$ cm^{-1}, $n_{\rm tot}/V_{\rm ac} = 5 \times 10^{15}$ cm^{-3}, $f_{23} = 0.8$, $\Delta\nu = 1.5$ THz, which results in a value of $\Delta n_{\rm th} = 0.27\, n_{\rm tot}$. Here, equation (2.33) is used to relate $g_{\rm th}$ to $\Delta n_{\rm th}$.

injector and the upper radiative levels for the overall current flow, noting that $3 \rightarrow 1$ scattering is a faster channel for current flow as compared to the $2 \rightarrow 1$ scattering for $2\tau_{31} < \tau_2 < \tau_{32}$. Note that $\tau_2 < \tau_{32}$ is a general condition to obtain a positive population inversion $n_3 - n_2$.

Even though the expressions for $\mathcal{R}$ are relatively complex, the expression for $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ is remarkably simple and no approximations have been made. These expressions hold true regardless of whether the tunneling transport through the injector barrier is coherent or incoherent (details to follow in section 2.4.3). The same expression is obtained even if the complete analysis is done with a non-unity injection efficiency ($\eta < 1$), provided η is a slowly varying function of I near threshold (since $\eta < 1$ can be thought of as being equivalent to $\eta = 1$ but with a smaller τ_{32}). Since typically $\tau_2 \ll \tau_{31}$, $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ gives an absolute measurement of the population inversion in the laser as a fraction of the combined populations of the laser levels 3 and 2, and the injector level 1 ($n_{\rm tot} = n_1 + n_2 + n_3$). For the 3-level QCL, $n_{\rm tot}$ is determined from the doping density; however, for a more complex design, populations of only the relevant levels contribute in its expression. It is clear from equation (2.95) that the maximum of $\Delta n_{\rm th} = n_{\rm tot}/2$ for $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th} = 1$, which corresponds to $n_1 = n_3$ and $n_2 = 0$.

In literature [148], it has been argued that $\mathcal{R}|_{I \gtrsim I_{\rm th}}$ will approach zero, and hence $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ will approach unity as the lower state lifetime $\tau_2 \rightarrow 0$, making $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ indicative of how fast the depopulation is. The underlying assumption for that argument is that $dn_3/dI \propto dV_{\rm bias}/dI$. However, this assumption is not strictly correct. Even though dn_3/dI above threshold, which is derived to be

$$\left.\frac{dn_3}{dI}\right|_{I>I_{\rm th}} = \frac{N_{\rm p}}{|e|}\frac{\tau_2}{\left(1+\frac{\tau_2}{\tau_{31}}\right)} \tag{2.96}$$

does indeed approach zero as $\tau_2 \rightarrow 0$, the resistance $dV_{\rm bias}/dI$ remains > 0 since the current flow in the device keeps increasing with increased bias as it is still dictated by resonant-tunneling through the injector barrier. Only for the case of coherent tunneling $\left(\frac{\Delta_0}{\hbar}\right)^2\tau_{\|}\tau_3 \gg 1$, does $(dV_{\rm bias}/dI)_{I>I_{\rm th}} \rightarrow 0$, since in that regime of operation $n1 \approx n3$ and populations in levels $1'$ and 3 are in equilibrium sharing the same quasi-Fermi level, and no extra voltage could be developed across the injector barrier once the threshold condition is met. The exact expressions for the differential resistance in the limit $\tau_2 \rightarrow 0$ are derived to be

$$\mathcal{R}|_{I<I_{\rm th},\tau_2\rightarrow 0} = \frac{1}{(-E_{1'3})f_{\rm V}n_{\rm tot}}\left(\frac{N_{\rm p}}{|e|}\frac{\hbar}{\tau_{\|}}\right)^2 \frac{\left(\left(\frac{\Delta_0}{\hbar}\right)^2\tau_{\|}\tau_3 + \left(\frac{E_{1'3}}{\hbar}\right)^2\tau_{\|}^2 + 1\right)^2}{\left(\frac{\Delta_0}{\hbar}\right)^2\tau_{\|}} \tag{2.97}$$

and

$$\mathcal{R}|_{I>I_{\rm th},\tau_2\rightarrow 0} = \frac{1}{(-E_{1'3})f_{\rm V}(n_{\rm tot}-2\Delta n_{\rm th})}\left(\frac{N_{\rm p}}{|e|}\frac{\hbar}{\tau_{\|}}\right)^2 \frac{\left(\left(\frac{E_{1'3}}{\hbar}\right)^2\tau_{\|}^2 + 1\right)^2}{\left(\frac{\Delta_0}{\hbar}\right)^2\tau_{\|}} \tag{2.98}$$

The differential resistance $\mathcal{R}$ will be plotted for some particular cases in section 2.5.1.

2.4.3 Coherent versus incoherent tunneling

In the most general case, the following condition must be satisfied to obtain the necessary population inversion $\Delta n_{\rm th}$ for lasing before the peak bias of $E_{1'3} = 0$

$$\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\tau_3 > \frac{2\Delta n_{\rm th}}{n_{\rm tot} - 2\Delta n_{\rm th} - (n_{\rm tot} + \Delta n_{\rm th})\tau_2/\tau_{32}} \tag{2.99}$$

This result is obtained from equation (2.92) as a necessary condition for $E_{1'3,\rm th}$ to exist. A rough estimate of the minimum value of Δ_0 required to meet this condition can be estimated from the plots in Fig. 2-11(b). Figure 2-11(b) can also be interpreted in a different way, since it determines the maximum value of population inversion that can be achieved (i. e. at $E_{1'3} = 0$) for a given Δ_0 in a non-lasing device. For example, in the limit $\tau_2 \rightarrow 0$, a maximum of $\Delta n = n_{\rm tot}/2$ can be obtained if $\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\tau_3 \gg 1$.

The resonant-tunneling through the injector barrier is *incoherent* if the expression $\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\tau_3 \ll 1$. In this regime, the Rabi oscillation of electron population between levels $1'$ and 3 is overdamped and the electron population on either side of the injector barrier (i. e. in levels $1'$ and 3 respectively) cannot maintain a coherent phase with respect to each other, even at resonance. Consequently, as can be deduced from equation (2.99), it is difficult to obtain population inversion to meet lasing threshold. For incoherent tunneling, the following expressions hold for the current at resonance ($E_{1'3} = 0$) and the corresponding level populations

$$\begin{aligned} I_{\rm max} &\approx \frac{n_{\rm tot}|e|}{2N_{\rm p}} \left(\frac{1}{\tau_3}\right) \left[\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\tau_3\right] \\ n_3 &\approx \frac{n_{\rm tot}}{2} \left[\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\tau_3\right] \\ n_1 &\approx n_{\rm tot} \\ &\ldots \text{in the limit} \quad \left[\left(\tfrac{\Delta_0}{\hbar}\right)^2 \tau_{\parallel}\tau_3 \ll 1,\ \tau_2 \ll (2\tau_{32})\right], I_{\rm max} < I_{\rm th} \end{aligned} \tag{2.100}$$

Thus, for incoherent injection, the maximum current is small and is $\propto \Delta_0^2$, which decreases exponentially with the thickness of the injector barrier. Also, almost the entire electron population becomes localized in the injector level $1'$.

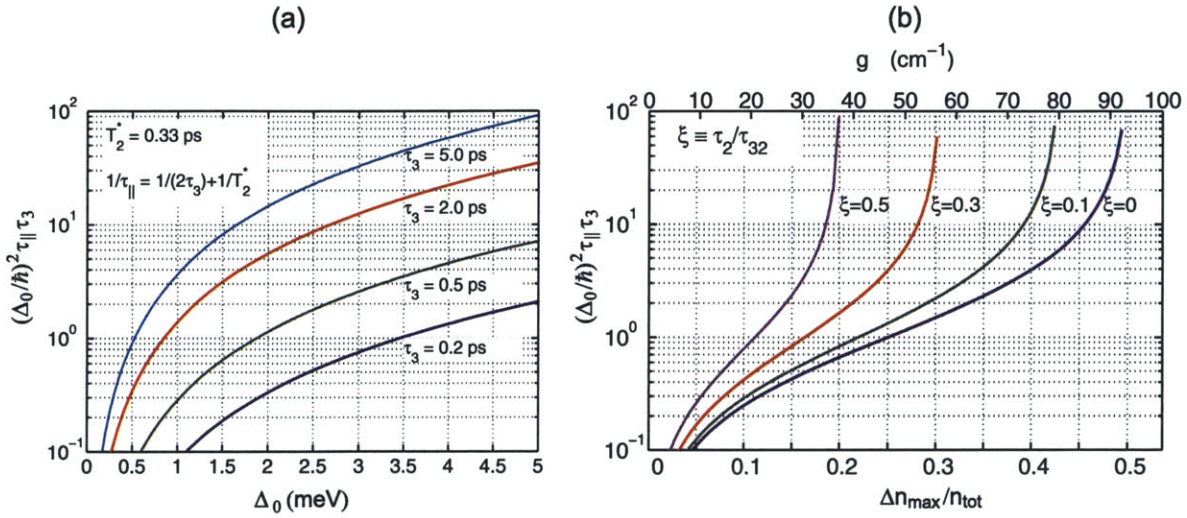


Figure 2-11: (a) Plot of the factor $\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\|}\tau_3$ to determine the regime of resonant-tunneling transport through a potential barrier ($\gg 1$ coherent, $\ll 1$ incoherent). $\tau_{\|}$ is determined from equation (2.84). (b) Plot to determine maximum value of population inversion that can be achieved (i. e. at $E_{1'3} = 0$) for a given Δ_0 in a non-lasing device, calculated using equation (2.99). The top axis indicates the conversion of population inversion $\Delta n_{\max}$ to a gain g for the typical parameters of $n_{\text{tot}}/V_{\text{ac}} = 5 \times 10^{15}$ cm^{-3}, $f_{23} = 0.8$, $\Delta\nu = 1.5$ THz, and using equation (2.33) for calculation.

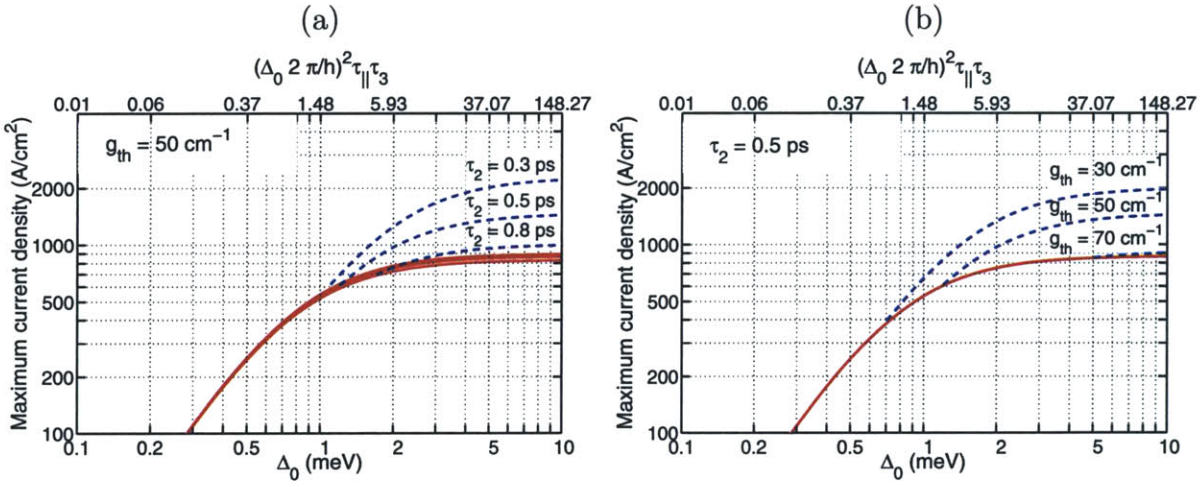


Figure 2-12: Plots of the maximum value of the current density ($I_{\max}/A_{\text{ac}}$) versus injection anticrossing Δ_0, where calculations for a lasing device are indicated by blue (dashed) lines, and for a non-lasing device they are indicated by red (solid) lines. The exact expressions for current as derived in equation (2.87) are used, for which the maximum value is obtained at resonance ($E_{1'3} = 0$). The following typical values are used for various parameters: $\tau_{31} = 7$ ps, $\tau_{32} = 3$ ps, $T_2^* = 0.33$ ps, $n_{\text{tot}}/V_{\text{ac}} = 5 \times 10^{15}$ cm^{-3}, $f_{23} = 0.8$, $\Delta\nu = 1.5$ THz. Equation (2.33) is used to relate g_{th} to Δn_{th}, and a module length $L_{\text{mod}} = 50$ nm is used to convert densities to numbers. The cross-sectional area of the active region A_{ac} can be chosen arbitrarily. For a fixed 3D *doping density*, the current density is $\propto L_{\text{mod}}$.

In the opposite regime of $\left(\frac{\Delta_0}{\hbar}\right)^2 \tau_{\|}\tau_3 \gg 1$, resonant-tunneling through the injector barrier is *coherent*. The maximum current is large and independent of Δ_0, and hence the thickness of the injector barrier. As derived from equation (2.87), the maximum current and the corresponding level populations for coherent transport are given by

$$\begin{aligned} I_{\max} &\approx \frac{n_{\mathrm{tot}}|e|}{2N_{\mathrm{p}}}\left(\frac{1}{\tau_3}\right) \\ n_3 &\approx n_1 \approx \frac{n_{\mathrm{tot}}}{2} \\ &\quad \ldots \text{in the limit} \qquad \left[\left(\tfrac{\Delta_0}{\hbar}\right)^2 \tau_{\|}\tau_3 \gg 1,\ \tau_2 \ll (2\tau_{32})\right], I_{\max} < I_{\mathrm{th}} \end{aligned} \tag{2.101}$$

when the population inversion could not be obtained at any bias in the device, and

$$\begin{aligned} I_{\max} &\approx \frac{n_{\mathrm{tot}}|e|}{2N_{\mathrm{p}}}\left(\frac{2}{3\tau_2}\right)\left(1 - \frac{2\Delta n_{\mathrm{th}}}{n_{\mathrm{tot}}}\right) \\ n_3 &\approx n_1 \approx \left(\frac{n_{\mathrm{tot}}}{3} + \frac{\Delta n_{\mathrm{th}}}{3}\right) \\ &\quad \ldots \text{in the limit} \qquad \left[\left(\tfrac{\Delta_0}{\hbar}\right)^2 \tau_{\|}(3\tau_2/2) \gg 1,\ \tau_2 \ll \tau_{31}\right], I_{\max} > I_{\mathrm{th}} \end{aligned} \tag{2.102}$$

for a lasing device. These expressions provides an important conclusion, in that for a *lasing* device the maximum current in the coherent transport regime is limited by the lower state lifetime, and not by the upper state lifetime as can be wrongly interpreted from the expression for a non-lasing device, which is often the one mentioned in literature [148, 166]. This is because after the onset of lasing, the upper state lifetime $(\tau_3\tau_{\mathrm{st}}/(\tau_3 + \tau_{\mathrm{st}}))$ is reduced due to increased stimulated emission ($\tau_{\mathrm{st}} \downarrow$), although it still has to be greater than $\sim \tau_2$ to sustain a population inversion. Note that in the coherent transport regime, the electron populations gets equally localized on the either side of the injector barrier, since tunneling through the barrier is not the limiting mechanism for current transport. Whereas the expressions derived above were for limiting cases, the exact expressions for the maximum current density obtained from equation (2.87) at $E_{1'3} = 0$ are plotted for some typical parameters in Fig. 2-12.

2.5 Calculation of optical gain spectrum using density matrices (below lasing threshold)

In the previous sections a Lorentzian linewidth was phenomenologically chosen for the optical gain spectrum in the laser. In the following two sections, the gain spectrum is analytically calculated by using the density matrices to describe the coherent interaction between light and the electronic system. While in most cases, the gain spectrum is Lorentzian; however, depending on the coherent or incoherent nature of resonant-tunneling and the dephasing lifetimes, the gain spectrum can be significantly modified. Moreover, the linewidth of the gain spectrum is shown to be sensitively dependent on level lifetimes and the level anticrossings. For example, in calculation of the I-V curves in Fig. 2-9, a constant value of 1.5 THz was assumed for the FWHM linewidth. However, for certain cases the linewidth depends on the parameters itself and the I-V curves need to be recalculated with correct linewidths to obtain a better understanding of the transport behavior in a QCL.

2.5.1 Optical gain spectrum as a function of injector anticrossing

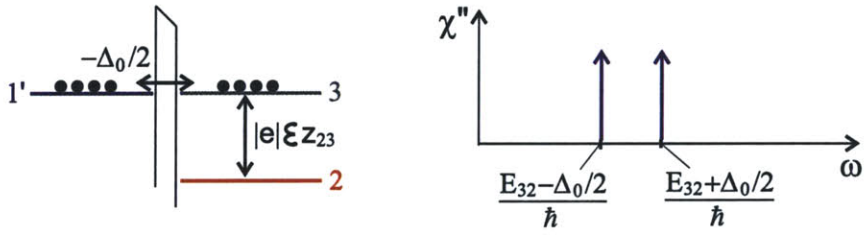


Figure 2-13: Schematic showing the optical response of the simple 3-level QCL system to an externally applied electric field $\mathcal{E}$ in the absence of any dephasing or non-radiative scattering. The optical gain (or loss) in the system is proportional to the imaginary component of the electrical susceptibility (χ''). The levels 3 and 2 are the upper and the lower radiative levels, respectively, which coherently interact with the radiation field $\mathcal{E}$ through the light-matter interaction term. Additionally, the injector level $1'$ coherently interacts with level 3 through resonant-tunneling, in the presence of which the optical response is split into two branches around the central frequency of $\omega = E_{32}/\hbar$ (the linewidth is infinitesimal in the absence of scattering or dephasing).

The optical gain spectrum for a given system can be calculated by solving the density matrix for an electronic system coherently coupled with the electromagnetic field through the light-matter perturbation term. The electromagnetic field can be treated semi-classically, as in section 2.1. Using equation (2.24) for the time-harmonic perturbation due to the electric field within the electric-dipole approximation, the time evolution of the density-matrix $\bar{\bar{\rho}}_{3\times3}$ for the 3-level QCL of Fig. 2-8 (operating below threshold) is written as

$$\frac{d}{dt}\begin{pmatrix} \rho_{1'1'} & \rho_{1'2}e^{-i\omega t} & \rho_{1'3} \\ \rho_{21'}e^{+i\omega t} & \rho_{22} & \rho_{23}e^{+i\omega t} \\ \rho_{31'} & \rho_{32}e^{-i\omega t} & \rho_{33} \end{pmatrix} = -\frac{i}{\hbar}\left[\begin{pmatrix} E_{1'} & 0 & \frac{-\Delta_0}{2} \\ 0 & E_2 & |e|z_{23}\mathcal{E}e^{+i\omega t} \\ \frac{-\Delta_0}{2} & |e|z_{23}\mathcal{E}e^{-i\omega t} & E_3 \end{pmatrix}, \bar{\bar{\rho}}_{3\times3}\right] + \begin{pmatrix} \frac{\rho_{33}}{\tau_{31}} + \frac{\rho_{22}}{\tau_2} & \frac{-\rho_{1'2}e^{-i\omega t}}{\tau_{\parallel,2}} & \frac{-\rho_{1'3}}{\tau_{\parallel,3}} \\ \frac{-\rho_{21'}e^{+i\omega t}}{\tau_{\parallel,2}} & \frac{\rho_{33}}{\tau_{32}} - \frac{\rho_{22}}{\tau_2} & \frac{-\rho_{23}e^{+i\omega t}}{\tau_{\parallel,23}} \\ \frac{-\rho_{31'}}{\tau_{\parallel,3}} & \frac{-\rho_{32}e^{-i\omega t}}{\tau_{\parallel,23}} & -\frac{\rho_{33}}{\tau_{31}} - \frac{\rho_{33}}{\tau_{32}} \end{pmatrix} \tag{2.103}$$

where, ρ_{ij} are slowly varying functions of time on the scale of the optical time period $2\pi/\omega$, and $\bar{\bar{\rho}}_{3\times3}$ in the commutator on the right is the same as the matrix operated upon by the derivative d/dt on the left. The density-matrix equation above essentially models the system shown in Fig. 2-13 but with scattering and dephasing terms included. The dephasing times $\tau_{\parallel,2}$, $\tau_{\parallel,23}$, and $\tau_{\parallel,3}$ are determined from the level lifetimes as in equation (2.84)

$$\frac{1}{\tau_{\parallel,2}} = \frac{1}{2\tau_2} + \frac{1}{T_2^*} \tag{2.104}$$

$$\frac{1}{\tau_{\parallel,23}} = \frac{1}{2\tau_2} + \frac{1}{2\tau_3} + \frac{1}{T_2^*} \overset{\tau_2 \ll \tau_3}{\approx} \frac{1}{2\tau_2} + \frac{1}{T_2^*} \tag{2.105}$$

$$\frac{1}{\tau_{\parallel,3}} = \frac{1}{2\tau_3} + \frac{1}{T_2^*} \tag{2.106}$$

where, T_2^* is the pure-dephasing time as described in section 2.4.2. T_2^* is a phenomenological parameter in the expressions above, and is taken to be the same for electrons in all levels assuming that all the electrons see a similar scattering environment [29].

The time-harmonic electric field is written as

$$\mathbf{E} = \mathbf{z}\,\mathcal{E}(e^{+i\omega t} + e^{-i\omega t}) = \mathbf{z}\,2\mathcal{E}\cos(\omega t) \tag{2.107}$$

where, $\mathcal{E}$ is taken as real, and the electric-field is assumed to be polarized in the growth (z) direction for simplicity. The dipole matrix element z_{23} between the radiative levels 3 and 2 is given by equation (2.28). It may be noted that the dipole matrix element between subbands $1'$ and 2 is zero since the subbands are calculated in the tight-binding scheme of Fig. 2-8(b).

The problem at hand is similar to that of the interaction between a two-level atom (say levels 3 and 2) and a time-harmonic electromagnetic field, which has been treated in numerous texts [185, 143]. As the field is turned on for such a system that has a finite dipole moment z_{23} between the two levels, a fraction of the level populations oscillate between the two-levels at a frequency Ω, which is close to the *Rabi-frequency* $\Omega_{\mathrm{R}} = |e|z_{23}\mathcal{E}/\hbar$. The exact frequency and the amplitude of the oscillations is a function of Ω_{R}, the detuning $\omega - \omega_{32}$, and also of the various dephasing times involved. Consequently, coherences ρ_{32} and ρ_{23} are induced between the two levels, which vary time-harmonically at the frequency of the driving field ω, modulated by a slowly varying amplitude at frequency Ω. The amplitude and phase of these coherences is also a function of Ω_{R}, $\omega - \omega_{32}$, and the various dephasing times, and additionally of the difference in populations $\rho_{33} - \rho_{22}$. The coherent oscillations at frequency Ω eventually decay due to various dephasing processes, including that of spontaneous emission. However a steady state effect persists, in which the level populations ρ_{22} and ρ_{33} become independent of time, whereas the coherences ρ_{32} and ρ_{23} vary as $e^{\pm i\omega t}$. The time-harmonic coherences lead to an induced polarization, which can either dissipate or enhance the energy of the electromagnetic field. Hence, even though the level populations remain independent of time, level transitions constantly take place to account for the energy transfer between the atomic system and the electromagnetic field. However, some secondary processes must be present to keep the level populations constant and allow for conservation of energy. In QCLs, those

secondary processes are the various transport processes such as resonant-tunneling and non-radiative scattering. It may be noted that the coherences ρ_{32} and ρ_{23} become zero as $\mathcal{E} \rightarrow 0$, even though z_{23} may be finite and the levels may have finite populations, since these coherences are defined for the ensemble and average to zero due to dephasing in the absence of the driving field. Also this is inherently a non-linear process and effects such as gain saturation can typically be observed at high field intensities.

In the ansatz $\bar{\bar{\rho}}_{3\times3}$ of equation (2.103), the coherences corresponding to levels $1'$ and 3 ($\rho_{1'3}$ and $\rho_{31'}$) vary slowly with time in the transient state, and become independent of time in the steady state. This is because these levels have resonant interaction characterized by the anticrossing energy Δ_0, and the phase of the electrons in levels $1'$ and 3 have a fixed relationship with respect to each other at and close to resonance. For this reason, the coherences between levels $1'$ and 2 must also be taken as time-harmonic, in addition to those for levels 3 and 2, despite the fact that the dipole matrix element $z_{1'2} = 0$. For the case of coherent tunneling through the injector barrier, the upper laser level becomes a superposition of levels $1'$ and 3 close to $1' - 3$ resonance. Consequently, an energy splitting takes place, which modifies the gain spectrum as will be seen below.

Whereas in transient state, ρ_{ij} are slowing varying with time, in steady state $d\rho_{ij}/dt = 0$. This, along with the *rotating-wave approximation* $|\omega - \omega_{32}| \ll \omega$ [143], which is essentially equivalent to ignoring the contribution of $e^{\pm i(\omega+\omega_{32})t}$ terms since those terms approximately average to zero over time scales longer than $1/\omega$, can be used to solve equation (2.103) analytically. The induced electrical polarization $\mathbf{z}\,\mathcal{P}(t)$ in the electronic system due the optical field is specified by

$$\begin{aligned} \mathcal{P} &\equiv \epsilon_0 \chi(\omega) \mathcal{E} e^{+i\omega t} + \text{c.c.} \\ &= \frac{-|e|\langle \hat{z} \rangle}{V_{\text{ac}}} \\ &= \frac{-|e| z_{23}}{V_{\text{ac}}} \left\{ [\rho_{32}(-\omega) + \rho_{23}(\omega)]\, e^{+i\omega t} + \text{c.c.} \right\} \end{aligned} \tag{2.108}$$

where, $\rho_{23}(\omega) = \rho_{32}^*(\omega)$, and V_{ac} is the volume of the active region. The electrical

susceptibility $\chi(\omega)$ becomes

$$\begin{aligned}\chi(\omega) &\equiv \chi'(\omega) + i\chi''(\omega) \\ &= \frac{-|e|z_{23}}{V_{ac}\epsilon_0\mathcal{E}}\left[\rho_{23}(\omega) + \rho_{23}^*(-\omega)\right] \end{aligned} \tag{2.109}$$

The optical gain coefficient $g(\omega)$ (in meter^{-1}) is related to $\chi''(\omega)$ as

$$g(\omega)|_{(\chi',\chi'')\ll n_r^2} = \frac{\chi''(\omega)}{n_r}\frac{\omega}{c} \tag{2.110}$$

Figure 2-14 shows plots of the calculated gain spectrum in the limit $\mathcal{E} \to 0$ (which ensures absence of any stimulated transitions, and any non-linear saturation effects) for various parameters. If transport through the injector barrier is not coherent, the FWHM linewidth for the gain spectrum becomes $\Delta\nu \approx (\pi\tau_{\|,23})^{-1}$ (in Hz), where $\tau_{\|,23}$ is given by equation (2.105). This is because $\tau_{\|,23}$ determines the decay rate for the coherences corresponding to the radiative levels 3 and 2. This result is derived for a simple case of a two level atom in Yariv [185], and is applicable to this analysis also. However, in the case of coherent tunneling such that $(\Delta_0/\hbar)^2\tau_{\|,3}\tau_3 \gg 1$, the linewidth gets additionally broadened due to energy splitting between levels $1'$ and 3. Additionally, if $(\Delta_0/h) \gtrsim (\Delta\nu/2)$, two separate peaks are resolved in the gain spectrum as can be seen from Fig. 2-14(a) for $\Delta_0 = 4$ meV.

Fig. 2-14(a) shows the effect of injection anticrossing on the gain spectrum for four different values of Δ_0. As determined from Fig. 2-11(a), for the values of $\tau_3 = 2.1$ ps and $T_2^* = 0.33$ ps assumed in these calculations, resonant-tunneling through the injector barrier is coherent for $\Delta_0 \gtrsim 1.5$ meV. Correspondingly, n_3 becomes approximately equal to n_1, and the population inversion $n_3 - n_2$ and hence the peak gain start to saturate with increasing Δ_0. For large values of Δ_0, when the energy splitting due to $1' - 3$ anticrossing is resolved in the gain spectrum, the peak gain reduces since the gain is spread over a larger frequency region. Although $\chi''(\omega)$ is symmetric about $\omega = \omega_{32} = 2.90$ THz, as can be seen from Fig. 2-15(a), the gain $g(\omega)$ is higher for the peak at $\omega > \omega_{32}$ since $g(\omega) \propto \chi''(\omega)\,\omega$ (equation 2.110).

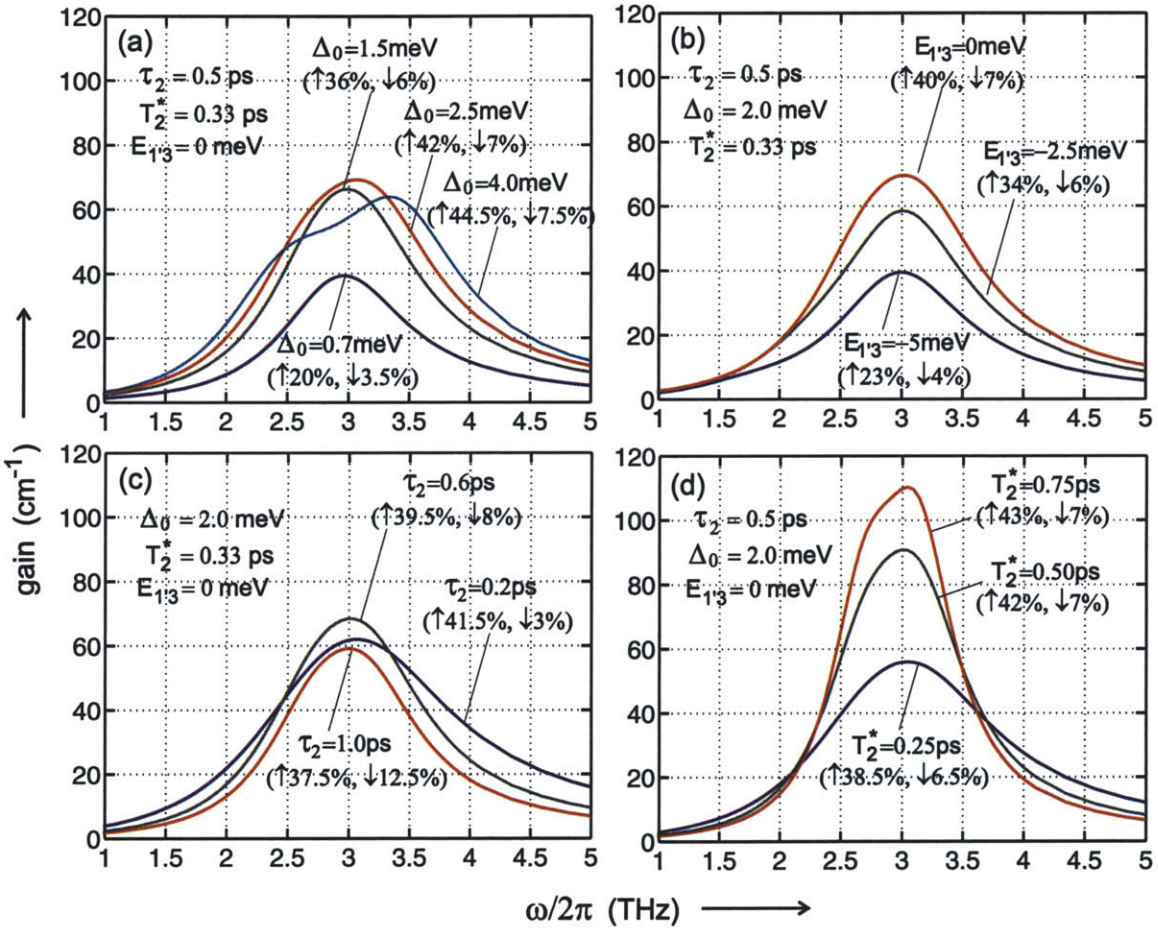


Figure 2-14: Calculated gain spectrum in the limit $\mathcal{E} \to 0$, for the 3-level QCL of Fig. 2-8(b). The values of other parameters remain the same as in Fig. 2-9, and $E_3 - E_2 = 12$ meV ($\nu_{32} = 2.90$ THz), which makes $z_{23} = 6.16$ nm for $f_{23} = 0.8$. The level populations n_3 and n_2 are indicated by $\uparrow$ and $\downarrow$, respectively, as a fraction of the total population n_{tot}. In the limit $\mathcal{E} \to 0$, there are no stimulated transitions, and the level populations are independent of ω.

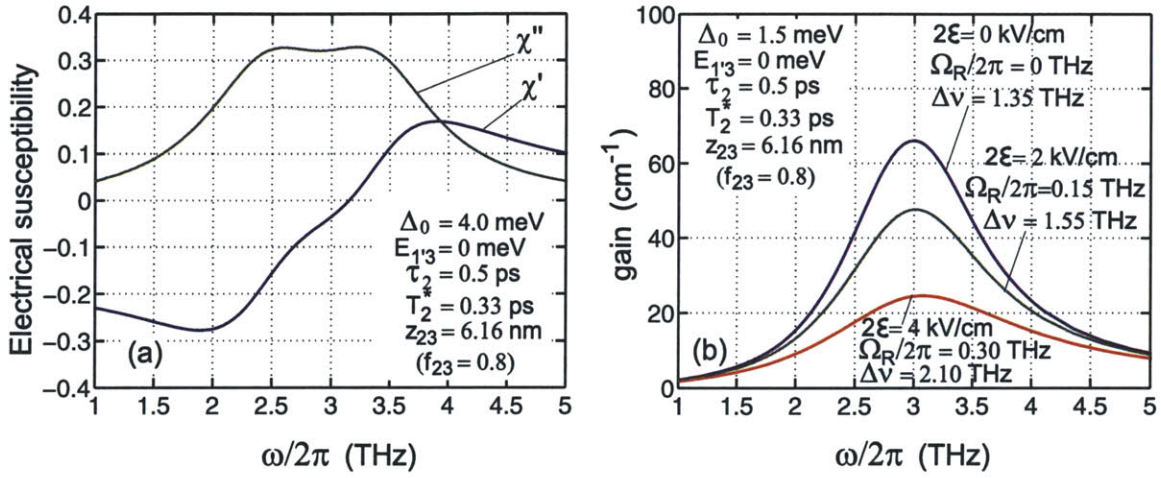


Figure 2-15: (a) Real and imaginary parts of the electrical susceptibility χ corresponding to the $\Delta_0 = 4$ meV calculation of Fig. 2-14(a). (b) Gain saturation and broadening of the gain linewidth due to increased field intensity corresponding to the $\Delta_0 = 1.5$ meV calculation of Fig. 2-14(a).

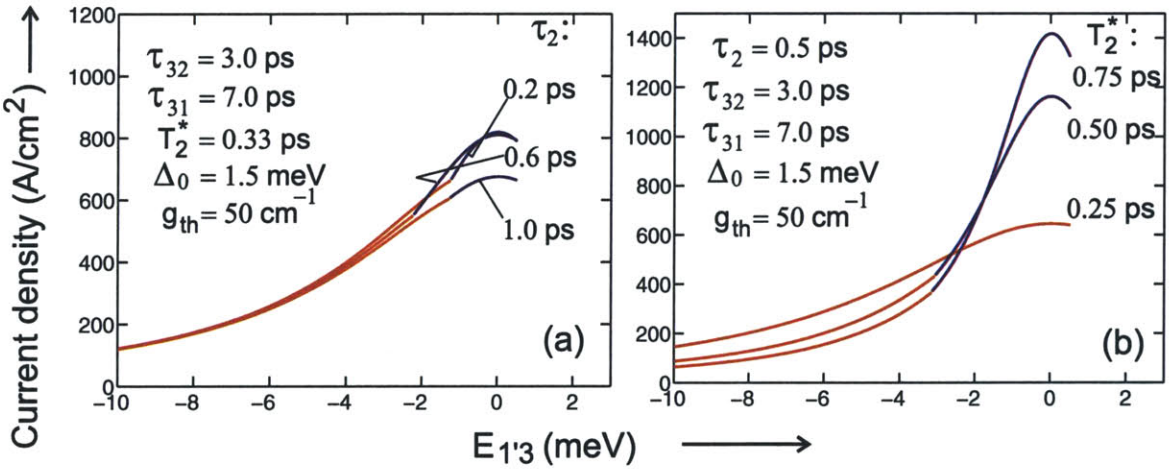


Figure 2-16: (a),(b) Current density versus bias plots of Figs. 2-9(e),(f) recalculated by including the variation of the gain linewidth as the parameters are changed. The FWHM gain linewidths, as calculated for the set of parameters mentioned in the plots, are: $\Delta\nu = 1.82$ THz, 1.30 THz, and 1.185 THz for $\tau_2 = 0.2$ ps, 0.6 ps, and 1.0 ps, respectively, in (a), and $\Delta\nu = 1.65$ THz, 1.065 THz, and 0.90 THz for $T_2^* = 0.25$ ps, 0.50 ps, and 0.75 ps, respectively, in (b).

The gain spectra shown in Figs. 2-14(a), (c) and (d) are for the peak bias $E_{1'3} = 0$. In a laser, the gain threshold should be met at an earlier bias $E_{1'3} < 0$. Beyond threshold, the population inversion gets pinned to a constant value and the extra current goes into stimulated transitions, which increase the amplitude of the radiation electrical field inside the cavity. With an assumption that the gain linewidth does not change with increased intensity, a linear optical power versus current behavior is expected, as derived in equation (2.62). However, in reality, the linewidth of the gain spectrum changes with increased field intensity as calculated in Fig. 2-15(b). This is due to the reduction in the upper state lifetime due to stimulated transitions, which causes additional lifetime broadening. Consequently, the output power versus current (L-I) slope is expected to deviate from its typical linear behavior near the peak bias, as is indeed observed experimentally[1]. Also, this will result in a slight modification of the $I - E_{1'3}$ curves in Fig. 2-9 by reducing the peak current densities from what plotted since the population inversion Δn needs to increase with an increased linewidth. The typical values of peak electric field amplitudes in terahertz QCL cavities are expected to be in the range of $2\mathcal{E}_{\max} \sim 1 - 3$ kV/cm. This rough estimate is obtained from the GaAs expression for $\tau_{\rm st}$ in equation (2.47) using the fact that $\tau_{\rm st}$ cannot possibly become much smaller than $\sim$ 1 ps (since typically $\tau_2 \gtrsim 0.2$ ps and the effective upper state lifetime $\tau_3\tau_{\rm st}/(\tau_3 + \tau_{\rm st})$ has to be kept greater than $\sim \tau_2$ to sustain any population inversion). This can also be confirmed from the experimentally measured output powers in terahertz QCLs that will be presented in chapter 3[2].

The $I - E_{1'3}$ curves in Fig. 2-9 were plotted for a constant FWHM linewidth of 1.5 THz. However, as can be noted from Fig. 2-14, the linewidth changes significantly with τ_2 and T_2^*. The $I - E_{1'3}$ curves of Fig. 2-9 are plotted again in Fig. 2-16 after including the variation in the linewidth for different parameters. The most notable change is in Fig. 2-16(a), whereby the range of lasing in current becomes higher for

[1]Note that the deviation of the L-I curves from linear behavior can also be caused by a change in the slope efficiency (equation 2.62) since the level lifetimes can change as the bias changes.

[2]The power flow density, $\epsilon_0 n_{\rm r} c(2\mathcal{E}_{\max})^2/2$, is $\sim 5 \times 10^7$ Watts/m^2 for $2\mathcal{E}_{\max} \sim 1$ kV/cm in GaAs, which, for a 10 μm $\times$ 100 μm sized aperture and 10 % out-coupling efficiency, is equivalent to 5 mW of output optical power. Note that the amplitude of the time-harmonic electric-field is taken as $2\mathcal{E}_{\max}$ in this section (equation 2.107).

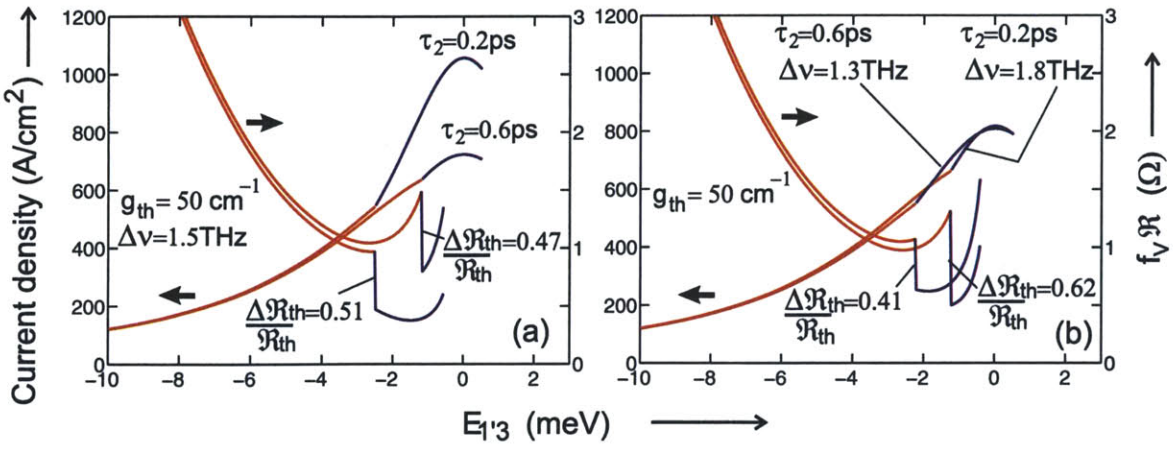


Figure 2-17: Current density (I/A_{ac}), and differential resistance ($\mathcal{R}$) versus bias ($E_{1'3}$) plots for the 3-level QCL for two different values of τ_2, to show the dependence of the discontinuity in differential resistance at threshold ($\Delta\mathcal{R}_{th}/\mathcal{R}_{th}$) on τ_2. Plot (a) is calculated with a constant gain linewidth as in Fig. 2-9(e), whereas plot (b) is calculated with correct linewidths calculated in this section as in Fig. 2-16(a). The factor $f_V\mathcal{R} \equiv f_V \dfrac{dV_{bias}}{dI}$ is equivalent to $\dfrac{N_p}{|e|}\dfrac{dE_{1'3}}{dI}$ as given by equation (2.90).

$\tau_2 = 0.6$ ps as compared to that of $\tau_2 = 0.2$ ps, for the chosen set of parameters. Additionally, the lasing threshold can be obtained even for a τ_2 as large as 1.0 ps since the population inversion Δn_{th} required to reach the gain threshold g_{th} is smaller for a smaller linewidth[1]. This behavior is analyzed further by plotting the differential resistance vs. $E_{1'3}$ in Fig. 2-17, corresponding to the plots of Fig. 2-9(e) and Fig. 2-16(a), respectively. As can be seen from Fig. 2-17(b), the discontinuity in differential resistance at threshold $\Delta\mathcal{R}_{th}/\mathcal{R}_{th}$ is larger for a smaller τ_2. This is because a greater value of population inversion Δn_{th} is needed to reach the gain threshold due to a larger gain linewidth, thereby making $\Delta\mathcal{R}_{th}/\mathcal{R}_{th}$ large according to equation (2.95).

The analyses presented in section 2.4.2 and the present section bring forth important features that are key to the understanding the operation of terahertz QCLs in general. Several of these will become more apparent in chapter 3. One of the

[1]A longer lower state lifetime τ_2 will only work if the upper state lifetime τ_3 remains sufficiently greater than τ_2. However, at higher temperatures, τ_3 decreases rapidly (as will be shown in section 3.3.4), in which case a design with a more coherent injection mechanism and a smaller τ_2 may be the only option to obtain lasing at high-temperatures. The longer lower state lifetime may be needed for designs where injection anticrossing Δ_0 cannot be made large, for example in the low-frequency ($\nu \lesssim 2.0$ THz) QCL designs (section 3.5).

most important parameters in the design process is the value of the injection anti-crossing Δ_0. As was shown in this section, for relatively small values of Δ_0 such that $(\Delta_0/\hbar)^2\tau_{\|,3}\tau_3$ being close to 1, making τ_2, the lifetime of the lower laser level, as small as possible ($\sim$ 0.2 ps) may not provide the maximum range of current operation beyond threshold, and a moderate value ($\sim$ 0.5 ps that is neither very large nor very small) may provide the best performance. This typically applies to long wavelength ($\nu \lesssim 2$ THz) resonant-phonon QCLs (section 3.5). However, for $\nu \gtrsim 2$ THz, a larger value of Δ_0 can be afforded (as will be discussed in chapter 3), in which case τ_2 can be made smaller. For $(\Delta_0/\hbar)^2\tau_{\|,3}\tau_3 \gg 1$, a value of τ_2 as low as possible is preferred, since the maximum lasing current is $\propto \tau_2^{-1}$ as is evident from equation (2.102).

For analyzing a QCL design, one of the most important measurable parameters is arguably the differential resistance $\mathcal{R}$. The behavior of $\mathcal{R}$ versus bias is in itself very informational as will be discussed in chapter 3. The more relevant parameter to the current discussion is $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$, which was shown to be a direct probe of the absolute value of the population inversion at threshold (equation 2.95). The variation of $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ with temperature, along with the corresponding behavior of the I-Vs, can tell about the parameters that determine the temperature performance of the QCL. An important finding in this section was the fact that $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ is large for a design with small τ_2, not because the lower state population may be smaller as has been conventionally believed, but because the population inversion required to reach the gain threshold in such a design will be greater due to a larger gain linewidth.

2.5.2 Optical gain spectrum as a function of collector anticrossing

Although the 3-level design of Fig. 2-8 captures the essentials of QCL operation, it needs to be modified further for a practical implementation. In that design as such, it is difficult to maintain a large dipole matrix element z_{23} for the radiative transition, while simultaneously keeping τ_{31} large. A modified technique with the so-called *resonant-phonon* depopulation scheme [168, 72], as shown in Fig. 2-18, has

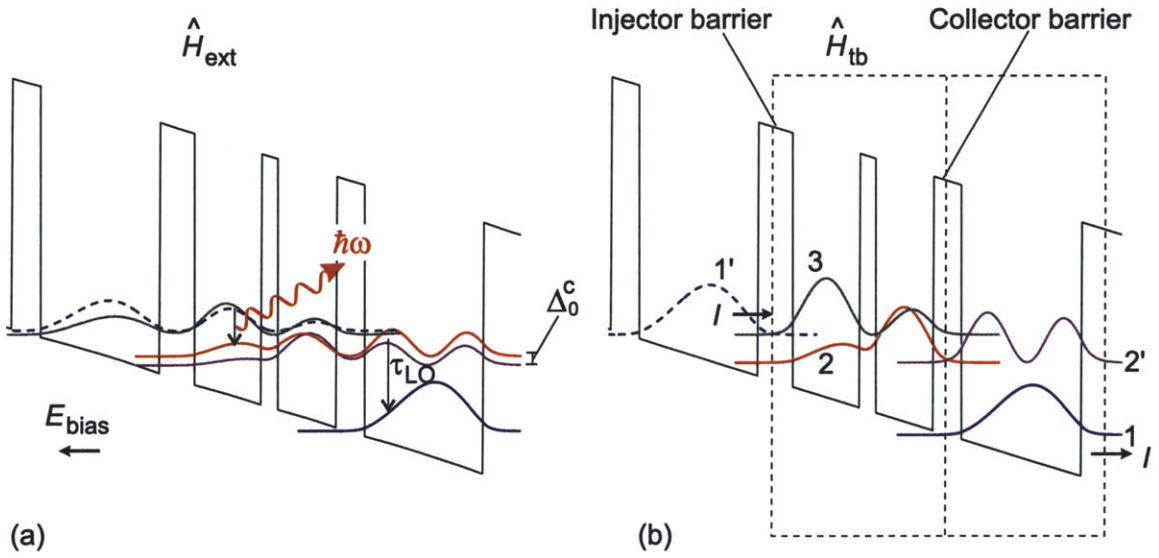


Figure 2-18: Plot of the magnitude squared envelope wavefunctions for a 4-level QCL in the (a) extended, and (b) tight-binding schemes similar to those of Fig. 2-8. In comparison to the design in Fig. 2-8, the lower laser level 2 in this design is depopulated by resonant-tunneling into level 2′, which itself is depopulated by fast electron-LO-phonon scattering into level 1. This scheme is termed as the *resonant-phonon* depopulation scheme [168, 72], and the resonant-tunneling through the collector barrier is characterized by the anticrossing energy Δ_0^c. For the tight-binding calculation in plot (b), one QCL module is further divided into two sub-modules to describe the resonant-tunneling transport through the collector barrier.

shown the best performance for terahertz QCLs until the time of writing of this thesis. In the tight-binding calculation of Fig. 2-18(b), the upper radiative level 3 is kept spatially isolated from the injector level 1 by introducing an additional well and a barrier, while it still maintains a large spatial overlap with the lower radiative level 2. This makes both τ_{31} and z_{23} large as desired. The depopulation of level 2 is facilitated by resonant-tunneling through the collector barrier into level 2′, which has an ultra-short lifetime due to fast electron-LO-phonon scattering. The energy spacing of the levels is such that the injector resonance $1' - 3$ happens at the same bias as the collector resonance $2 - 2'$ to maximize the current flow at the design bias.

When comparing the 3-level QCL design of Fig. 2-8, which features a diagonal

interwell radiative transition, to the 4-level QCL of Fig. 2-18, which has a more vertical radiative transition, it should be pointed out that the former may offer a better performance at higher temperatures, once the LO phonon scattering mechanism is thermally activated from level 3 to 2. In that case, the 3 level design can maintain a larger value of τ_{32} due to its diagonality, and thus might be able to sustain enough population inversion to reach the gain threshold. However, the temperature degradation mechanisms in terahertz QCLs are not yet fully understood and there could be other issues that are important. More will be discussed about this in chapter 3, where experimental results from both types of terahertz QCLs will be presented.

The main objective of this section is to analyze the effect of resonant-tunneling transport through the collector barrier on QCL operation below threshold. In a simplified model, it is assumed that $(\tau_{32'}, \tau_{31}) \gg \tau_{32}$, such that $\tau_3 \approx \tau_{32}$. The current I flowing in the device is taken as a parameter, rather than modeling the transport through the injector barrier as in section 2.4.2, since only the effect of collector transport is sought after here. The gain spectrum calculated by making this simplification will not be able to incorporate the coherent effects due to a large injector anticrossing, as in Fig. 2-14(a) for $\Delta_0 = 4$ meV. Even though some mid-infrared QCLs may operate in that regime, the terahertz QCLs that have so far been demonstrated do not have such large values for the injector anticrossing. Analogous to equation (2.103), the density matrix for levels 2', 2, and 3 of the tight-binding model of Fig. 2-18(b), when coupled with the electromagnetic field, is written as

$$\frac{d}{dt}\begin{pmatrix} \rho_{2'2'} & \rho_{2'2} & \rho_{2'3}e^{+i\omega t} \\ \rho_{22'} & \rho_{22} & \rho_{23}e^{+i\omega t} \\ \rho_{32'}e^{-i\omega t} & \rho_{32}e^{-i\omega t} & \rho_{33} \end{pmatrix} = -\frac{i}{\hbar}\left[\begin{pmatrix} E_{2'} & \frac{-\Delta_0}{2} & 0 \\ \frac{-\Delta_0}{2} & E_2 & |e|z_{23}\mathcal{E}e^{+i\omega t} \\ 0 & |e|z_{23}\mathcal{E}e^{-i\omega t} & E_3 \end{pmatrix}, \bar{\bar{\rho}}_{3\times 3}\right] + \begin{pmatrix} -\frac{\rho_{2'2'}}{\tau_{2'}} & \frac{-\rho_{2'2}}{\tau_{\parallel,2'}} & \frac{-\rho_{2'3}e^{+i\omega t}}{\tau_{\parallel,2'3}} \\ \frac{-\rho_{22'}}{\tau_{\parallel,2'}} & \frac{\rho_{33}}{\tau_3} & \frac{-\rho_{23}e^{+i\omega t}}{\tau_{\parallel,3}} \\ \frac{-\rho_{32'}e^{-i\omega t}}{\tau_{\parallel,2'3}} & \frac{-\rho_{32}e^{-i\omega t}}{\tau_{\parallel,3}} & I\frac{N_p}{|e|} - \frac{\rho_{33}}{\tau_3} \end{pmatrix} \tag{2.111}$$

The definition of the symbols in the equation above is similar to that in sec-

tion 2.5.1, equation (2.103). The corresponding dephasing times are

$$\frac{1}{\tau_{\parallel,2'}} = \frac{1}{2\tau_{2'}} + \frac{1}{T_2^*} \tag{2.112}$$

$$\frac{1}{\tau_{\parallel,2'3}} = \frac{1}{2\tau_{2'}} + \frac{1}{2\tau_3} + \frac{1}{T_2^*} \overset{\tau_{2'} \ll \tau_3}{\approx} \frac{1}{2\tau_{2'}} + \frac{1}{T_2^*} \tag{2.113}$$

$$\frac{1}{\tau_{\parallel,3}} = \frac{1}{2\tau_3} + \frac{1}{T_2^*} \tag{2.114}$$

Following the argument in section 2.5.1, the coherences corresponding to the levels $2'$ and 3 should be taken as time-harmonic, even though the dipole matrix element $z_{2'3} = 0$. Expressions for the electrical susceptibility $\chi(\omega)$ and the optical gain coefficient $g(\omega)$ remain the same as in equation (2.109), and equation (2.110), respectively.

Figure 2-19 shows plots of the calculated gain spectrum, analogous to those in Fig. 2-14, for different values of the collector anticrossing Δ_0^c, and the energy alignment $E_{22'}$ (which is a function of applied bias V_{bias} similar to that in equation 2.90). For a typical collector anticrossing of $\Delta_0^c = 4$ meV in Fig. 2-19(a), it can be noticed that the peak gain for a bias corresponding to $E_{22'}$ as low as -4 meV is almost the same as that for the resonant condition of $E_{22'} = 0$ meV. This suggests that in a resonant-phonon QCL design, it may be advantageous to design for the collector anticrossing energy alignment ($E_{22'} = 0$) to occur at a bias smaller than that for the injector anticrossing energy alignment ($E_{1'3} = 0$). This will allow the threshold condition to be reached earlier in bias, leading to a greater lasing range from I_{th} to I_{max}. For such a design, the output power versus current slope can deviate from its linear behavior near the peak bias of $E_{1'3} \sim 0$, when $E_{22'} > 0$, which leads to a reduction in the peak gain as seen from Fig. 2-19(a). However, this effect is expected to be smaller than the one caused by broadening of the gain linewidth due to increased field intensity, as discussed in section 2.5.1.

The effect of the collector anticrossing Δ_0^c on the gain spectrum in shown in Fig. 2-19(b). For a small Δ_0^c such that $(\Delta_0^c/\hbar)^2 \tau_{\parallel,2'} \tau_{2'} \lesssim 1$ (which can be estimated from Fig. 2-11(a) by using $\tau_{2'}$ instead of τ_3), the resonant-tunneling transport through the collector barrier is increasingly incoherent. This is reflected in the relatively

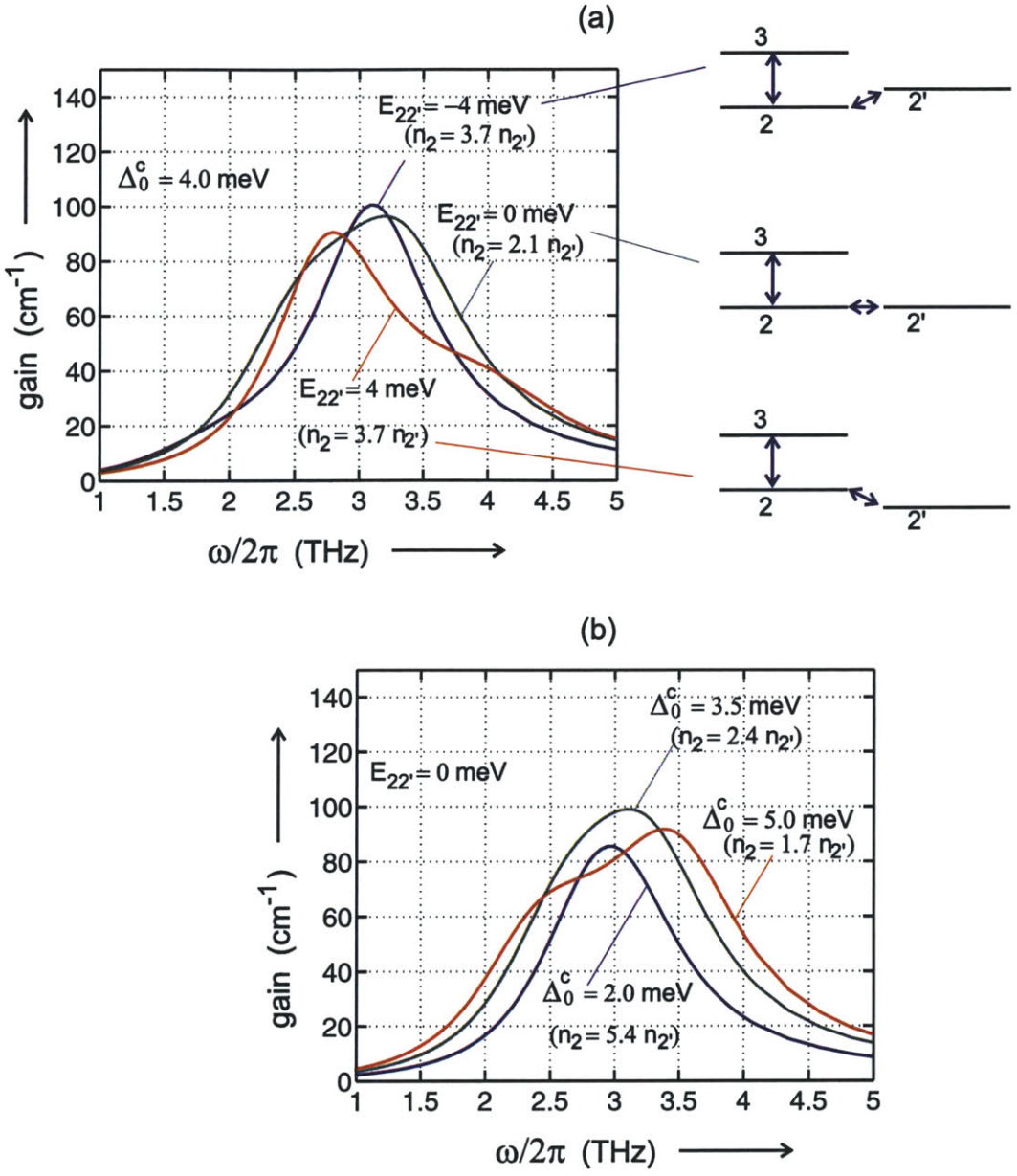


Figure 2-19: Calculated gain spectrum versus various parameters, for the 4-level QCL of Fig. 2-18(b), in the limit $\mathcal{E} \to 0$. Following values are chosen for the various parameters: $E_3 - E_2 = 12$ meV$\Leftrightarrow$ 2.90 THz, $\tau_3 = 3$ ps, $\tau_{2'} = 0.25$ ps, $T_2^* = 0.33$ ps, $f_{23} = 0.8$, and $f_{2'3} = 0$. The values of $m^* = 0.067\, m_0$ and $n_r = 3.6$ corresponding to those for GaAs are used. A current density of $I/A_{ac} = 800$ A/cm^2 is assumed, which, for a chosen module length of $L_{mod} = 50$ nm and the above mentioned level lifetimes, gives $n_3/(\text{Volume}) = 3 \times 10^{15}$ cm^{-3} and $n_{2'}/(\text{Volume}) = 0.25 \times 10^{15}$ cm^{-3}. The population of the lower laser level n_2, which is a function of the energy alignment $E_{22'} \equiv E_2 - E_{2'}$ and the collector anticrossing Δ_0^c, is indicated alongside the plots.

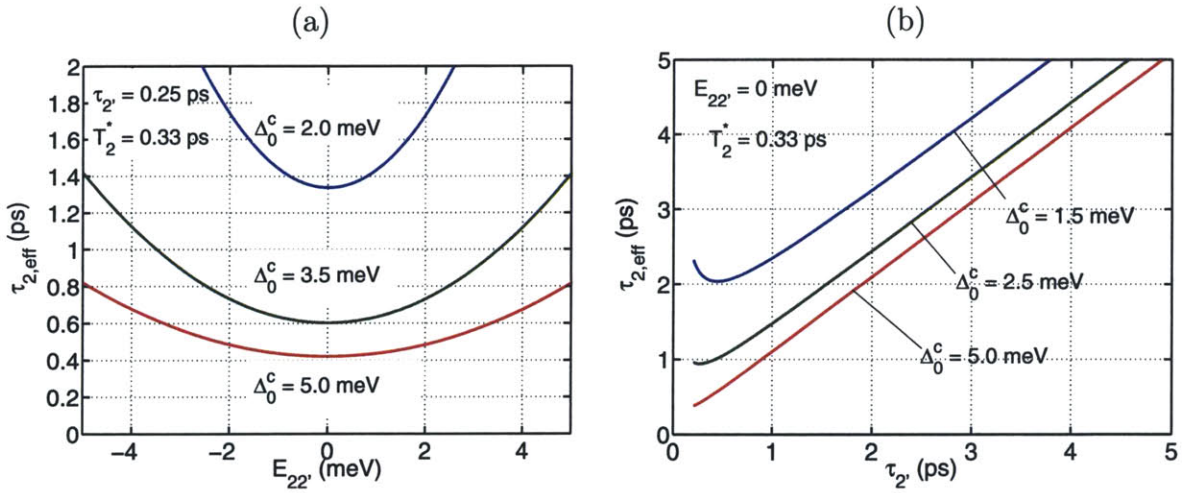


Figure 2-20: Plots for the effective lifetime of the lower laser level 2, which is also equivalent to tunneling time through the collector barrier due to the resonant-tunneling transport, as expressed by equation (2.115).

large population of level 2 as compared to that of level 2′, as shown in Fig. 2-19(b). However, the peak gain does not change as drastically, since for a smaller Δ_0^c the gain linewidth is narrower, which keeps the peak gain relatively unchanged even for a reduced population inversion. This is analogous to the behavior in Fig. 2-14(c) since a small Δ_0^c reduces the *effective lifetime* of the lower laser level, which is also equivalent to the *tunneling time* through the collector barrier. The expression for $\tau_{2,\text{eff}}$ in the limit $\mathcal{E} \rightarrow 0$ becomes

$$\begin{aligned} \tau_{2,\text{eff}} &= \tau_{2'} \frac{n_2}{n_{2'}} \\ &= \tau_{2'} \left[\frac{\left(\frac{\Delta_0^c}{\hbar}\right)^2 \tau_{\parallel,2'} \tau_{2'}/2 + \left(\frac{E_{22'}}{\hbar}\right)^2 \tau_{\parallel,2'}^2 + 1}{\left(\frac{\Delta_0^c}{\hbar}\right)^2 \tau_{\parallel,2'} \tau_{2'}/2} \right] \end{aligned} \tag{2.115}$$

This expression is derived assuming that the transport through the collector barrier is only through the $2 \rightarrow 2'$ resonant-tunneling. It may be noted that levels 2 and 2′ are calculated in the tight-binding sense as shown in Fig. 2-18.

The expression derived in equation (2.115) is true in general for resonant-tunneling transport through a potential barrier, and thus, also holds true for transport through the injector barrier to determine $\tau_{1',\text{eff}}$, once the relevant parameters are replaced.

Although for the case of injector transport, the reduction in lifetime of level 3 due to stimulated transitions should also be taken into account. The expression for $\tau_{2,\text{eff}}$ is plotted for various values of Δ_0^c in Fig. 2-20. This value can be used instead of τ_2 in section 2.5.1 to maintain the validity of the derived expressions, provided $\tau_{2,\text{eff}}$ does not change rapidly with V_{bias}. As can be noted, even for a value of Δ_0^c as large as 5 meV, the minimum value of $\tau_{2,\text{eff}}$ is about a factor of ~ 2 larger than the raw LO-phonon scattering time $\tau_{2'}$ of ~ 0.25 ps. This justifies the use of a typical value of $\tau_2 = 0.5$ ps for most of the plots in section 2.5.1. The best resonant-phonon QCL designs have typically been obtained with values of $\Delta_0^c \sim 4$ meV, which maintains a small $\tau_{2,\text{eff}}$ over a relatively large bias range. For terahertz QCLs, Δ_0^c is not preferred to be designed much larger than 4 meV, the reasons for which will become apparent in chapter 3.

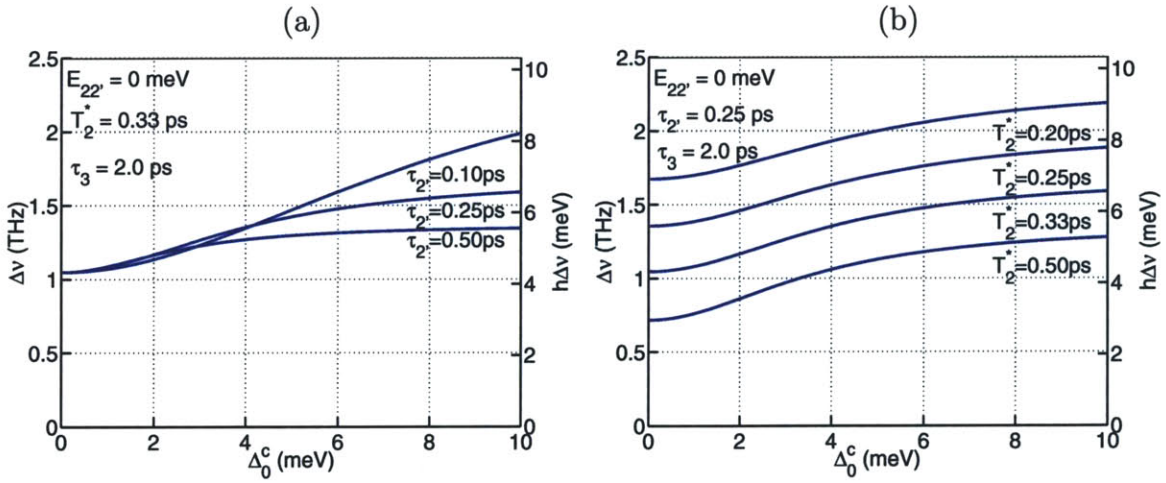


Figure 2-21: FWHM gain linewidth for some typical parameters, plotted at the collector anticrossing resonance ($E_{22'} = 0$). Equation (2.116) is used for this plot, which is valid when tunneling through either the injector or the collector barrier is not coherent, and the lineshape can be approximated by a single Lorentzian.

If tunneling through both the injector and collector barriers is not coherent, the gain linewidth (below threshold) will have the shape of a Lorentzian. Its FWHM value is modified from the one mentioned in section 2.5.1 by replacing τ_2 with $\tau_{2,\text{eff}}$ as

$$\Delta\nu \approx \frac{1}{\pi}\left[\frac{1}{2\tau_{2,\text{eff}}} + \frac{1}{2\tau_3} + \frac{1}{T_2^*}\right] \tag{2.116}$$

This expression is plotted for some typical parameters in Fig. 2-21 (note that $\tau_{2,\text{eff}}$ varies with Δ_0^c according to equation 2.115). The gain linewidth is additionally broadened for $(\Delta_0^c/\hbar)^2\tau_{\|,2'}\tau_{2'} \gg 1$, similar to the case of coherent tunneling through the injector barrier. In that case if $(\Delta_0^c/h) \gtrsim (\Delta\nu/2)$, the collector anticrossing is resolved in the gain spectrum, as shown for $\Delta_0^c = 5$ meV in Fig. 2-19(b). In general, equation (2.116) will always give a lower limit for the gain linewidth.

2.6 Summary

In this chapter, starting from a description of the light-matter interaction Hamiltonian, various radiative scattering times due to intersubband optical transitions were derived from first principles. Transport in QCLs was first described using simple rate equations, which allowed a description of photon and electron number densities below and above lasing threshold as a function of the current flowing in a device. The current flow in a QCL below and above the lasing threshold was then determined using a density matrix formalism. Density matrices were further used as a tool to analytically evaluate the gain spectrum in a QCL, by incorporating coupling of the optical field with the electronic system. Some of the results derived in sections 2.4.2, 2.5.1, and 2.5.2 are relatively new and have not yet been applied. They may provide additional insight into the operation of terahertz QCLs and their further development.

In all the aforementioned transport and gain analyses, the various non-radiative lifetimes were taken as parameters. In general, these lifetimes have a complex dependence on the electronic temperature, the electron distribution in the subbands, the wavefunctions corresponding to those subbands, and dephasing rates that that describe the coherence of the wavefunctions. The wavefunctions, when described in a tight-binding formalism, do not change significantly with bias in the range of interest, which allows a reasonable description of transport behavior with bias using the techniques presented in this chapter. However, the behavior of transport with temperature cannot be described within this scheme. The calculation of non-radiative lifetimes needs a detailed treatment, and a mathematical formulation can be found in the

previous work by Smet [152, 151] as applied to intersubband transitions. A detailed description of QCL transport in the form of Monte-Carlo simulations was undertaken by Callebaut [28, 29], which allows a calculation of electron temperatures and electron distribution self-consistently using the aforementioned mathematical formulation of the non-radiative scattering-rates, while some parameters related to dephasing are included phenomenologically. A non-equilibrium Green functions based model for QCLs has recently been developed by the group of Andreas Wacker [18, 104, 103], to which more functionality was added by Callebaut [28]. This formalism self-consistently incorporates a description of the wavefunction coherences; however, much still remains to be done before such computational results could reliably be used for QCL design.

Even though the simplistic 3 or 4 level models described in this chapter can be used to describe gain in a particular QCL design since only the radiative and the injector levels need be modeled, the calculation of loss is a much more complicated problem. A QCL design can in general have many subbands per repeated module, and a density matrix formulation of such a multi-level system, which includes coherent interaction between all the levels, may become very computationally intensive [29]. For this reason, loss due to various intersubband resonances can be phenomenologically treated separately [14]. The "free-carrier" contribution of the loss, which is affected by electron scattering in the plane of the layers of the quantum-wells, cannot in general be described by bulk semiconductor mobility parameters due to two-dimensional confinement of the electrons in the quantum-wells. Hence, a simplistic Drude-model [16] cannot be used to estimate the free-carrier contribution to the loss. For QCLs this problem is further complicated by the fact that the electric-field is polarized perpendicular to the quantum-wells (due to the intersubband polarization selection rule), and hence does not directly couple to the in-plane electronic motion. Although some work has been done on estimation of the free-carrier absorption loss in QCLs [158, 110], uncertainties still remain due to the complexity of the problem.

Chapter 3

Terahertz QCL active region: designs, experiments, and analyses

This chapter describes the various terahertz QCL active region designs that have been experimentally realized during the course of this thesis. This includes designs that made a laser, and also others that did not. As will become apparent, the improvement of an existing design is intrinsically linked to the experimental analyses of previous designs that are characteristically different, and that differ only marginally from it alike. A computational tool for modeling the electrical transport in a QCL to the desired accuracy is not yet available; however, the simplistic theory of electrical transport presented in chapter 2 was routinely used to analyze QCL operation. Some additional insight that has been developed during the course of writing of this thesis will also be discussed in this chapter. Although significant progress has been made in the development of terahertz QCLs since their inception in the year 2001, further improvement, particularly in their temperature performance, will possibly require a paradigm shift from the design methodology of the present best performing QCLs. This will become progressively more evident from the following discussions.

Numerical calculations of the bound state wavefunctions and their eigenenergies for the one-dimensional multiple quantum-well (MQW) problem were performed using a Schrödinger solver called SEQUAL [113], and a Poisson solver called FISH1D [109]. The operation of these softwares is described in a greater detail in chapter 2 of

Ref. [182]. Band-bending, which arises due to spatial separation of the positively charge dopant impurities and the negatively charged electrons, is neglected due to relatively low doping densities in the terahertz QCLs discussed in this thesis. This is equivalent to assuming $\phi(z) \approx 0$ in equation (2.14), and a self-consistent solution of the Poisson equation (2.6) with the Schrödinger equation (2.14) is not necessary. The experimentally measured laser frequencies often correspond well with those calculated by neglecting band-bending, which is the primary basis for this assumption.

A uniform conduction band (Γ valley) electron effective mass of $m^* = 0.067\,m_0$, corresponding to that of GaAs, is assumed for the GaAs/$Al_xGa_{1-x}As$ MQW structure following the introductory remarks in chapter 2. The potential barrier height for the heterostructure is taken to be $\Delta E_c = 0.894\,x$ eV [150]. This value is equivalent to $\Delta E_c = 0.72\,\Delta E_g$, where $\Delta E_g = 1.247\,x$ eV [8] is the difference in the bandgaps of GaAs and $Al_xGa_{1-x}As$. The factor of 0.72 relating ΔE_c to ΔE_g is not a definitive value since several different values in the range of $0.6 - 0.9$ have been used in literature for this material system. For the designs presented in this thesis, this value was found to agree well with the experimental results, although values of 0.65 and 0.80 have also been used in the past as detailed in chapter 6 of Ref. [166]. The barrier height affects several important design parameters, such as energy anticrossings and energy alignment of the subbands, which directly influence both the optical and electrical transport properties of a terahertz QCL. It is even more critical for long wavelength ($\nu \lesssim 2$ THz, $\lambda \gtrsim 150$ μm) terahertz QCLs, as will become more apparent in section 3.5. Barring a few particular cases, non-parabolicity is ignored for calculation of the wavefunctions and their eigenenergies, as per the introductory discussion in chapter 2. Calculations that deviate from the aforementioned assumptions will be explicitly stated in the following discussions.

The earliest of the terahertz QCL designs developed in Qing Hu's group at MIT were by Bin Xu and are discussed in his PhD thesis [182]. Subsequently, Benjamin S. Williams carried the work forward and demonstrated the first terahertz QCL from this group in the year 2002 [168]. Chapters 6 and 7 of Williams's PhD thesis [166] discuss the non-lasing, and the lasing terahertz QCL designs, respectively, developed

in this group up to August of the year 2003. The present thesis focuses on terahertz QCL designs developed from the year 2004 onwards up to the end of the year 2006 in the same group.

The layer thicknesses in the following sections will be mentioned in the units of a monolayer (ML). In the GaAs/$Al_xGa_{1-x}As$ material system, 1 ML = $a_{GaAs}/2$ = 2.825 Å [8], where a_{GaAs} is the lattice constant of GaAs. The wavefunction plots at the "design-bias", which is defined as one of the bias points at which some injector level lines up in energy with the upper radiative level, will be shown for a tight-binding calculation similar to that in section 2.4.2, Fig. 2-8(b). This is because resonant-tunneling through the injector barrier in terahertz QCLs is not coherent, and thus, the wavefunctions calculated in the extended scheme are a less accurate representation of the spatial extent of the electron distribution near the injector barrier. A Schrödinger solver developed by Hans Callebaut [29] is used for the tight-binding calculations. However, for the "parasitic-bias", which will be defined in section 3.2, the wavefunction plots will be shown in the extended scheme despite the parasitic-bias resonant-tunneling transport being even more incoherent than the design-bias transport through the injector barrier. This is just to "visibly" show the parasitic anticrossing so it is more apparent.

The MBE wafers V0416 (OWI180B) and V0418 (OWI185-M3) discussed in section 3.5.2 were grown by Dr. Z. R. Wasilewski at the National Research Council, Canada. All the other MBE wafers discussed in the chapter were grown by Dr. J. L. Reno at the Sandia National Laboratories, USA. The experimental setup and fabrication details for the measurement results presented in this chapter are discussed in detail in chapter 5 of Williams [166]. Additional experimental details pertaining to the fabrication of metal-metal waveguides using the Cu-Cu thermocompression wafer bonding technique will be discussed in chapter 4. For the results reported in this chapter, the substrate thickness for the metal-metal waveguide devices is typically (200 ± 50) μm. The metal-metal waveguides are fabricated using a dry-etching technique [170], which yields a vertical profile for the sidewalls of the ridges (chapter 4). The number of modules in the active region are chosen to keep its

thickness to approximately 10 μm for all the devices (except for the ones discussed in section 3.3.2 where the thickness will be explicitly noted) so that the processing recipes need not to be altered from wafer to wafer. All the optical power measurements are for power as collected at the detector (ScienTech model AC2500H) without correcting for power lost to non-unity collection efficiency, or due to losses in the optical path including those due to non-unity transmission from the polypropylene dewar window (Ref. [166], Fig. 5-1). All results are for Fabry-Pérot ridge cavity lasers with cleaved facets at both ends of the cavity unless high-reflectivity (HR) coating is explicitly mentioned. Any other exceptions to the aforementioned will be explicitly noted.

3.1 A sample of two early designs: M100 and T65

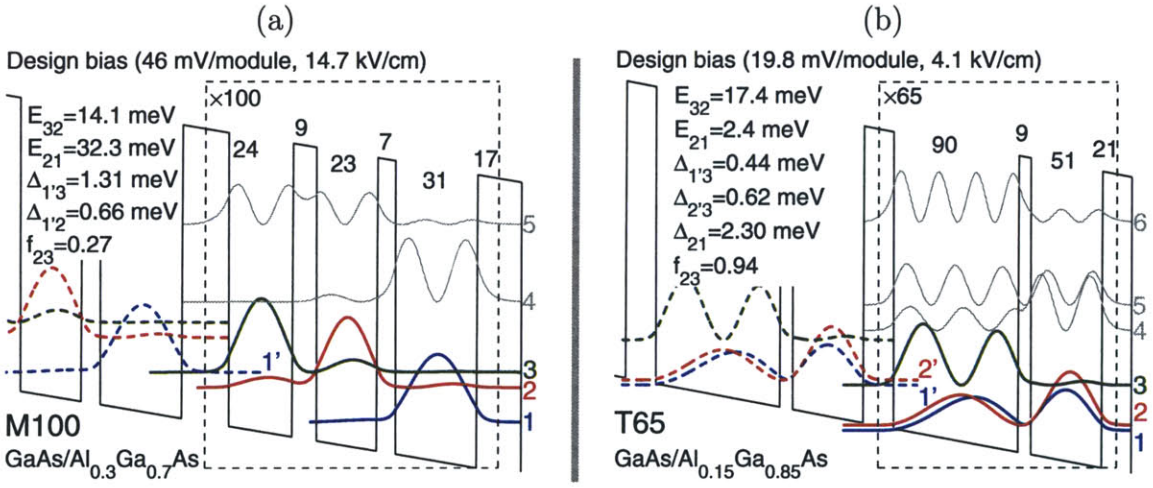


Figure 3-1: (a) Conduction band diagram for a design named M100 [183]. The barrier height $\Delta E_c = 0.72\,\Delta E_g = 0.27$ eV for the $GaAs/Al_{0.3}Ga_{0.7}As$ material system. Various relevant parameters along with the layer thicknesses (in ML) are indicated. (b) Conduction band diagram for a design named T65 [167]. The barrier height $\Delta E_c = 0.72\,\Delta E_g = 0.135$ eV for the $GaAs/Al_{0.15}Ga_{0.85}As$ material system.

Whereas a comprehensive survey of the early terahertz QCL designs that were developed in this group is presented in chapter 6 of Williams [166], this section briefly discusses two contrastingly different designs that were developed before a working terahertz QCL was demonstrated.

Figure 3-1(a) shows one of the very first designs that were developed in this group [183]. This design, named as M100, has 3-levels per module where the lower radiative level 2 is depopulated by LO-phonon scattering ($E_{21} \approx \hbar\omega_{\mathrm{LO}}$). This design is identical to the simple *model* design discussed in chapter 2 (Figs. 2-1 and 2-8) albeit the $2 \rightarrow 1$ transition is a diagonal interwell transition. This makes $\tau_{21,\mathrm{LO}} \sim 1$ ps, a value that is relatively long compared to a value of < 0.25 ps, which can be obtained with a vertical intrawell transition. The upper state lifetime τ_3 is expected to be > 10 ps for this design [183]. It appears that a population inversion should be established in this design at the indicated bias; however, lasing was not observed in any of the measured devices. There are two likely reasons for this. First, these devices were implemented in plasmon waveguides of the type shown in Fig. 1-9, which have a high loss at terahertz frequencies. Second, and perhaps a more critical reason is that enough gain is unlikely to be established in this design, even with a population inversion, due to a low oscillator strength for the $3 \rightarrow 2$ radiative transition ($f_{\mathrm{rad}} \equiv f_{23} = 0.27$), which is due to the very diagonal nature of the transition. Additionally, the more diagonal a radiative transition is, the greater is the dependence of the radiative energy separation $E_{\mathrm{rad}} \equiv E_{32}$ on the applied bias due to the Stark effect. Since a QCL typically has growth fluctuations from module to module, different modules get biased differently, which will cause a greater *inhomogeneous* linewidth broadening for a more diagonal radiative transition. A relatively broad spontaneous emission linewidth $\Delta\nu_{\mathrm{FWHM}} \sim 2$ THz was measured experimentally for this design, which confirms this argument. In comparison, successful terahertz QCL designs with $E_{\mathrm{rad}} \sim 12$ meV ($\nu_{\mathrm{rad}} \sim 3$ THz) and LO-phonon based depopulation have typically been designed with a radiative oscillator strength $f_{\mathrm{rad}} \gtrsim 0.7$ [168, 174].

The conduction band diagram in Fig. 3-1(b) is for T65, which is one of the few designs developed in this group that does not use LO-phonon scattering as a depopulation mechanism [167, 166]. Although a narrow linewidth spontaneous emission ($\Delta\nu_{\mathrm{FWHM}} \sim 1$ THz) was measured for devices fabricated into metal-metal waveguides for this design, no lasing was observed. It was believed originally that population inversion could be established between the radiative levels 3 and 2 due to fast electron-

electron scattering (as will be explained below). However, as a better understanding of resonant-tunneling transport through the injector barrier has been developed, it is finally concluded that population inversion will be difficult to establish in this design in a regime where transport through the injector barrier is not coherent [29]. In that case, the population in upper radiative level 3 will always be smaller than the populations in the injector levels $1'$ and $2'$ (which are same as those in levels 1 and 2 due to the periodicity of the structure). This can be deduced from the expressions

$$n_3 \approx \begin{cases} n_1 \left[\dfrac{\left(\frac{\Delta_{1'3}}{\hbar}\right)^2 \tau_{\parallel,3}\tau_3/2}{\left(\frac{\Delta_{1'3}}{\hbar}\right)^2 \tau_{\parallel,3}\tau_3/2 + 1} \right] & \ldots \text{for} \quad E_{1'3} = 0 \\ n_2 \left[\dfrac{\left(\frac{\Delta_{2'3}}{\hbar}\right)^2 \tau_{\parallel,3}\tau_3/2}{\left(\frac{\Delta_{2'3}}{\hbar}\right)^2 \tau_{\parallel,3}\tau_3/2 + 1} \right] & \ldots \text{for} \quad E_{2'3} = 0 \end{cases} \tag{3.1}$$

where equation (2.115) is used to relate the population in level 3 to that in levels $1'$ and $2'$ assuming that only a single injector level contributes to majority of the current transport at the specified anticrossing bias, and the dephasing time $\tau_{\parallel,3}$ is expressed as $1/\tau_{\parallel,3} = 1/2\tau_3 + 1/T_2^*$. In contrast, for the case of coherent resonant-tunneling, n_3 is approximately equal to n_1, or n_2, depending on the bias. For a thermalized electron distribution in levels 1 and 2 (due to fast intersubband electron-electron scattering inbetween the closely spaced injector levels), $n_2 < n_1$, and hence population inversion can be obtained at the bias $E_{1'3} = 0$ for which $n_3 \approx n_1$ (provided the device could be biased up to this point). Realizing this, the injector barrier was thinned down to 10 ML in a subsequent design named T150C [166]; however, T150C had an unreasonably high current density and no meaningful optical emission was observed. This behavior was ascribed to the presence of an upper level parasitic current channel in this design (section 3.2.1, Fig. 3-2), which reduces the upper state lifetime and makes it less likely to obtain population inversion.

The problem of parasitic current channels, which can occur below or close to the design bias, is a major challenge towards design of terahertz QCLs. That is the primary reason why resonant-tunneling transport through the injector barrier cannot be

made fully coherent in a terahertz QCL, as will become apparent in the next section.

3.2 The parasitic current channels

For terahertz QCLs, the radiative energies involved ($\hbar\omega \sim 4-20$ meV, $\nu \sim 1-5$ THz, $\lambda \sim 300-60$ μm) are much smaller than for mid-infrared QCLs ($\hbar\omega \sim 60-415$ meV, $\lambda \sim 20-3$ μm). Consequently, the bias at which any spurious level alignments may take place are not significantly below the design bias, which makes the said resonant interactions (characterized by the anticrossing energy at resonance) of the same order as the anticrossings at the design bias. Those unwanted level alignments constitute parasitic leakage channels for current flow in the device, and can severely limit the amount of population inversion that could be achieved. In the worst case, such parasitic current channels may prevent the device from reaching the design bias due the occurrence of an early negative differential resistance (NDR) region and the consequent formation of high-field domains, as was described in section 1.2.2. However, for mid-infrared QCLs the parasitic level couplings (particularly those due to injector levels coupling with lower energy levels in an adjacent module) are relatively insignificant since they happen at a bias much lower than the design bias. In that case, the parasitic anticrossings are negligible since the wavefunction energies are much lower in the quantum wells as compared to what they are at the design bias, which makes the potential barriers more effective in localizing the wavefunctions in their respective quantum wells.

Primarily, there are two different types of parasitic current channels that can exist in terahertz QCLs. These will be discussed next in the following sections.

3.2.1 Parasitic current channels due to higher energy levels

Figure 3-2 shows the conduction band diagram of T150C, a design similar to T65 except with a much thinner injector barrier. The calculation in Fig. 3-2(a) is for the design bias, whereas the calculation in Fig. 3-2(b) is for a bias slightly above the design bias, at which the upper radiative level 3′ anticrosses with level 4 in the

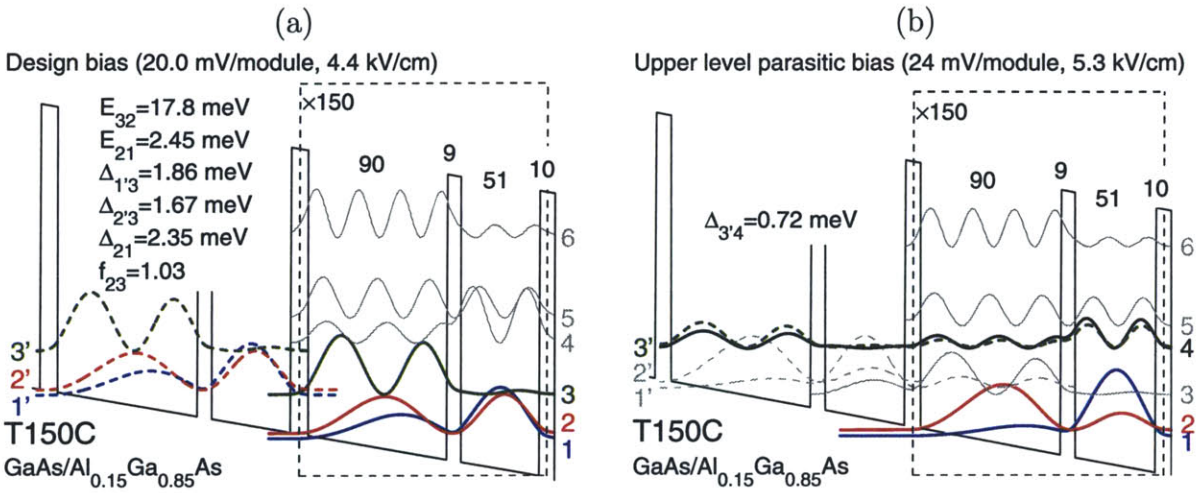


Figure 3-2: Conduction band diagrams of T150C at (a) the design bias, and (b) the upper level parasitic bias, respectively.

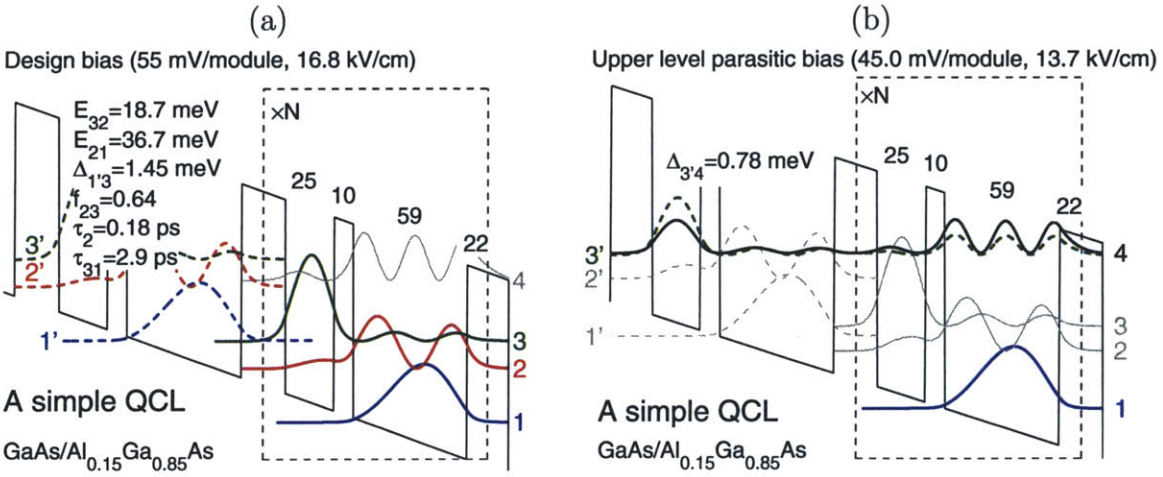


Figure 3-3: Conduction band diagrams of the simple 3-level QCL design of section 2.4.2 at (a) the design bias, and (b) the upper level parasitic bias, respectively.

adjacent module. Due to a thin injector barrier, this parasitic coupling is enhanced and constitutes an additional leakage channel from level $3'$ to the injector levels 2 and 1, thereby reducing its lifetime. Moreover, since resonant-tunneling due to such an energy alignment transports the electrons to the next module, the lifetime of level 3 is effectively reduced by twice that of the calculated value, as will be shown below. Since any current due to level 4, which lies above the radiative level 3, constitutes leakage current, this channel is attributed as an *upper level parasitic channel.*

A similar parasitic current channel due to $3' - 4$ coupling exists for the simple 3-level QCL design of chapter 2 (section 2.4.2, Fig. 2-8) as shown in Fig. 3-3(b). For this design the upper level parasitic coupling happens at a bias slightly lower than the design bias. The effective lifetime of level $3'$ due to resonant-tunneling into the parasitic level 4, $\tau_{3,\text{para}}$, can be written similar to that in equation (2.115) as

$$\tau_{3,\text{para}} = \tau_4 \left[\frac{\left(\frac{\Delta_{3'4}}{\hbar}\right)^2 \tau_{\parallel,34}\tau_4/2 + \left(\frac{E_{3'4}}{\hbar}\right)^2 \tau_{\parallel,34}^2 + 1}{\left(\frac{\Delta_{3'4}}{\hbar}\right)^2 \tau_{\parallel,34}\tau_4/2} \right] \tag{3.2}$$

where the dephasing time $\tau_{\parallel,34} = (1/2\tau_3 + 1/2\tau_4 + 1/T_2^*)^{-1} \approx (1/2\tau_4 + 1/T_2^*)^{-1}$, and the level lifetimes τ_3 and τ_4 are calculated in the tight-binding scheme of Fig. 3-3(a).

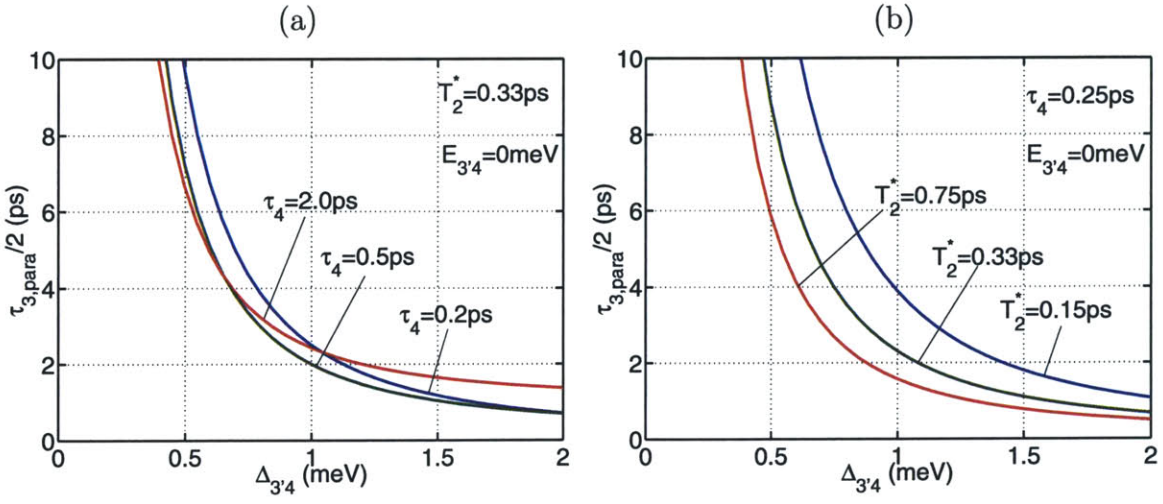


Figure 3-4: The effective lifetime of the upper radiative state $\tau_{3,\text{para}}/2$ due to coupling with a higher energy parasitic level, where $\tau_{3,\text{para}}$ is given by equation (3.2).

Close to the parasitic bias, assuming only the $1' - 3$ and $3' - 4$ resonant-tunneling

channels are active, the current I in the device can be written by modifying equation (2.87) as

$$I = \frac{|e|}{N_{\mathrm{p}}}\left(\frac{\rho_{33}}{\tau_{31}} + \frac{\rho_{22}}{\tau_2} + \frac{\rho_{33}}{\tau_{3,\mathrm{para}}} + \frac{\rho_{33}}{\tau_{3,\mathrm{para}}}\right) \tag{3.3}$$

where $\rho_{33}/\tau_{3,\mathrm{para}}$ is included twice since it has to account for current flow through both the channels across the injector barrier. In physical terms, the $3' - 4$ channel transports the electrons into the injector level of the next module instead of the same module, which makes its contribution to the current flow twice as much. For below threshold operation $\rho_{22}/\tau_2 = \rho_{33}/\tau_{32}$. In that case, the population in the upper radiative level 3 becomes

$$\rho_{33} = I\frac{N_{\mathrm{p}}}{|e|}\tau_{3,\mathrm{eff}} \tag{3.4}$$

where $\tau_{3,\mathrm{eff}}$ is specified by the expression

$$\frac{1}{\tau_{3,\mathrm{eff}}} = \frac{1}{\tau_{31}} + \frac{1}{\tau_{32}} + \frac{1}{\tau_{3,\mathrm{para}}/2} \tag{3.5}$$

Hence, $\tau_{3,\mathrm{para}}/2$ determines the effective reduction in the lifetime of the radiative level 3 due to a coupling with the upper parasitic level 4. Figure 3-4 shows plots of $\tau_{3,\mathrm{para}}/2$ for some typical parameters at the resonance condition $E_{3'4} = 0$. As seen from the plot (a), its value is approximately independent of the lifetime of level 4 because the $3' - 4$ resonant-tunneling is incoherent (section 2.4.3).

The injector transport in terahertz QCLs cannot typically be made coherent, as will become apparent in section 3.2.2. Therefore, the current flow in the device depends less critically on the lifetime of the upper radiative level. However, as can be noted from equation (3.4), for a given current I, the population in the radiative level 3 is proportional to its lifetime. Consequently, any reduction in the lifetime of the upper level directly limits the possibility of obtaining population inversion in a given design. This is reflected in the plots of Fig. 2-9(b), where the performance of a laser is shown to depend critically on τ_{31}, which is essentially the parameter that becomes smaller due to $\tau_{3,\mathrm{para}}/2$. From a practical perspective $\Delta_{3'4}$ is best kept $\lesssim 0.5$ meV for any such parasitic couplings in GaAs/$\mathrm{Al}_x\mathrm{Ga}_{1-x}$As terahertz QCL designs.

3.2.2 Parasitic current channels due to lower energy levels

For injector transport that is not coherent, a parasitic current channel due to coupling with higher energy levels, as discussed in section 3.2.1, primarily reduces the lifetime of the upper radiative state without increasing the current flow in the device significantly. Hence, even if the parasitic resonance happens below the design bias, as shown in Fig. 3-3, the device could still operate at the design bias. This is in contrast to a yet another, and often the more critical, type of parasitic coupling that exists for terahertz QCLs, which was first identified for resonant-phonon terahertz QCLs with the aid of Monte-Carlo simulations [30]. As the applied bias is increased from zero in a QCL, the injector level(s) first get aligned with energy levels that are below the upper radiative level in the adjacent module, before the desired alignment with the upper radiative level could be obtained. If at some particular lower bias the current flow in the device is greater than that at the design bias, the operating bias point for the QCL will lie in a NDR region, which is intrinsically unstable for a MQW structure as discussed in section 1.2.2. Even if such a *lower level parasitic channel* does not cause an early NDR, it may still conduct appreciable current close to the design bias, thereby *artificially* increasing the current density required to reach the lasing threshold. If such a channel is due to coupling of the injector level to the lower radiative level, it makes the threshold current density *actually* higher since the population of the lower radiative level is increased by direct injection. Moreover, any parasitic current close to the design bias makes fewer number of electrons available for current transport through the desired channel (i. e. the injector level to the upper radiative level tunneling), thereby reducing the maximum possible operating current beyond the lasing threshold.

The conduction band diagram at the lower level parasitic bias for the simple 3-level QCL design of Fig. 3-3 is shown in Fig. 3-5. For this design the parasitic coupling, which is characterized by the anticrossing energy $\Delta_{1'2}$, is relatively large due to only two barriers and a well that lie between the injector level $1'$ and the lower radiative level 2. The value of the current due to this parasitic channel follows the result derived

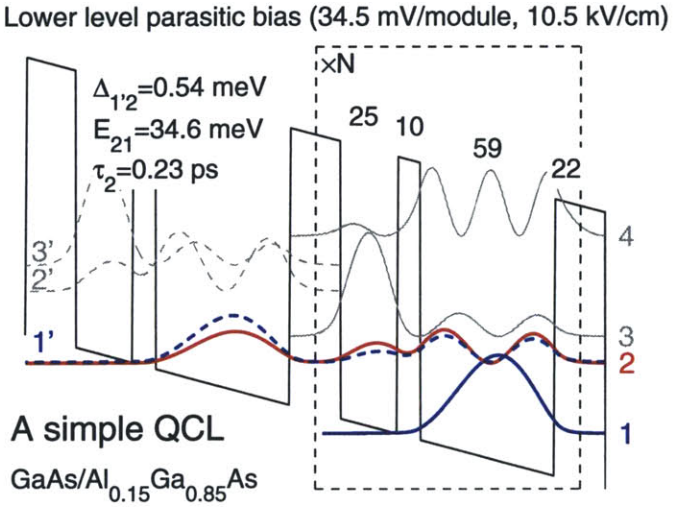


Figure 3-5: Conduction band diagram of the simple 3-level QCL design of section 2.4.2 at the lower level parasitic bias. The design bias diagram is shown in Fig. 3-3(a).

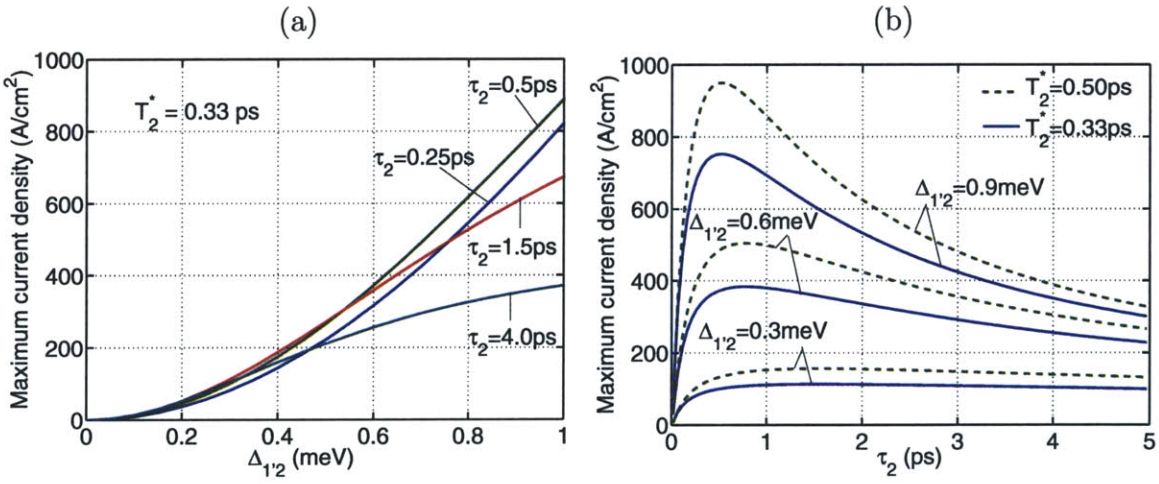


Figure 3-6: Current density (I/A_{ac}) at resonance ($E_{1'2} = 0$) for the lower level parasitic current channel, plotted for some typical parameters and a doping density $n_{\text{tot}}/V_{\text{ac}} = 5 \times 10^{15}$ cm^{-3} according to equation (3.6). A module length $L_{\text{mod}} = 50$ nm is used. The cross-sectional area of the active region A_{ac} can be chosen arbitrarily. For a fixed 3D *doping density*, the current density is $\propto L_{\text{mod}}$.

in equation (2.88)(b), and is written as

$$I_{\text{para}} = \frac{(n_1 + n_2)|e|}{2N_{\text{p}}} \left(\frac{1}{\tau_2}\right) \left[\frac{\left(\frac{\Delta_{1'2}}{\hbar}\right)^2 \tau_{\parallel,2}\tau_2}{\left(\frac{\Delta_{1'2}}{\hbar}\right)^2 \tau_{\parallel,2}\tau_2 + \left(\frac{E_{1'2}}{\hbar}\right)^2 \tau_{\parallel,2}^2 + 1} \right] \tag{3.6}$$

where, $n_1 + n_2 \approx n_{\text{tot}}$ at this bias assuming the population in level 3 is negligible, $\tau_{\parallel,2} = (1/2\tau_2 + 1/T_2^*)^{-1}$, and the level lifetime τ_2 is to be determined for wavefunctions calculated within a tight-binding scheme. The current density at resonance ($E_{1'2} = 0$) as calculated from this expression is plotted from some typical parameters in Fig. 3-6. For terahertz QCLs in GaAs/$Al_xGa_{1-x}As$ material system with optical phonon based depopulation, the anticrossing energy for the lower level parasitic current channels is typically kept $\lesssim 0.4$ meV. For a small value of anticrossing energy $\Delta_{1'2}$, the current density is relatively independent of τ_2 for $\tau_2 \gtrsim 0.5$ ps, as can be seen from Fig. 3-6(b). This is because resonant-tunneling in this case is incoherent, which makes the current density depend more on $\Delta_{1'2}$ rather than τ_2 (section 2.4.3). In general, the operation of a QCL depends critically on the anticrossing energies of various level couplings and there is always a trade-off between how low the parasitic anticrossings could be made without compromising on the peak current density and thus the maximum operating temperature for a given design.

The simple 3-level QCL of Fig. 3-3(a) has two relatively strong parasitic current channels as discussed in this and the previous section. Although lasing in this design cannot be ruled out (this design has not been experimentally realized yet), it is unlikely that this design will make a laser with a better temperature performance than the more complex designs that will be presented in the following sections. All of the designs discussed next are based on an optical phonon scattering based depopulation mechanism. The best performing designs out of those have so-far been with the *resonant-phonon* extraction scheme of Fig. 2-18, which allows an efficient selective depopulation of the lower radiative level while still maintaining a relatively large oscillator strength for the radiative transition. The primary challenge then is to keep the various parasitic couplings low, while still maintaining a large peak current

density, as will become more apparent from the upcoming discussions.

3.3 Terahertz QCLs with two-well injector, two-well active region, and resonant-phonon depopulation: the "FL" design

3.3.1 The FL design optimized for temperature performance: FL-C series

This first terahertz QCL from this group was named as FL175C and was developed by Benjamin S. Williams [168] after a few iterations of the so called "FL" designs where "F" stands for four wells per module, and "L" for LO-phonon depopulation (although this acronym has since lost this meaning, and subsequent modifications to the FL series have also had more than four wells per module). The FL175C design, and the subsequent designs in the FL-C series discussed in this section, are based on the resonant-phonon depopulation scheme as described in section 2.5.2. The conduction band diagram of the design is shown in Fig. 3-7. This design is a modified version of the one in Fig. 2-18 with an additional injector well (and correspondingly an additional injector level), which primarily helps in reducing the parasitic coupling of the injector levels $1'$ and $2'$ to the phonon-depopulation level 3. These couplings correspond to the lower level parasitic current channels of the type discussed in section 3.2.2, and are shown in Fig. 3-8. As will become apparent in the following discussions, each of the four barriers, as named in Fig. 3-7(b), play an important role in the optical and electrical behavior of the QCL and even 1 ML thickness variations can significantly change the laser characteristics.

Table 3.1 summarizes the experimental results obtained from various FL-C designs. Each entry in the table corresponds to a different MBE wafer and provides a range of values, which represent the majority of devices that were fabricated and tested for that wafer. "Mn" signifies the n-th modification in the FL-C series of

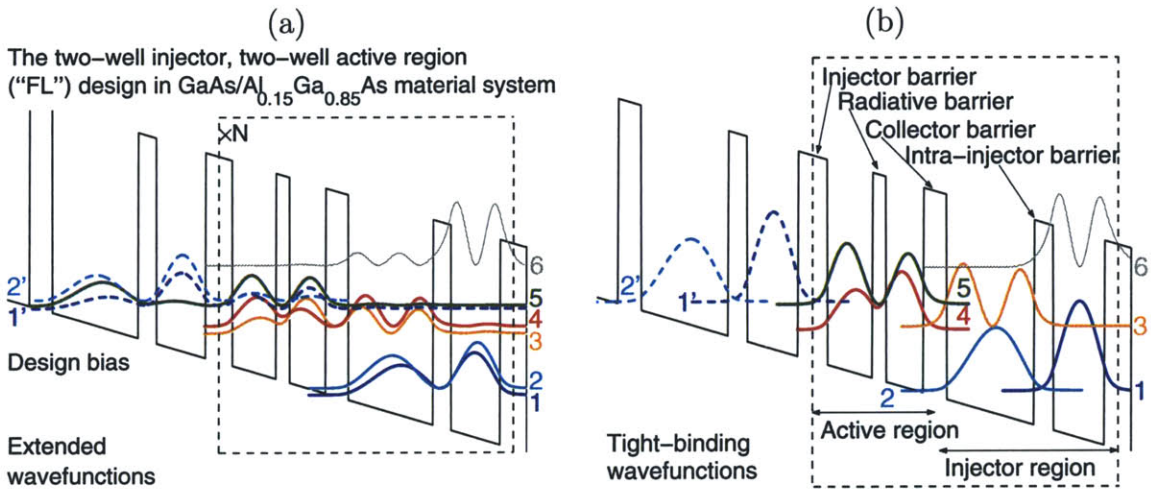


Figure 3-7: Conduction band diagram for the two-well injector, two-well active region "FL" design calculated in (a) the extended scheme, and (b) the tight-binding scheme at the design bias. For the tight-binding scheme, a module is further divided into three submodules, one for each of the two injector wells and the third for the two wells that form the active region.

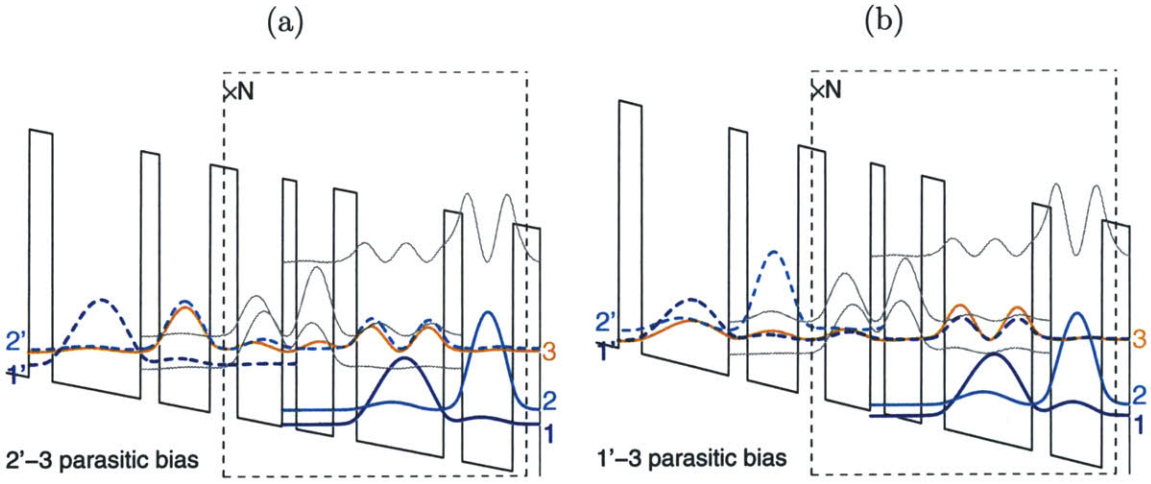


Figure 3-8: Conduction band diagrams for the two dominant (lower-level) parasitic current channels, corresponding to the (a) $2' \rightarrow 3$ and (b) $1' \rightarrow 3$ transport, for the FL design in Fig. 3-7.

FL-C series summary						
Device	$J_{th,5K}$ (A/cm^2)	J_{max} (A/cm^2)	$T_{max,pul}$ (K)	$T_{max,cw}$ (K)	ν (THz)	Waveguide
FL175C	800–900	950–1000	87	– –	3.35–3.55	SISP
FL178C-M1	625–675	1050–1150	137	67	3.75–3.90	MM (In-Au)
FL176C-M4	375–425	600-700	130	84	2.80–3.05	MM (Cu-Cu)
FL177C-M5	400–475	650-750	134	93	3.10–3.30	MM (In-Au)
FL176C-M4-2	325–400	450–600	119	98	2.70–2.95	MM (Cu-Cu)
FL178C-M6	400–450	700–775	131	83	2.80–3.10	MM (In-Au)
FL178C-M7	425–500	650–850	164	117	2.85–3.20	MM (Cu-Cu)
FL178C-M7-P	425	625	115	n.a.	2.90	MM (In-Au)
FL177C-M8	350	500	100	50	3.05	MM (In-Au)
FL179C-M9	650	1000	135	n.a.	2.65–2.90	MM (In-Au)
FL179C-M9-2	500–600	800–1000	163	85	2.50–3.10	MM (Cu-Cu)
FL178C-M10	350-450	575–825	169	109	2.60–3.00	MM (Cu-Cu)

Table 3.1: Summary of experimental results from FL-C series of terahertz QCLs. Each entry in the table corresponds to a different MBE wafer. "-2" implies second growth of the same design. The devices are listed in chronological order of development from top to bottom. FL175C was never fabricated in MM waveguides, hence only SISP waveguide data is reported. $T_{max,cw}$ for FL178C-M7-P and FL179C-M9 is not available, although these devices did operate in continuous-wave (cw) mode. FL178C-M7-P is the same as FL178C-M7 except that the former had layers that were growth thicker by 2.5% as compared to the latter, and that two growth interrupts of 15 seconds each per module were incorporated during its growth. For all other wafers there were no such growth interrupts. FL176C-M4 is the same as FL176C-M4-2 except that its layer thicknesses were smaller by 3.5% to that of FL176C-M4-2, whereas its doping was kept the same. FL179C-M9-2 was grown with similar layer thicknesses to FL179C-M9 except that the former was doped to 2.3×10^{16} cm^{-3} in the wide-well of the injector, as opposed to 1.9×10^{16} cm^{-3} for the latter; however, there seems to be an ambiguity in the doping values since both devices had similar experimental results. All the results shown are for a ~ 10 μm thick active region. It may be noted that the doping values used for these designs (typically 5.5×10^{15} cm^{-3} average 3D doping or 3×10^{10} cm^{-2}/module average 2D doping) are a carry-over from previous work by Xu [182] and Williams [166] with non-lasing designs, although these values have since been shown to be close to optimum for obtaining the best temperature performance in an experimental study, which was conducted on terahertz QCLs based on the FL177C-M5 design, albeit with different doping values [107].

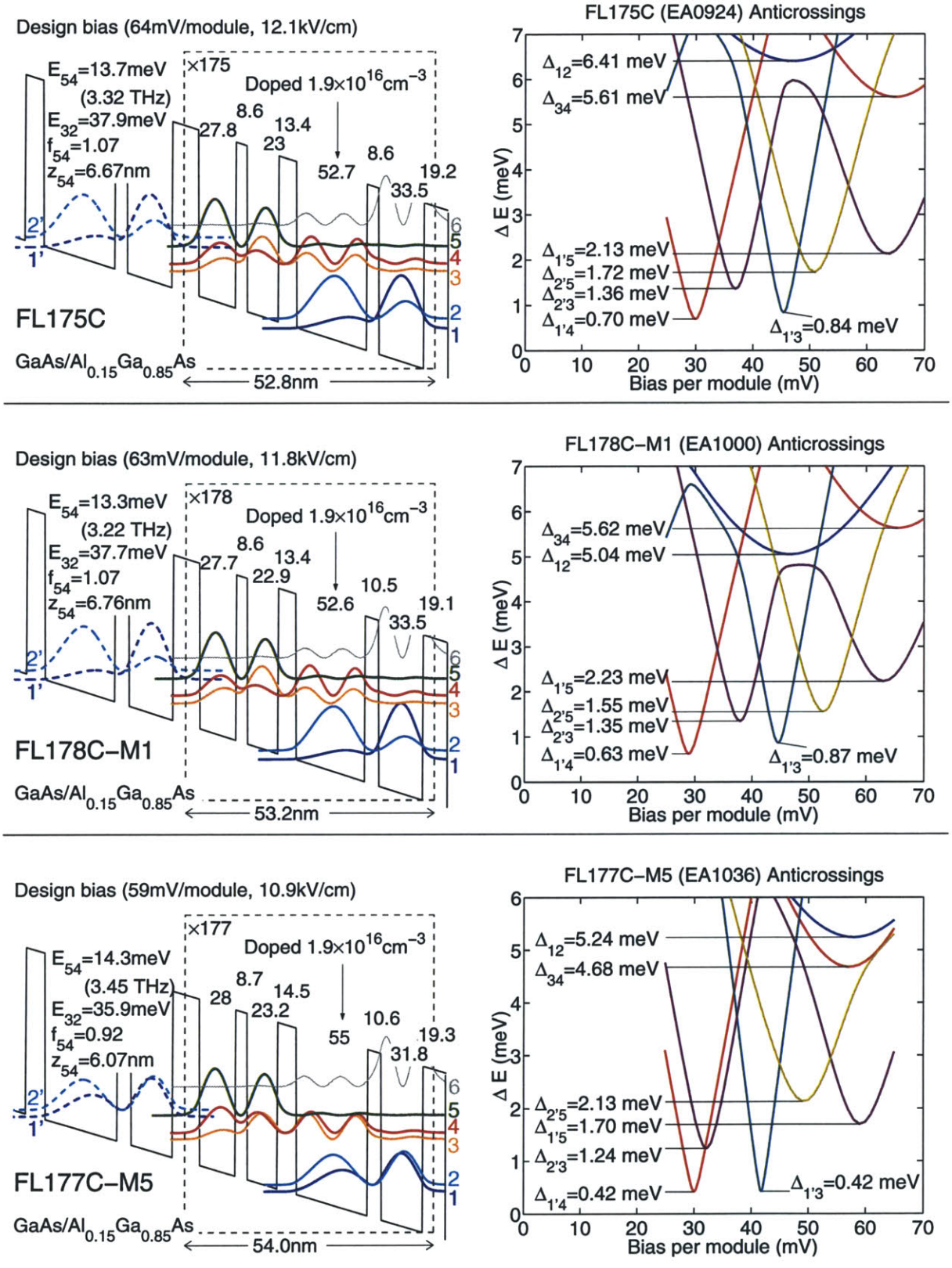


Figure 3-9: Conduction band diagram and the anticrossing plots for the FL-C series of devices for layer thicknesses as grown.

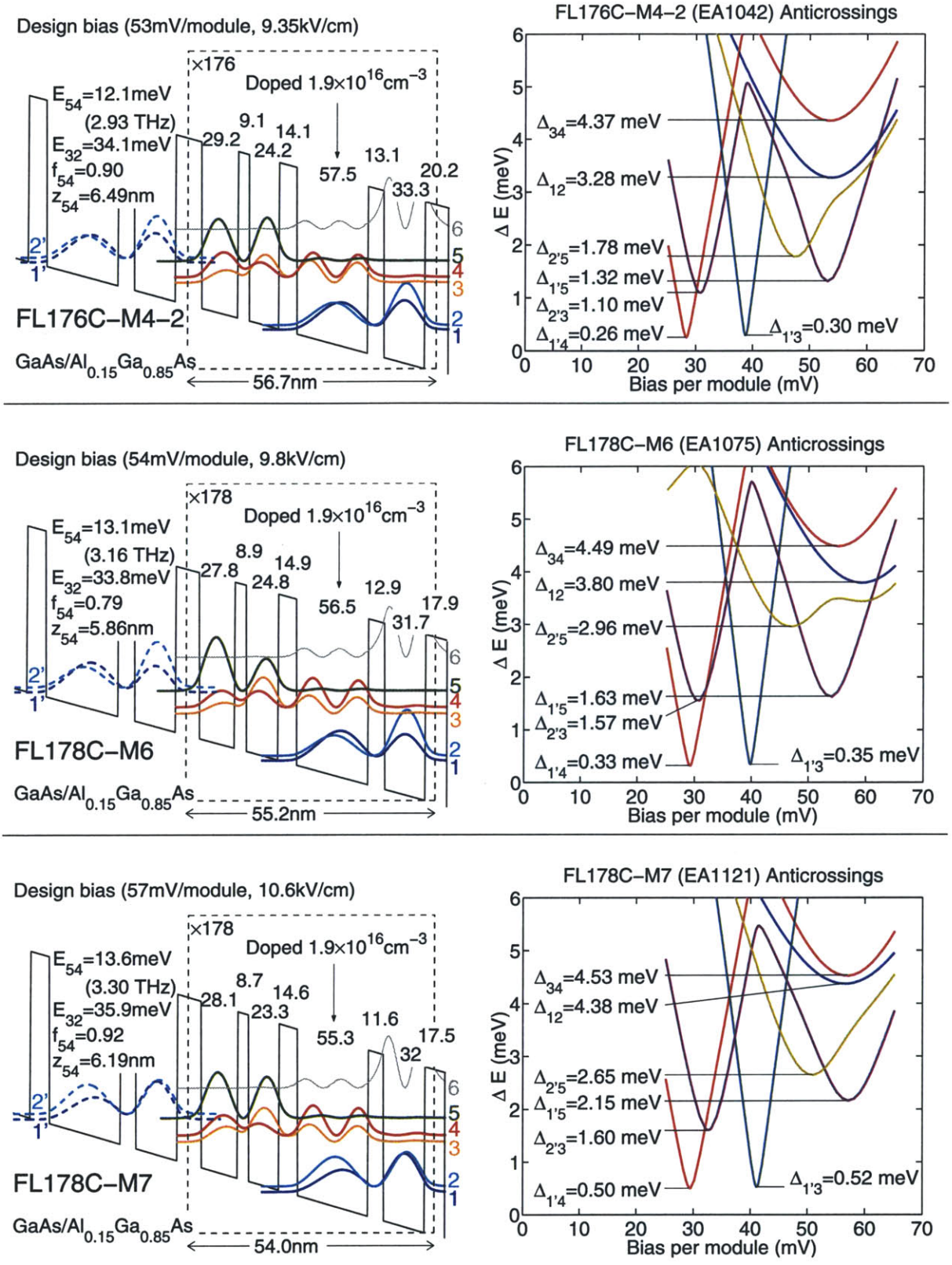


Figure 3-10: Conduction band diagram and the anticrossing plots for the FL-C series of devices for layer thicknesses as grown (contd.).

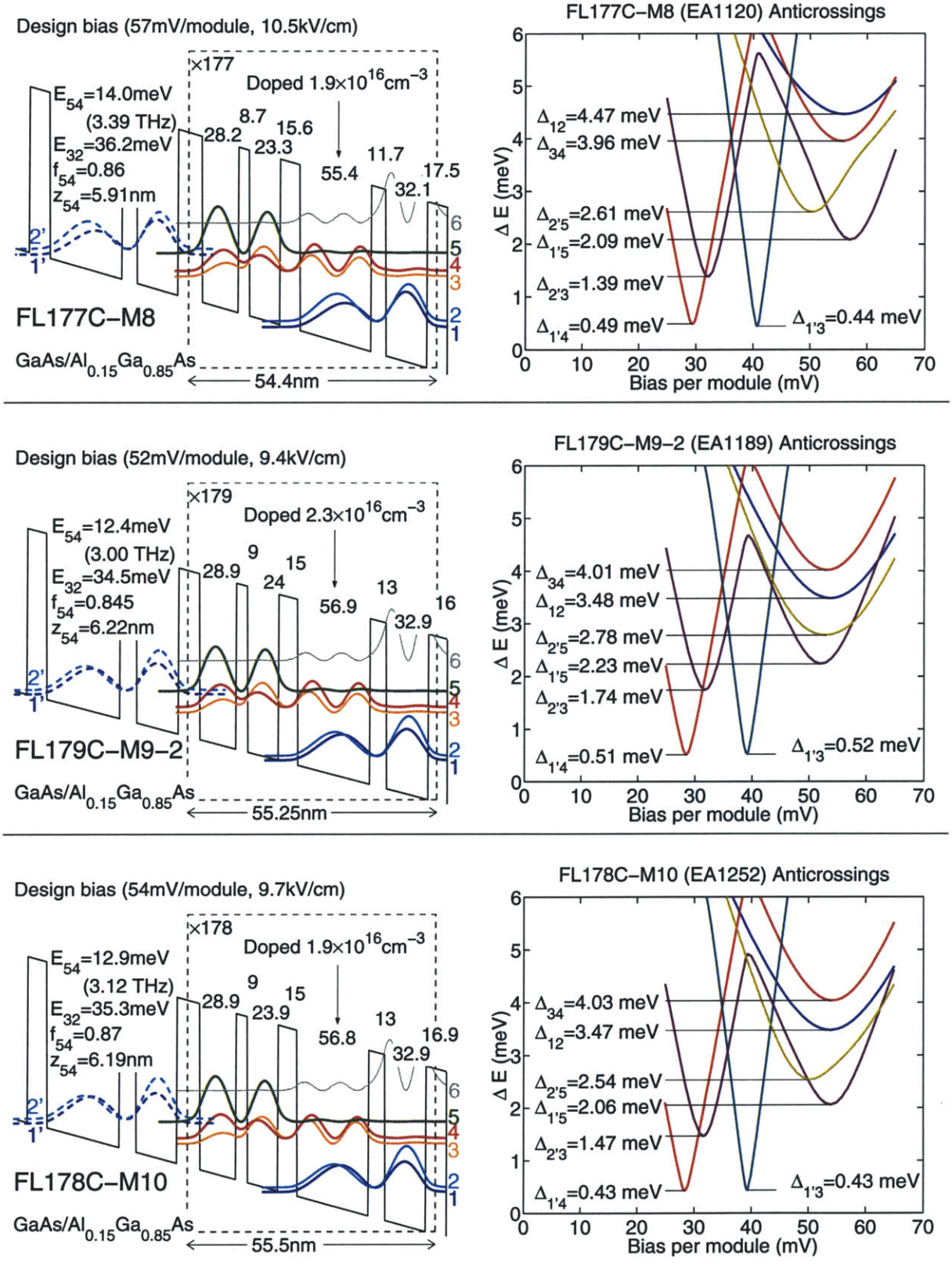


Figure 3-11: Conduction band diagram and the anticrossing plots for the FL-C series of devices for layer thicknesses as grown (contd.).

designs[1]. The conduction band diagrams at the design bias and the corresponding anticrossing plots for each design are shown in Figs. 3-9, 3-10, and 3-11. The anticrossing plots of the type shown in these figures were first used by Williams [166] to show the behavior of various level-anticrossings with respect to the applied bias compactly in a single figure. Such plots are often useful in the design process since the location of various anticrossings in bias space is one of the more important criteria, in addition to their values, which determine the performance of the laser.

Whereas various different designs from the FL series have achieved lasing in the spectral range of $2.1-5.0$ THz, the FL-C series of designs were all designed to operate close to 3 THz since they were based on successive modifications to improve their temperature performance. The center frequency of the gain bandwidth for these designs is approximately determined by the radiative energy E_{54} in the tight-binding scheme of Fig. 3-7(b). The radiative transition from level $5 \rightarrow 4$ is close to being vertical since the design bias is close to the $5-4$ anticrossing bias. Consequently, E_{54} ($\sim \Delta_{54}$) depends strongly on the thickness of the radiative barrier. All the FL-C series designs were designed with a 9 ML radiative barrier, but due to overgrowth/undergrowth in MBE, the actual thicknesses for each design were different and so are the experimentally measured lasing frequencies. All these designs were grown in the same MBE machine; consequently, a very good correlation can be noticed between the measured and the designed laser frequencies. The measured values are typically centered at a value $0.1-0.2$ THz smaller than that predicted by the design bias calculation (this is true for all designs except in FL175C and FL178C-M1, which have a considerably different level alignment as compared to the other designs). Note that the various design parameters such as level anticrossings and their location in the bias-space depend critically on the barrier height, which can vary from one MBE machine to another due to slightly varying calibrations of Al concentration in the $Al_xGa_{1-x}As$ barriers. Thus, a terahertz QCL design optimized for growth in one MBE machine may not work as well when grown in a different one.

[1]Note that a design named as FL177C-M3 was never grown and another design named as FL180C-M2 was grown but the processed devices did not conduct current for an unexplored reason.

The results reported in Table 3.1 are for devices fabricated with metal-metal (MM) waveguides [170], except for FL175C devices which were only fabricated with semi-insulating surface-plasmon (SISP) waveguides [89]. The electromagnetic characteristics of both types of waveguides for terahertz QCLs are discussed in greater detail in Ref. [88]. Most of the results in this thesis focus on MM waveguides since they have been shown to provide lower waveguide loss and better mode confinement for terahertz QCLs as compared to SISP waveguides, and hence a better temperature performance. The fabrication and operating principles of MM waveguides will be discussed in chapter 4. It may be noted, however, that much greater power can be out-coupled from SISP waveguides, which is why terahertz QCLs for higher optical power are developed using SISP waveguides, as will be discussed in section 3.6.

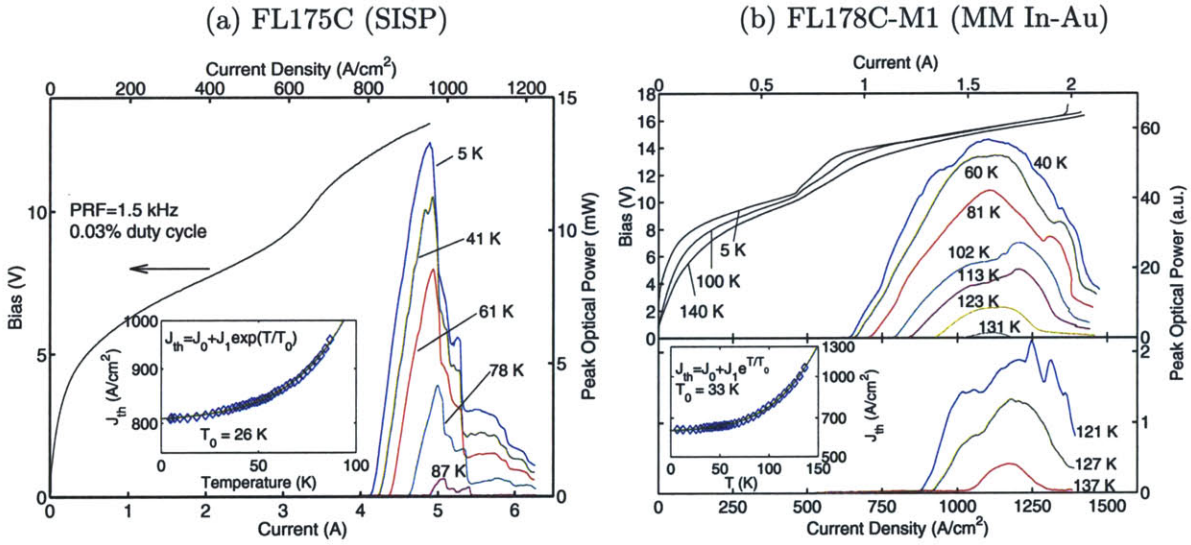


Figure 3-12: Pulsed (duty-cycle $< 0.05\%$) light-current (L-I) and current-voltage (I-V) characteristics measured from (a) a 200 μm $\times$ 2.56 mm SISP FL175C ridge laser with its rear facet high-reflectivity (HR) coated [169], and (b) MM FL178C-M1 ridge lasers of different dimensions [171] (without HR coating). The L-Is were recorded with a He-cooled Ge:Ga photodetector. The peak power in plot (a) was measured using a calibrated thermopile detector (ScienTech Model 360203).

Figure 3-12 summarizes the experimental results from the best of FL175C and FL178C-M1 lasing devices. Detailed results for these devices can be found in chapter 7 of Williams [166]. It was determined from Monte-Carlo electrical transport simulations [30] that the relatively high threshold current densities in these devices

($J_{\mathrm{th,5K}} \sim 800$ A/cm^2 for FL175C and $J_{\mathrm{th,5K}} \sim 625$ A/cm^2 for FL178C-M1) were due to the strong lower level parasitic current channel caused by $1' - 3$ coupling as shown in Fig. 3-8(b). This can be seen as a shoulder in the I-Vs just before lasing threshold for both of these devices. The intra-injector barrier in FL178C-M1 was thickened by $\sim$ 2 ML as compared to FL175C, which reduced this parasitic coupling slightly, thereby reducing the threshold current densities ($J_{\mathrm{th,5K}}$) and allowing a greater range of lasing in current ($J_{\mathrm{max}} - J_{\mathrm{th,5K}}$). Consequently, a modest increase in temperature performance was observed for FL178C-M1 SISP devices, which had a $T_{\mathrm{max}} = 92$ K in pulsed mode as compared to $T_{\mathrm{max}} = 87$ K for FL175C. However, the pulsed T_{max} for FL178C-M1 increased radically to 137 K by the use of more efficient metal-metal waveguides, as shown in Fig. 3-12(b). Note that a phenomenological fit of the variation of J_{th} with the heat-sink temperature T to the expression $J_{\mathrm{th}} = J_0 + J_1 \exp(T/T_0)$ is also indicated. Although no physical justification for such a fit has yet been determined, such an expression typically expresses $J_{\mathrm{th}} - T$ variation well for both mid-infrared and terahertz QCLs, and the characteristic temperature T_0, along with the lasing range in current $J_{\mathrm{max}} - J_{\mathrm{th,5K}}$ are the two key parameters that characterize the temperature performance of a given design. The low temperature threshold current density ($\sim J_0 + J_1$) for mid-infrared QCLs is usually a good measure of the waveguide and the mirror losses. However, for the phonon-depopulated terahertz QCLs discussed in this thesis, this value is artificially enhanced by the parasitic current channels, and hence is not indicative of the losses in the structure. Consequently, devices with varying cavity dimensions typically have the same low-temperature threshold current densities. For a given terahertz QCL design, the value of T_0 is larger for a lower frequency design, the reasons for which will be explained in section 3.3.4.

The original two of the FL-C lasers (FL175C and FL178C-M1) had relatively high threshold current densities, which, in addition to the fact that SISP waveguides typically needed to be wider than 100 μm to keep the waveguide losses low [88], did not allow continuous-wave operation in SISP waveguides for both of these lasers. The maximum duty-cycle obtained from FL175C SISP devices was $\sim$ 60%, whereas some

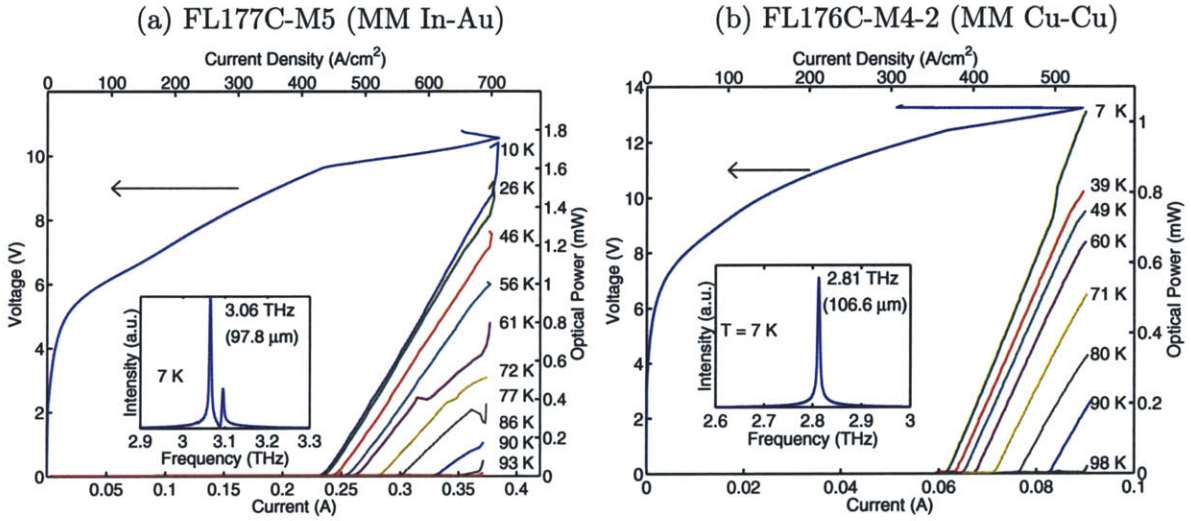


Figure 3-13: CW *I-V* and *L-I* characteristics measured from (a) a FL177C-M5 ridge laser of dimensions 40 μm $\times$ 1.35 mm fabricated with a In-Au MM waveguide, and (b) a FL176C-M4-2 ridge laser of dimensions 25 μm $\times$ 0.67 mm fabricated with a Cu-Cu MM waveguide. The *L-I* characteristics are measured with a room-temperature pyroelectric detector (Molectron model P4-42), and the peak power is calibrated using a room-temperature thermopile detector (ScienTech model AC2500). Typical cw spectra are also shown in the inset on a linear scale, which are measured with a Nicolet Magna-IR 850 spectrometer in linear scan mode using a room-temperature deuterated triglycine sulfate (DTGS) pyroelectric detector. The linewidth of the laser modes is limited by the instrument resolution of 0.125 cm^{-1} (3.75 GHz). The substrate thickness was 140 μm for the FL177C-M5 laser, and 210 μm for the FL176C-M4-2 laser. Without any corrections in the detected power, the single-facet wall-plug efficiency for the FL177C-M5 laser is $\sim$ 0.05%, and the slope-efficiency is $\sim$ 12 mW/A ($\sim$ 1 photon/electron). For the FL176C-M4-2 laser, the wall-plug efficiency is $\sim$ 0.1%, and the slope efficiency is $\sim$ 35 mW/A ($\sim$ 3.2 photons/electron).

of the FL178C-M1 SISP lasers operated in cw mode for a few seconds only before eventually being destroyed due to heating. The MM FL178C-M1 devices (processed using the newly developed In-Au fabrication technique [170]) typically stopped lasing after the first $10-20$ μs of the applied pulses due to a high thermal resistance of the In-Au metal-metal bonding interface.

In the next batch of FL-C designs, which included FL177C-M5 and FL176C-M4, the injector doublet consisting of levels $1'$ and $2'$ was brought more into anticrossing at the design bias, rather than being "downwards-diagonal" (as can be seen from the anticrossing plots in Figs. 3-9 and 3-10). Moreover, the collector barrier in both these new designs was thickened by ~ 1 ML. As a consequence of these modifications, the parasitic anticrossing $\Delta_{1'3}$ became smaller, which in turn reduced the threshold current densities in these designs. This, in addition to an improved In-Au wafer bonding technique, which will be discussed in chapter 4, helped realize the first QCL operating above the technologically important liquid nitrogen temperature [94] with the FL177C-M5 design, as shown in Fig. 3-13(a). The original FL178C-M1 devices also lased in cw mode when reprocessed with this improved In-Au bonding technique; however, the maximum temperature of cw operation was low due to their higher current densities as indicated in Table 3.1. Subsequently, an even more robust technique using Cu-Cu thermocompression wafer-bonding, which had been already developed elsewhere for Silicon wafer boding [38], was used to realize MM waveguides for terahertz QCLs. As compared to In-Au, Cu-Cu fabrication is more demanding yet a more reliable fabrication process and has an inherent advantage due to the higher thermal conductivity of Copper. Hence, an improved cw performance was obtained from FL176C-M4-2 devices, which were fabricated with the Cu-Cu bonded MM waveguides, as is evident from Fig. 3-13(b). Note that the threshold current densities for FL176C-M4-2 lasers were typically lower than for FL177C-M5, since the intra-injector barrier in FL176C-M4-2 was thicker by ~ 2.5 ML as compared to FL177C-M5, which further reduced the parasitic anticrossings. However, the advantage of low $J_{\mathrm{th,5K}}$ in FL176C-M4-2 was offset by lowering of J_{max} as well due to the occurrence of an early NDR, the reasons for which will be explained later in this section.

(a) FL177C-M5 (In-Au), 40 μm $\times$ 1.35 mm (b) FL176C-M4-2 (Cu-Cu), 40 μm $\times$ 0.47 mm

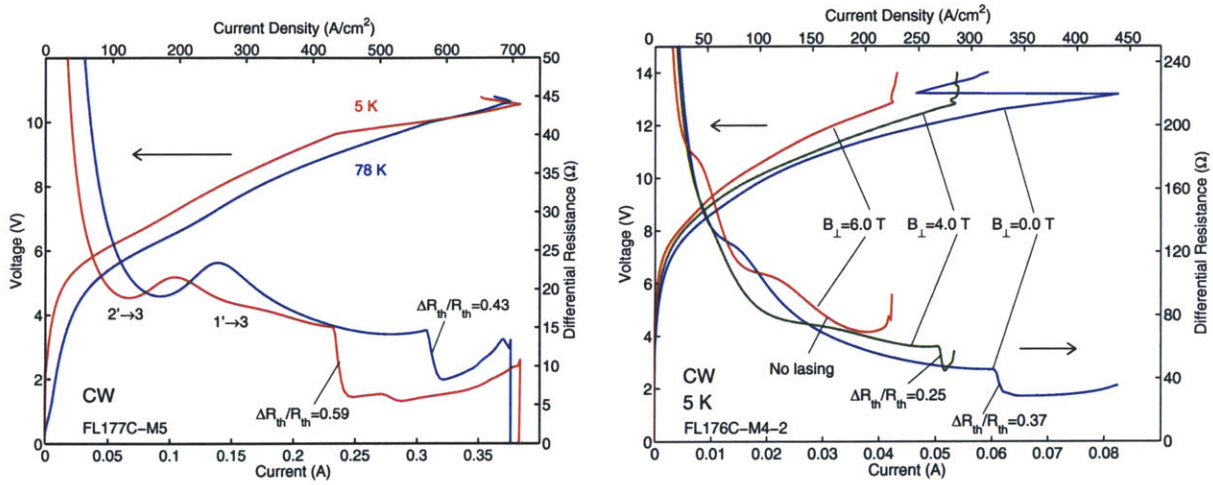


Figure 3-14: (a) V-Is and $\mathcal{R}$-Is ($\mathcal{R} \equiv$ differential resistance) for the same FL177C-M5 laser as in Fig. 3-13 measured at heat-sink temperatures of 5 K, and 78 K, respectively. $\mathcal{R}$ is experimentally measured by superimposing a small ($\sim$ 10 mV peak-peak), low-frequency (40 kHz) sinusoidal signal on the applied DC (zero-frequency) bias, and by measuring the time-varying small-signal current flowing through the device with a lock-in-amplifier. The discontinuity in $\mathcal{R}$ at the lasing threshold is also indicated, which is characterized by the parameter $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ as defined in section 2.4.2.
(b) V-Is and $\mathcal{R}$-Is for a FL176C-M4-2 laser immersed in liquid Helium in a superconducting magnetic dewar, measured at different values of the magnetic field $B_\perp$ applied perpendicular to the layers of the MQWs.

The differential resistance $\mathcal{R}$ in a laser could be measured much more easily in cw operation as compared to pulsed operation. A direct experimental measurement considerably reduces the noise in the data when compared to the values obtained from a mathematical derivative of the measured I-Vs, thereby revealing finer features in the slope of the I-V. The results obtained from one such measurement of a FL177C-M5 laser are shown in Fig. 3-14(a). The discontinuity in differential resistance at threshold, which was discussed in sections 2.4.2 and 2.5.1, can be directly measured, thus providing valuable insight into the performance of the QCL design. The variation of $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ with temperature will be discussed further in section 3.3.4. The valleys in the $\mathcal{R}$ plots that occur below the lasing threshold are indicative of the various parasitic current channels. These valleys happen at a bias slightly below the bias at which the corresponding parasitic resonance happens, as determined by the bias at which the slope of the Lorentzian that models the resonant-tunneling current goes

to zero. The two of the dominant parasitic current channels of Fig. 3-8 that happen for the FL-C designs, are believed to be the ones as indicated in Fig. 3-14(a). As can be noticed, the device does not need much more current past the parasitic bias points before it starts to lase, which indicates that the parasitic current channels put a lower limit on the threshold current densities that could be obtained in this design. As can be noticed from the 78 K measurement in Fig. 3-14(a), the fine features in the $\mathcal{R}$-I disappear at the higher temperature since the valleys becomes broad. Possible reasons for this will be discussed in section 3.3.4.

Figure 3-14(b) shows results from a similar measurement for a FL176C-M4-2 laser, when a magnetic field is applied perpendicular to the layers of the MQWs. For this measurement, only the transport results are available since the magnetic dewar used in this measurement did not have provisions for coupling out the light. The in-plane electronic dispersion is split into discrete energy levels (Landau levels) upon application of $B_\perp$ field due to in-plane quantization. The modification of the scattering times due to Landau levels has been exploited to achieve lasing in QCL structures down to a frequency of 1.39 THz [138, 139]. Landau level splitting has also been used to do magnetotunneling spectroscopy of the energy anticrossings in the T65 structure of section 3.1 [167, 166]. However, such a spectroscopy is difficult to perform on the more complex FL design since many different resonances are active at a given bias. Nevertheless, the enhancement of the features due to parasitic current channels upon application of a high magnetic field, which is possibly due to reduced intersubband scattering and dephasing, can be seen from the $\mathcal{R}$-I plot for $B_\perp = 6.0$ T in Fig. 3-14(b). Even though the upper state lifetime may be enhanced for some particular $B_\perp$, the performance of this laser was found to monotonically degrade as $B_\perp$ increases, and eventually lasing ceased, as seen from the $B_\perp = 6.0$ T measurement. The sharp decline in $J_{\max}$ with $B_\perp$, possibly due to the pinching of the design bias current channel, is believed to be the cause of this behavior. Hence, for an upper radiative state that is populated by resonant-tunneling, application of $B_\perp$ is less likely to aid lasing in a QCL structure. In contrast, the transport in the low frequency terahertz QCL of Ref. [138] is more likely determined by intersubband scattering rather

than resonant-tunneling, in which case in-plane quantization can help in obtaining a longer upper state lifetime and narrower linewidths, which aids lasing.

(a) FL176C-M4-2 (Cu-Cu)
Pulsed and cw L-Is

(b) FL176C-M4-2 (Cu-Cu)
CW I-V, $\mathcal{R}$-V, and L-Vs

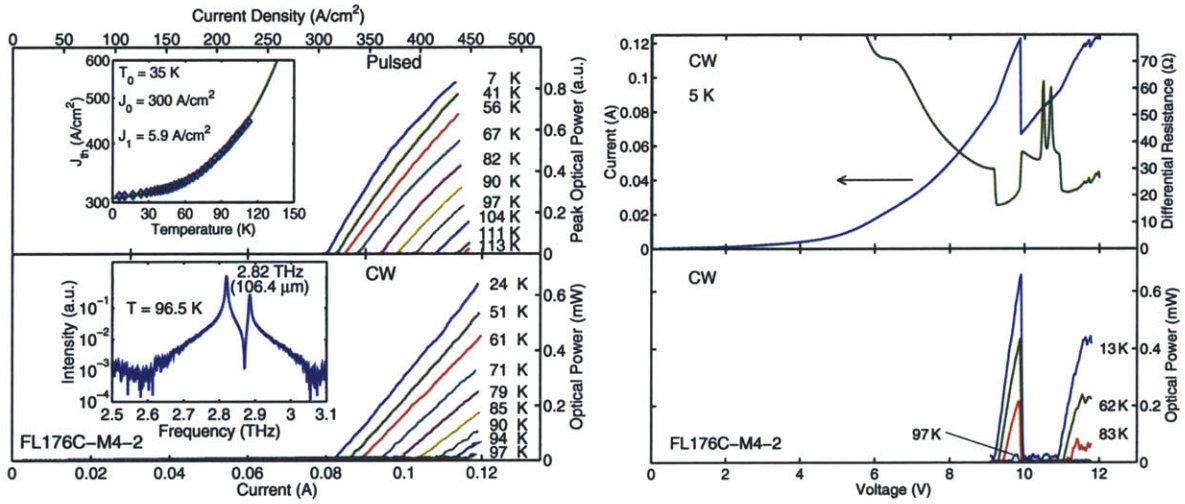


Figure 3-15: (a) Pulsed and cw L-Is measured from a 25 μm $\times$ 1.04 mm FL176C-M4-2 (MM Cu-Cu) ridge laser. A cw spectrum taken close to the maximum operating temperature with a room-temperature DTGS detector is shown in the inset of the lower panel on a log scale. (b) The upper panel shows cw I-V and $\mathcal{R}$-V measured at 5 K for the same device as in (a). The lower panel shows light-voltage (L-V) characteristics, as opposed to the L-I characteristics of plot (a). Note that the L-I curves in plot (a) correspond only to the lower-bias (< 10 V) lasing region.

FL176C-M4-2 had a relatively thick intra-injector barrier, which made the injector doublet consisting of levels $1'$ and $2'$ "tighter" as characterized by a smaller anticrossing energy Δ_{12}. To obtain the maximum lasing range in current, the current should keep increasing monotonously as the device is biased up from $2' - 5$ anticrossing to the $1' - 5$ anticrossing, although it is not easy to quantify the minimum values of these anticrossings unless full transport calculations are done with the inclusion of optical feedback, similar to that in section 2.4.2. It is believed that FL176C-M4-2 devices had an early NDR that occurred before the $1' - 5$ alignment could be reached, thereby limiting $J_{\max}$ to a low value. This can be seen from Fig. 3-15, which shows the light-voltage (L-V) characteristics of one of the FL176C-M4-2 lasers that had two separate lasing regions. This was typical of many such devices tested and is attributed to the fact that once NDR is reached even for only a few of the modules

in a QCL, a uniform bias cannot be maintained across the structure due to high-field domain formation within the multiple-quantum wells. Consequently, beyond the first NDR, the device behavior is unpredictable and any lasing afterward, even if it is due to $1' - 5$ injection, does not last to higher temperatures since all the modules do not contribute to gain beyond the first occurrence of the NDR region.

The FL177C-M5 and FL176C-M4-2 lasers achieved record high cw operating temperatures at the time of their development [94]. This was primarily because the threshold current densities were lowered by cutting down the parasitic current channels, and also because of significant improvement in MM waveguide fabrication. Moreover, MM waveguides could be made much narrower without significantly affecting either the waveguide or the mirror losses, thereby improving heat-removal from the cavities during cw operation. For example, for a 10 μm thick active region, MM waveguides down to $\sim$ 40 μm widths do not cause any significant reduction in $J_{\max} - J_{\text{th,5K}}$. However, $J_{\max} - J_{\text{th,5K}}$ typically reduces by $50 - 150$ A/cm^2 for $\sim$ 25 μm wide waveguides, which is less so because of increase in $J_{\text{th,5K}}$ but more because of a reduction in $J_{\max}$. These and other results from very small MM cavities will be discussed further in chapter 4. In contrast, SISP waveguides are typically designed to be wider than 100 μm to keep their waveguide losses low [88], and thus their temperature performance in cw operation is limited as will be shown in section 3.6.2.

Despite improvements in cw performances of FL177C-M5 and FL176C-M4-2 lasers, their maximum temperatures in pulsed operation (FL177C-M5: $T_{\max} \sim 125 - 135$ K, FL176C-M4-2: $T_{\max} \sim 110 - 120$ K) were worse than those for the previous generation FL178C-M1 design, since their peak current densities were also reduced in conjunction with the threshold current densities. In the next generation of lasers, FL178C-M7, FL179C-M9-2, and FL178C-M10, the injector barrier was thinned by $2 - 3$ ML, whereas the collector barrier was further thickened by $0.5 - 1$ ML to make the injection process more coherent, while still keeping the parasitic current channels in check. Significant improvement in temperature performance has since been obtained for all the three latter designs as can be noted from Table 3.1, although it is debatable whether the improved results were entirely due to the design modifications

as will be discussed in section 3.3.3.

Pulsed and cw measurement results from the best performing FL178C-M7 laser are shown in Fig. 3-16 [174]. The inset in Fig. 3-16(b) shows the cw spectra taken at $T \sim 10$ K for different bias points. It is typical to see changes in the measured spectrum corresponding to the changes in the slope of the L-Is. These are possibly due to a shift in the peak-gain frequency as the injection switches from $2'-5$ to $1'-5$ with higher bias, or they can also be caused by changes in the gain spectrum due to the injector and collector anticrossings as described in section 2.5. Any such changes in the optical behavior also induce relatively small variations in the I-V that are caused due to re-adjustments of the population inversion. Such variations are more clearly noticed in the differential resistance measurements (for this laser the $\mathcal{R}$-I plots are shown in Fig. 3-23). The value of $T_{\mathrm{max,cw}} = 117$ K for FL178C-M7 is the highest such reported value for any terahertz QCL. CW measurement results from a yet another FL178C-M7 laser are shown in Fig. 3-17. As will become apparent in section 3.6.1, a shorter cavity device yields better values of the slope-efficiency and the wall-plug efficiency. The corresponding values of ~ 55 mW/A and $\sim 0.2\%$ for the laser in Fig. 3-17 are the highest measured values for resonant-phonon QCLs with metal-metal waveguides. The subsequent modifications to the FL178C-M7 design did not yield any significant improvements. In pulsed operation a value of $T_{\mathrm{max,pul}} = 169$ K for a FL178C-M10 laser (section 3.3.2, Fig. 3-18a) is the highest such reported value for any terahertz QCL. It is difficult to pick the best design out of FL178C-M7, FL179C-M9-2, and FL178C-M10 since the variations in their results (Table 3.1) are well within the processing and growth fluctuations.

3.3.2 Terahertz QCLs with active region thinner than 10 μm

Terahertz QCLs are typically made with a much thicker active region ($10-15$ μm) as compared to that for mid-infrared QCLs ($2-4$ μm) due to the longer wavelengths involved. For SISP waveguides, a thicker active region obtains a greater value for the mode-confinement factor Γ, which keeps the mirror losses $\alpha_{\mathrm{m}}/\Gamma$ low, and it also limits the waveguide losses α_{w} by minimizing mode overlap with the semi-insulating

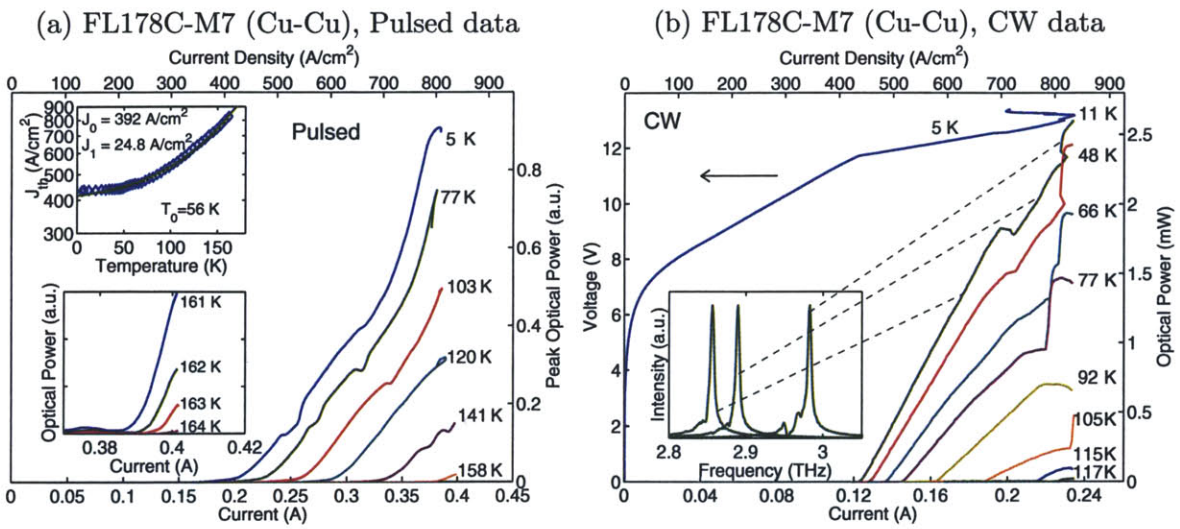


Figure 3-16: (a) Pulsed L-Is measured from a 48 μm $\times$ 0.99 mm FL178C-M7 (MM Cu-Cu) ridge laser. This device lased up to 98 K in cw mode with maximum cw power of 2.1 mW at 5 K. (b) CW L-Is, I-V, and linear-scale spectra measured from a 23 μm$\times$1.22 mm FL178C-M7 (MM Cu-Cu) ridge laser. The values of $T_{\mathrm{max,cw}} \sim 117$ K for this laser is the highest such value reported for any terahertz QCL. The optical power was collected from a single facet, and the reported value is as measured by the thermopile detector without correcting for any losses from the device to the detector. This device lased up to 158 K in pulsed mode. The substrate was lapped down to a thickness of 170 μm to improve heat-sinking. Without any corrections in the detected power, the single-facet wall-plug efficiency for this laser is $\sim 0.1\%$, and the slope-efficiency is ~ 27 mW/A (~ 2.3 photons/electron).

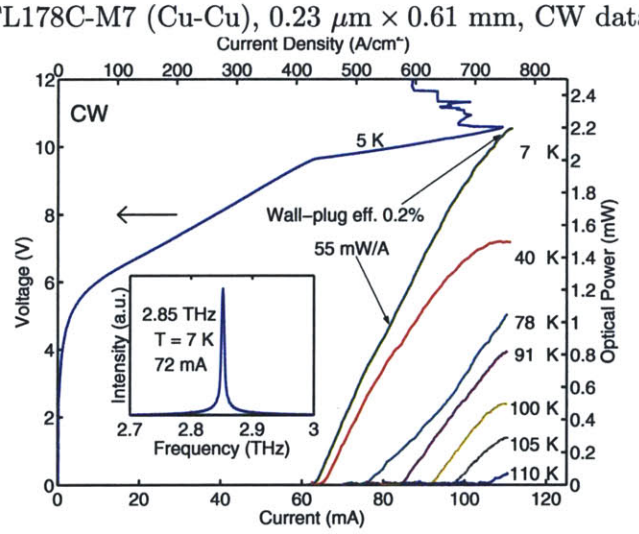
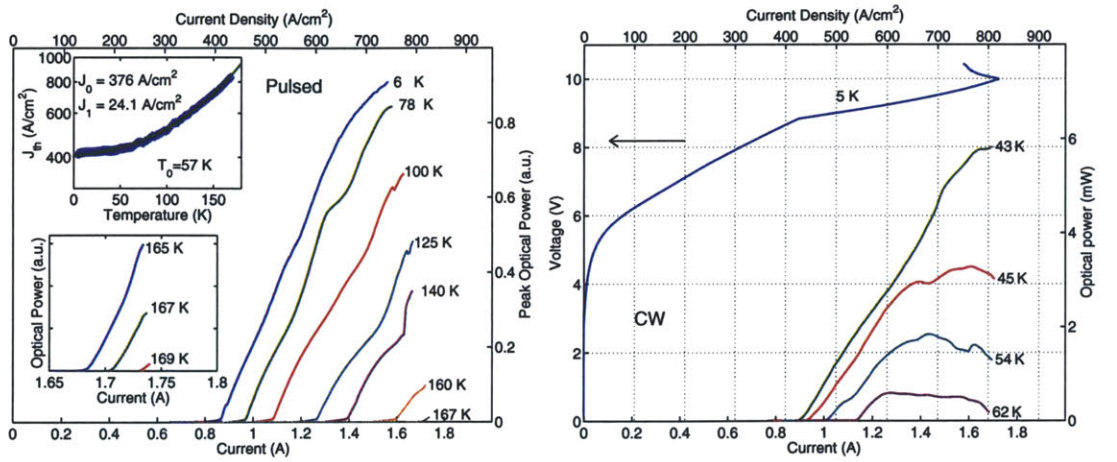
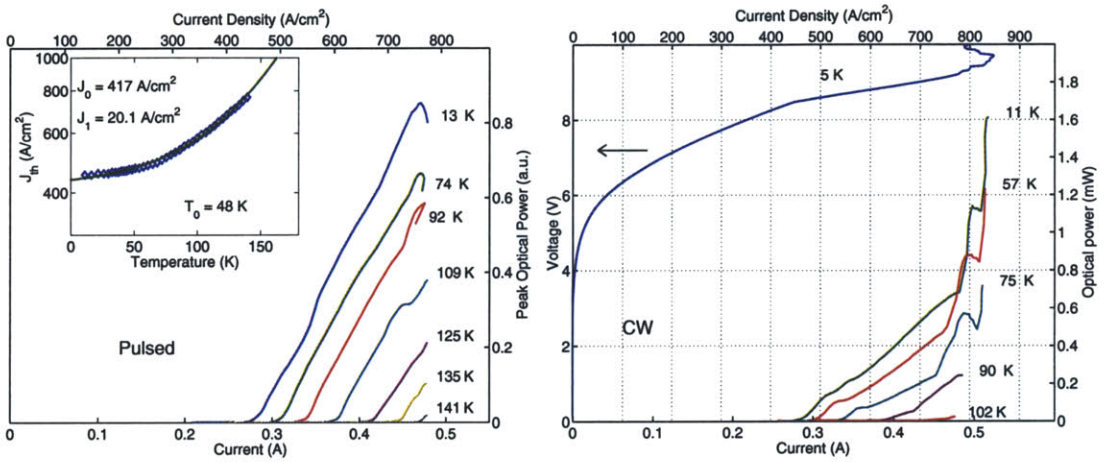


Figure 3-17: CW L-Is from a shorter cavity FL178C-M7 laser ($T_{\mathrm{max,pul}} \sim 147$ K and $T_{\mathrm{max,cw}} \sim 111$ K) that had higher slope-efficiency and wall-plug efficiency. Without any corrections in the detected power, the single-facet wall-plug efficiency for this laser is $\sim 0.2\%$, and the slope-efficiency is ~ 55 mW/A (~ 4.6 photons/electron).

(a) FL178C-M10 (Cu-Cu), 100 μm $\times$ 2.1 mm, 10 μm thick



(b) FL178C-M10 (Cu-Cu), 60 μm $\times$ 1.03 mm, 5.1 μm thick



(c) FL178C-M10 (Cu-Cu), 80 μm $\times$ 1.03 mm, 2.8 μm thick

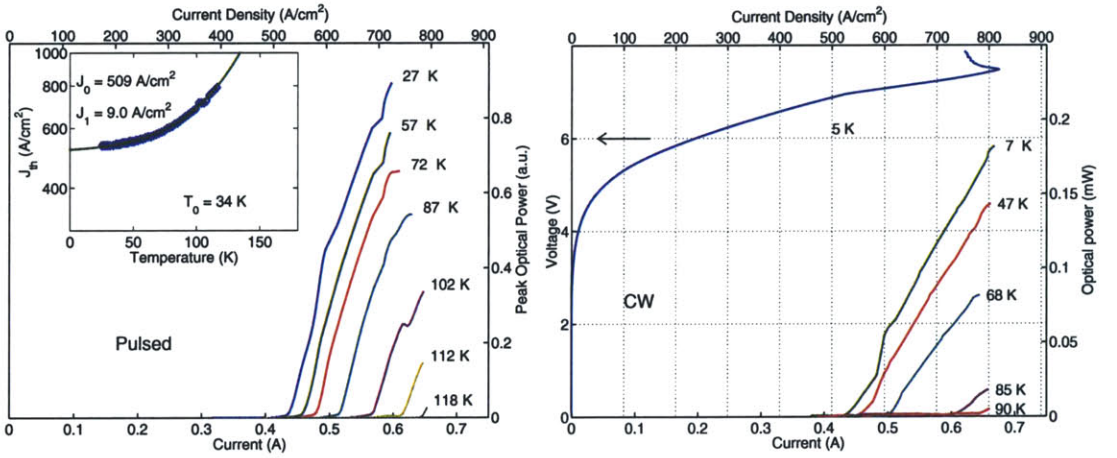


Figure 3-18: Experimental results from three different FL178C-M10 (MM Cu-Cu) lasers with 10 μm, 5.1 μm, and 2.8 μm thick active regions, respectively. The 10 μm thick laser has the highest reported value of $T_{\max,\mathrm{pul}} \sim 169$ K for any terahertz QCL. The lasers with active regions thinner than 10 μm were processed by wet-etching in a $H_3PO_4 : H_2O_2 : H_2O$ 1 : 1 : 25 solution (GaAs etch rate: 200 – 250 nm/min), and a mechanical profilometer was used to measure the final thickness.

substrate that could be lossy at terahertz frequencies due to impurity absorptions. For MM waveguides, a smaller fraction of the mode overlaps with the lossy metal for a thicker active region, thereby minimizing the waveguide losses α_w (as will be shown in chapter 4). To determine the effect of the active region thickness on laser performance, FL178C-M10 lasers with different active region thicknesses were processed in MM (Cu-Cu) waveguides. Experimental results for three such lasers are shown in Fig. 3-18. For the 10 μm thick active region, several different lasing devices were tested, which had a range of $T_{max,pul} \sim 160 - 169$ K. For the thinner active regions, only one of each type of lasers were tested; hence, the $T_{max,pul}$ reported in Fig. 3-18 should not be taken as an upper limit for those devices. Nevertheless, it is apparent that for a thinner active region, the threshold current density is higher, and correspondingly $T_{max,pul}$ is smaller. A device with active region as thin as 1.35 μm lased only barely before becoming electrically shorted during operation, possibly due to a non-uniform etch profile for the thickness of the active region. These experiments prove experimentally the absence of any *microcavity effect* for gain enhancement, as was discussed in section 2.2 (the vertical dimensions of the metal-metal waveguides, and correspondingly the vertical modal confinement in these lasers is much smaller than the wavelength since $\lambda_{GaAs} \sim 30$ μm at the lasing frequencies). However, the performance of the 2.8 μm thick active region laser is encouraging since a thinner active region has better heat removal, which directly improves cw performance of the laser. Moreover, the thinner active regions are easier to process, which is beneficial for demanding applications such as implementation of distributed feedback in the waveguide structures. It may be noted that the bias voltage for the devices in Figs. 3-18(b) and (c) does not scale down according to their thickness even after subtracting a possible $\sim 2 - 3$ V voltage drop across the metal-semiconductor Schottky contact. The reasons for this are not clear; it could possibly be related to a non-uniform thickness profile after etching.

3.3.3 Effect of the highly doped top contact layer on temperature performance

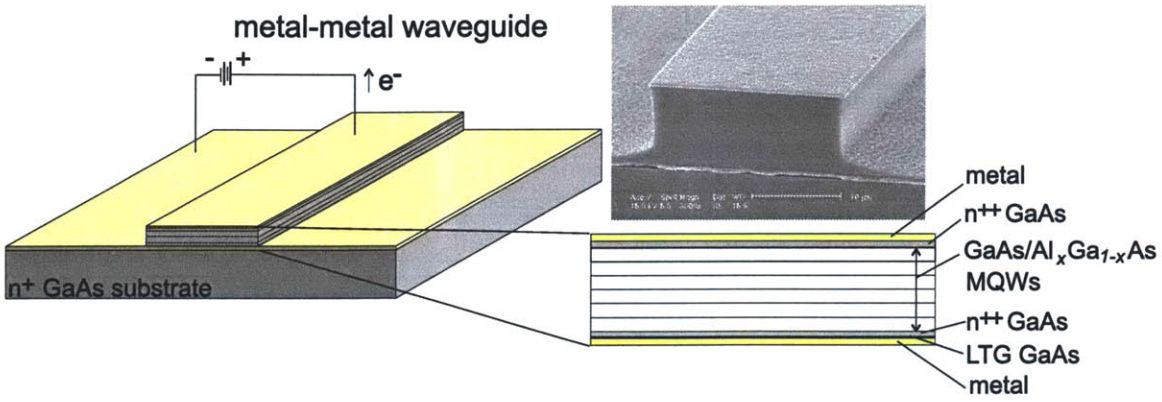


Figure 3-19: Schematic showing the layer sequence in a terahertz QCL fabricated with a MM waveguide. A scanning electron-microscope (SEM) image of the cleaved facet (10 μm $\times$ 25 μm active region) of a MM Cu-Cu laser is shown on the upper right.

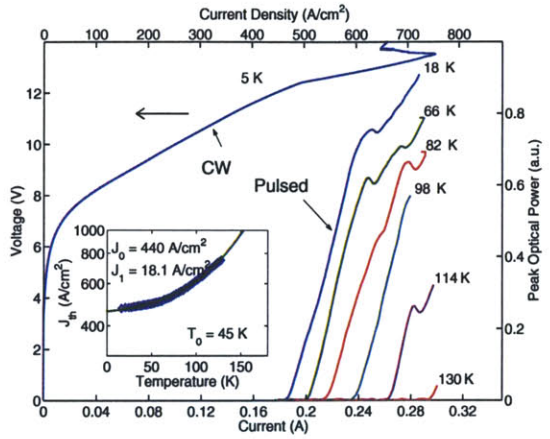
Low resistance ohmic contacts to n-GaAs are difficult to obtain due to a high energy depletion layer Schottky barrier ($\sim 0.8 - 0.9$ eV) associated with the metal-GaAs interface [114, 134, 17]. Alloyed ohmic contacts [25], where thin metal film(s) are deposited on the surface and then annealed to allow dopants to diffuse into the semiconductor, are often used; however, it is difficult to control the extent of dopant diffusion and there can be potential loss due to high free carrier losses at terahertz frequencies. To obtain a non-alloyed ohmic contact, the following two techniques are most commonly used. In the first technique, a layer of very highly doped n^{++} GaAs is grown on the top of the active region (for which the maximum doping density is typically limited to $\sim 5 \times 10^{18}$ cm^{-3}). This gives rise to a very narrow depletion barrier width ($\sim$ 100 Å) between the metal and the semiconductor. Increased conduction is then obtained due tunneling through the barrier. In another technique, a n-In$_x$Ga$_{1-x}$As epitaxial layer is grown with a graded composition to obtain a near ohmic structure, since the Schottky barrier height for the metal-In$_x$Ga$_{1-x}$As interface is nearly zero [179].

Figure 3-19 shows the layer structure for the terahertz QCLs with MM waveguides discussed in this thesis. The fabrication procedure will be described in detail in

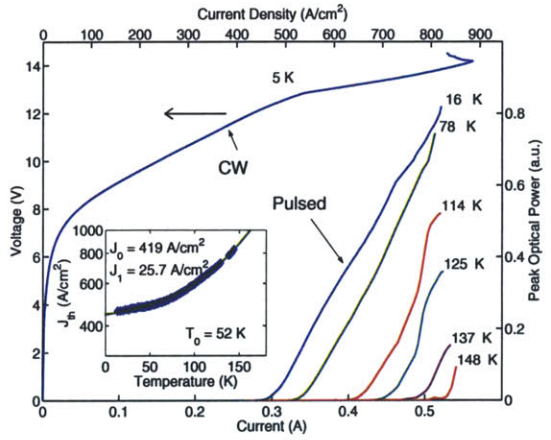
chapter 4 (section 4.2). A MBE machine dedicated only to GaAs/$Al_xGa_{1-x}As$ growth is typically used to grow terahertz QCLs to keep the machine "clean" and to minimize the background doping in the grown material. This is because terahertz QCLs are usually doped low ($< 7 \times 10^{15}$ cm^{-3}) to minimize free-carrier losses in the active region. Hence, the option of $In_xGa_{1-x}As$ to form non-alloyed ohmic contacts is not available. As indicated in the figure, the bottom most layer of the stack is a thin ($\sim$ 35 Å), low-temperature (250 °C) grown (LTG) GaAs layer on top of which there is a $\sim$ 500 Å thick n^{++} GaAs layer. The LTG layer has mid-gap defect states, which assist in the tunneling process to create a low-resistance contact [127]. However, there is no LTG layer for the top-contact since a low-temperature is not available during the growth of this layer (the MBE growth sequence is from top to bottom in Fig. 3-19 since the MBE stack gets inverted for MM waveguide fabrication). Thus, only a n^{++} GaAs layer (which is ~ 0.1 μm thick and is doped to $\sim 5 \times 10^{18}$ cm^{-3}) is used for the top contact.

The role of the top contact layer (and to some extent, even the LTG layer for the bottom contact) on the laser performance is not clearly understood. This is due to some of the more recent observations from the experimental results as shown in Fig. 3-20. In the figure, experimental data from lasers of roughly the similar widths, and fabricated with, and without the heavily doped top contact layer, respectively, are reported. The temperature performance of lasers with MM waveguides depends weakly on the cavity dimensions (this will be explained in chapter 4), and hence length variations in the devices can be disregarded. As is evident from the figure, the devices without the top contact layer consistently have a $J_{max} - J_{th,5K}$ value larger by $50 - 150$ A/cm^2 as compared to that for devices with the contact layer, which leads to their $T_{max,pul}$ being higher by $15 - 30$ K correspondingly. It may be noted that due to etching uncertainties, exact widths of the cavities are not known, hence, it is more appropriate to compare $J_{max} - J_{th,5K}$ (or $J_{max}/J_{th,5K}$ to be even more accurate), rather than just J_{max} or $J_{th,5K}$, for different devices to compare their temperature performance. It is also likely that QCLs with the top contact layer have a larger waveguide loss, since large fringing fields can exist at the top corners of the waveguide

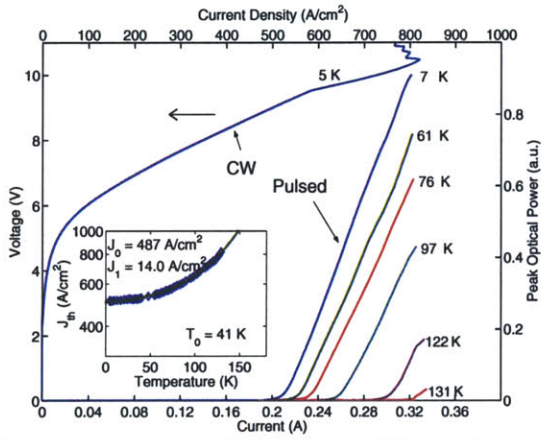
(a1) FL178C-M7 (In-Au), 40 μm × 0.99 mm top contact layer present



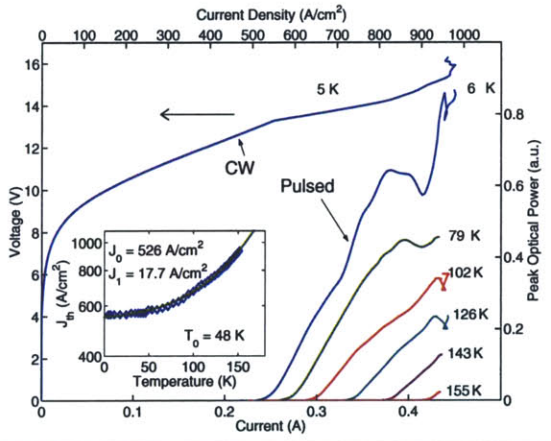
(a2) FL178C-M7 (In-Au), 50 μm × 1.27 mm top contact layer removed completely



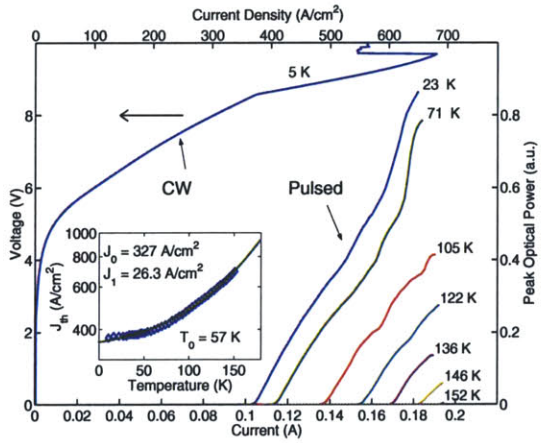
(b1) FL179C-M9-2 (In-Au), 40 μm × 1.0 mm top contact layer present



(b2) FL179C-M9-2 (Cu-Cu), 50 μm × 0.92 mm top contact layer removed completely



(c1) FL178C-M10 (Cu-Cu), 40 μm × 0.7 mm top contact layer removed only partially



(c2) FL178C-M10 (Cu-Cu), 40 μm × 2.1 mm top contact layer removed completely

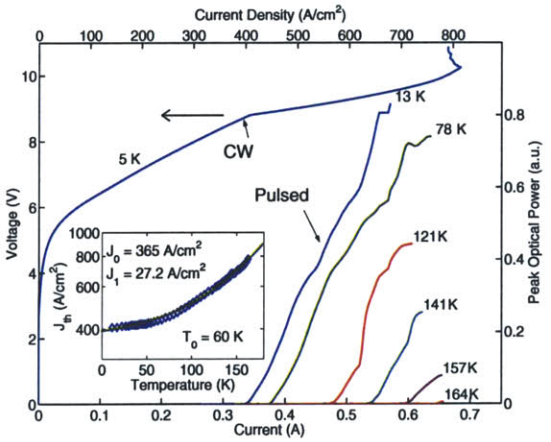


Figure 3-20: Experimental results from terahertz QCLs with MM waveguides, fabricated with, and without the highly doped top contact layer, respectively, for three different QCL designs. Wheres the L-Is are for pulsed operation, the I-Vs are for cw operation since pulsed I-Vs for all devices were not available.

if any part of the contact layer is not well covered by metal at the edges, which then induces extra propagation loss. In general, the contribution to the waveguide loss by the doped contact layer, when it is covered by metal, is an order of magnitude smaller than due to the metal itself as shown through Drude model calculations in chapter 4 of Williams [166].

The reasons for the aforementioned increase in $J_{\max} - J_{\text{th,5K}}$ for devices without the top contact layer are, as of yet, unclear. One possibility could be dopant diffusion from the highly doped top contact layer into the adjacent active region QCL modules, which could make the doping profile for the first few modules different as compared to the rest. Since the current density has to be same across all the modules during device operation, this would then cause a non-uniform static electric field to develop across the first few modules (since each module will have a different J-V characteristic due to a different doping profile), which could trigger an early NDR in the cascade of modules, thereby limiting the value of $J_{\max}$.

For devices without the top contact layer, the parasitic voltage drop is typically higher, which causes greater electrical power to be dissipated in the device during cw operation. This behavior can possibly be explained by the fact that the tunneling barrier will be thicker for a direct metal contact on the lightly doped active region, which therefore will increase the contact resistance. However, this behavior has been observed to be more random with some lasers being affected by a larger increase in operating voltages while the others behaving the same as those with the top contact layer, the reasons for which are also unclear. Finally, it may be noted that no definitive record exists for the FL178C-M7 lasers whose experimental results were shown in Fig. 3-16, about whether or not the top contact layer was removed for those devices. It is believed that at least a part, if not full, of the top contact layer was removed for those devices, since many devices tested with the top contact layer had similar results to those in Fig. 3-20(a1), which are significantly different than those presented in Fig. 3-16. All the devices in Fig. 3-18, however, are without the top contact layer due to their specific fabrication procedure that involved etching of the active region.

3.3.4 Temperature degradation mechanisms in terahertz QCLs with the FL design

This section describes the various mechanisms perceived to be the cause of temperature degradation in QCLs based on the FL design. Most of the discussion in this section is also applicable to terahertz QCLs in general. The discussion on the issue of thermal backfilling of the lower radiative level as the electrons in the injector level(s) become hotter, which is also one of the possible temperature degradation mechanisms, will be deferred until section 3.4.1, since it has been experimentally determined that it is not the primary temperature degradation mechanism in the single-phonon depopulated FL designs at their present operating temperatures.

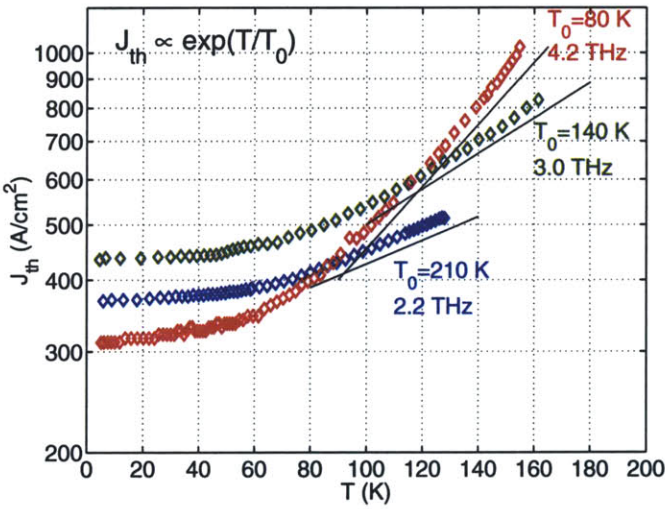


Figure 3-21: Semi-logarithmic plot of the J_{th}-T variation for three different terahertz QCLs fabricated with metal-metal (Cu-Cu) waveguides and based on the FL design, operating close to the frequencies of 2.2 THz (FL175M-M3), 3.0 THz (FL178C-M7), and 4.2 THz (FL183R-2), respectively. A phenomenological fit of the high-T data points to the expression $J_{\text{th}} \propto \exp(T/T_0)$ is also indicated. The data is for pulsed mode operation, when $T_{\text{active}} \sim T$, where T is the heat-sink temperature.

Figure 3-21 shows the variation of the threshold current density J_{th} with the heat-sink temperature T ($\sim T_{\text{active}}$ in pulsed operation) for three different terahertz QCLs based on the FL design of Fig. 3-7. Whereas FL178C-M7 ($\nu \sim 3.0$ THz) was discussed in section 3.3.1, FL175M-M3 ($\nu \sim 2.2$ THz) and FL183R-2 ($\nu \sim 4.2$ THz), which are minor variations of the FL-C series of designs will be discussed in section 3.5.1, and

section 3.6.2, respectively. Since the lower level parasitic current channels determine $J_{\rm th}$ at low-temperature in these designs, rather than fitting the entire curve to the phenomenological expression $J_{\rm th} = J_0 + J_1 \exp(T/T_0)$, only the high-temperature data points are fitted to determine a more representative value for the characteristic temperature T_0. It may be noted that even though T_0 is the highest for the 2.2 THz laser, the $T_{\rm max,pul}$ ($\sim$ 130 K) for it is relatively small since its $J_{\rm max} - J_{\rm th,5K}$ is much smaller compared to the other two designs, the reasons for which will be explained in section 3.5.1.

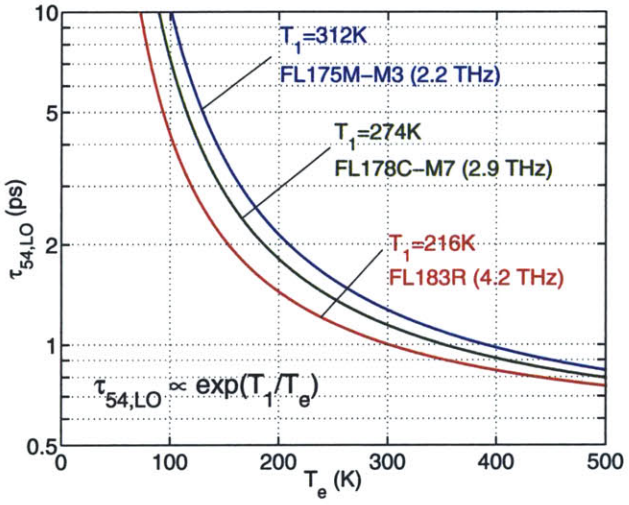


Figure 3-22: Semi-logarithmic plot of the $\tau_{54,\rm LO}$-$T_{\rm e}$ variation for three different FL designs with E_{54} ($\sim \hbar\omega$) values of 9.1 meV, 12.4 meV, and 17.4 meV corresponding to frequencies of 2.2 THz, 3.0 THz, and 4.2 THz, respectively. The value $\tau^{\rm hot}_{54,\rm LO}$ is calculated to be $\sim$ 0.5 ps for all the three designs, and equation (3.7) is used for the plots in the figure with $\hbar\omega_{\rm LO} \sim 36$ meV corresponding to the value in bulk GaAs. Here, $T_{\rm e}$ is the electronic temperature that characterizes the thermal (Maxwell-Boltzmann like) distribution of the electrons in the upper radiative state 5.

One of the primary temperature degradation mechanisms in LO-phonon depopulated terahertz QCLs is believed to be the reduction in the upper state lifetime due to thermally activated LO-phonon scattering from the upper radiative state u to the lower radiative state l. This is because the radiative energy separation $E_{ul} < \hbar\omega_{\rm LO}$ for terahertz QCLs, and hence, only the electrons in level u with kinetic energy $E(k_{\|})$ $(= \hbar^2 k_{\|}^2/2m^*)$ that is greater than $\hbar\omega_{\rm LO} - E_{ul}$ can scatter into the lower level l by emitting an LO-phonon. The dependence of the LO-phonon scattering rate on the

electron wavevector $\mathbf{k}_{\|}$ is weak, and is calculated in chapter 2 of Williams [166] for some specific cases. In that case, the lifetime $\tau_{ul,\mathrm{LO}}$ due to LO-phonon emission from the level $u \rightarrow l$ can be determined approximately from the following simple expression

$$\tau_{ul,\mathrm{LO}} \approx \tau_{ul,\mathrm{LO}}^{\mathrm{hot}} \exp\left(\frac{T_1}{T_\mathrm{e}}\right) \tag{3.7}$$

where,

$$T_1 \equiv \frac{\hbar\omega_{\mathrm{LO}} - E_{ul}}{k_\mathrm{B}} \tag{3.8}$$

In the expression above, state-blocking in level l is neglected and T_e is the characteristic temperature for the electronic distribution in level u, which is typically Maxwell-Boltzmann like, as was discussed in section 2.2 (Fig. 2-2). $\tau_{ul,\mathrm{LO}}^{\mathrm{hot}}$ is the lifetime for the electrons with the kinetic energy $E(k_{\|,\mathrm{LO}}) = \hbar\omega_{\mathrm{LO}} - E_{ul}$, which is the lowest kinetic energy in level u for which such a scattering is possible, and $\exp(T_1/T_\mathrm{e})$ is the fraction of the electrons that have a kinetic energy greater than $E(k_{\|,\mathrm{LO}})$. $\tau_{ul,\mathrm{LO}}^{\mathrm{hot}}$ can be analytically calculated [151, 166], and is a function of the overlap of the electron wavefunctions and the momentum of the LO-phonons that can be emitted. It may be noted that T_e increases with the heat-sink temperature T although it is a much difficult problem to determine its exact value. Numerical simulations as well as experimental measurements suggest that T_e may be $50-100$ K higher than T_{active} (which is $\sim T$ in pulsed operation) during device operation [31, 29, 157].

The reduction in the upper state lifetime $\tau_{54,\mathrm{LO}}$ with the electronic temperature T_e for three different terahertz QCLs based on the FL design, and operating at different frequencies is shown in Fig. 3-22. Even though the $\tau_{54,\mathrm{LO}}$-T variation in Fig. 3-22 might qualitatively explain the J_{th}-T variation in Fig. 3-21, it is not quite consistent quantitatively. Note that $\tau_{54,\mathrm{LO}} \propto \exp(T_1/T_\mathrm{e})$, whereas, $J_{\mathrm{th}} \propto \exp(T/T_0)$ as obtained experimentally. Firstly, both these expressions have a different dependence on the temperature in the exponent, which is difficult to explain mathematically for any algebraic relation between J_{th} and $\tau_{54,\mathrm{LO}}$. Secondly, J_{th} increases more rapidly with T for a higher frequency design; however, $\tau_{54,\mathrm{LO}}$ decreases less rapidly with T_e for the same, thus contradicting the observed experimental behavior. Clearly, then, some

other complex temperature degradation mechanisms are also at play.

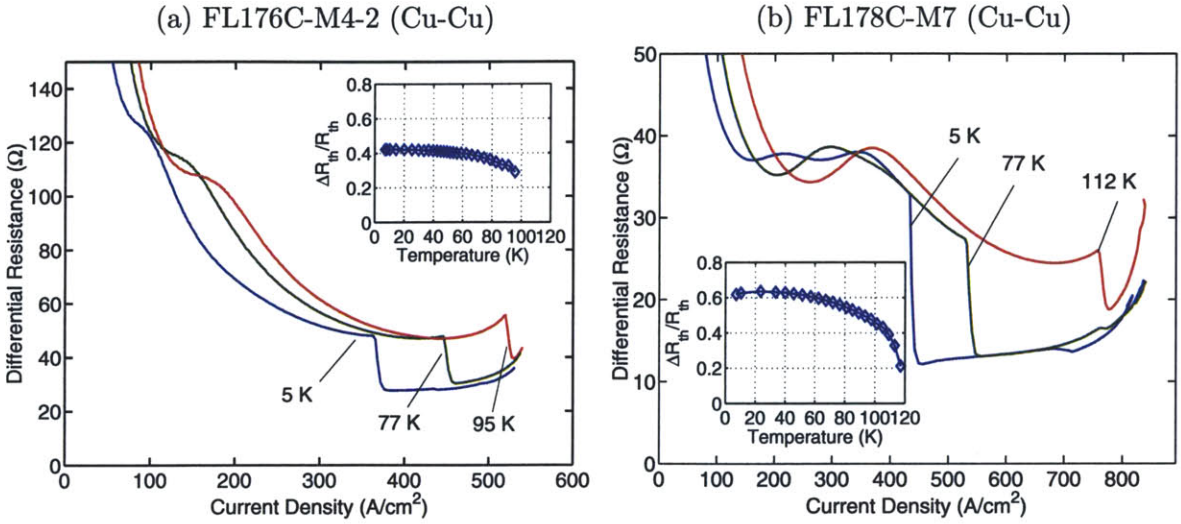


Figure 3-23: Differential resistance $\mathcal{R}$ versus current I plots measured at three different heat-sink temperatures in cw operation. Plot (a) is for a FL176C-M4-2 laser, same as the one in Fig. 3-13(b), and plot (b) is for a FL178C-M7 laser, same as the one in Fig. 3-16(b). The discontinuity in $\mathcal{R}$ at threshold specified by the expression $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ (as defined in section 2.4.2) is also plotted in the corresponding insets.

Figure 3-23 shows the variation of the differential resistance $\mathcal{R}$ with the heat-sink temperature T for two different FL-C series lasers in cw operation. Note that the features corresponding to the lower level parasitic current channels occur at a larger value of $\mathcal{R}$ for FL176C-M4-2 as compared to FL178C-M7 since the related parasitic anticrossings are smaller for FL176C-M4-2 (Fig. 3-10). The active region temperature $T_{\rm active}$ in cw operation can be much greater than T due to Joule heating caused by the electrical power dissipated in the device. For example, $T_{\rm max,cw} = 117$ K and $T_{\rm max,pul} = 158$ K for the FL178C-M7 laser in Fig. 3-23(b), which suggests that $T_{\rm active} \sim T + (30 - 50)$ K for this particular laser in its operating range. Nevertheless, the $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$-$T$ variation, as plotted in the insets, is directly indicative of how $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ changes with the active region temperature.

An expression for $\Delta\mathcal{R}_{\rm th}/\mathcal{R}_{\rm th}$ for the FL design of Fig. 3-7 can be derived similar to the one derived for a simple 3-level QCL in section 2.4.2. For the tightbinding wavefunctions of Fig. 3-7(b), a density matrix similar to that in equation (2.85) can be written for the levels 5, 4, and 1′ by replacing τ_{31} with $\tau_{5,\rm par}$ ($\equiv \tau_{5\rightarrow(6,3,2,1)}$), and τ_2

with $\tau_{4,\text{eff}}$ (which can be derived as in equation (2.115) assuming that the current flow through the collector barrier is primarily by $4 \rightarrow 3$ resonant-tunneling). The result in equation (2.95) is then modified for the FL design to be

$$\frac{\Delta\mathcal{R}_{\text{th}}}{\mathcal{R}_{\text{th}}} = \frac{2\Delta n_{\text{th}}}{(n_1 + n_4 + n_5)} \left(\frac{1 - \frac{\tau_{4,\text{eff}}}{2\tau_{5,\text{par}}}}{1 + \frac{\tau_{4,\text{eff}}}{\tau_{5,\text{par}}}} \right) \tag{3.9}$$

Note that the expression for $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}}$ is independent of τ_{54} since the discontinuity in $\mathcal{R}$ occurs only when the required population inversion is achieved, even though the current density that is required to reach the threshold depends critically on τ_{54}.

For the tightbinding wavefunctions of Fig. 3-7(b), $n_2 > n_1 > n_5 > n_4 > n_3$ during laser operation, assuming resonant-tunneling is the primary mechanism for current conduction across the intra-injector, the injector, and the collector barriers, respectively. Note that the total doping n_{tot} in equation (2.95) is replaced by a smaller value $(n_1+n_4+n_5)$ in equation (3.9). To find the value for n_1, a 5 level density matrix must be solved; however, an estimate can be obtained from the three level solution (i. e. for levels 1, 4, and 5) and the fact that $n_{\text{tot}} \approx (2n_1+n_4+n_5)$ since $n_1 \approx n_2 \gg n_3$. Also note that $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}}$ is larger for FL178C-M7 as compared to FL176C-M4-2. This is because FL178C-M7 has a thinner injector barrier, which makes the gain linewidth broader, thus requiring a larger amount of population inversion Δn_{th} to obtain a given value of the peak gain (as was shown in section 2.5.1, Fig. 2-14a).

The expression for $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}}$ in equation (3.9) is derived without any major approximations when using the tightbinding wavefunctions of Fig. 3-7(b). $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}}$ can be measured experimentally and is plotted in the insets of Fig. 3-23. As temperature increases Δn_{th} needs to increase due to increased losses in the metal claddings. Also, any change in the value of $(n_1+n_4+n_5)$ is likely to be small. Hence, $\tau_{4,\text{eff}} \uparrow$ and $\tau_{5,\text{par}} \downarrow$ remain the only two mechanisms that can cause $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}} \downarrow$ as $T \uparrow$. As can be deduced from Fig. 2-10(b), a rather large reduction in $\tau_{5,\text{par}}$ is needed to explain the observed behavior in Fig. 3-23. This could then very well be an equally important temperature degradation mechanism in terahertz QCLs, in addition to the reduction in the lifetime τ_{54} due to thermally activated LO-phonon scattering.

Whereas an increase in $\tau_{4,\text{eff}}$ can be caused due to increased dephasing for resonant-tunneling across the collector barrier, the causes for the reduction in $\tau_{5,\text{par}}$ are more ambiguous and difficult to explain. With an increased temperature the electrons in a subband acquire larger values for the in-plane wavevectors $k_{\|}$. In the simplest case, the tunneling probability through a barrier is considered independent of $k_{\|}$, as in the original model proposed by Esaki and Tsu [49]. However, the in-plane and z-directed electron motion cannot be de-coupled in a heterostructure as simply as was assumed in equations (2.16) and (2.17), which makes tunneling $k_{\|}$ dependent in the most general case. Simple models for tunneling transport that incorporate the effect of different effective masses in the wells and the barriers, which also makes tunneling $k_{\|}$ dependent, have been reported [141]. However, much more complex phenomena, which involve interface-related effects lying outside the scope of conventional envelope-function theory [56], are likely to make resonant-tunneling dependent on $k_{\|}$ additionally. Considering these effects, a reduction in the lifetime $\tau_{5,\text{par}}$ as the average $k_{\|}$ increases in the level 5 could be due to enhanced tunneling into the phonon level 3 and the upper parasitic level 6, which would have otherwise been negligible due to relatively large detuning energies E_{56} and E_{53} close to the operating bias range.

Figure 3-24 shows the typical variation of the I-Vs and differential resistance $\mathcal{R}$ with the heat-sink temperature for a FL178C-M10 laser in cw operation. The lasing threshold is characterized by a kink in the I-V, which also corresponds with a discontinuity in $\mathcal{R}$. As was discussed in section 3.3.1, the low temperature threshold current densities (J_{th}) are artificially enhanced due to the presence of lower bias parasitic current channels. Besides $\tau_{54} \downarrow$, $\tau_{5,\text{par}} \downarrow$, and $\tau_{4,\text{eff}} \uparrow$, which all cause $J_{\text{th}} \uparrow$, an enhancement in the current due to the lower bias parasitic current channels is also believed to be a major cause for the J_{th} to increase with temperature. This is because an increase in parasitic current causes a pinch-off in the lasing bias range, and therefore, a corresponding pinch-off in $J_{\text{max}} - J_{\text{th}}$. This is deduced from the fact that the current in the device increases uniformly at all bias, rather than just close to the design bias as $T \uparrow$. A higher average value of the in-plane wavevectors in each subband can cause tunneling into levels that are otherwise sufficiently detuned

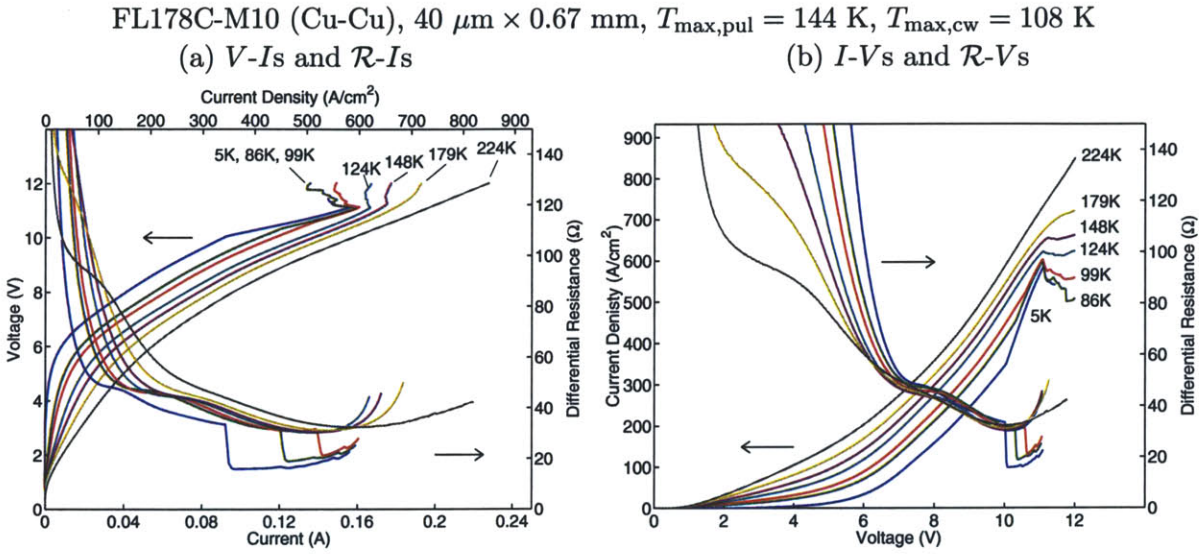


Figure 3-24: (a) Variation of the I-Vs and the differential resistance $\mathcal{R}$ with the heat-sink temperature T for a FL178C-M10 (Cu-Cu) laser in cw operation. Plot (a) is the same as plot (b), except that $\mathcal{R}$ is plotted versus current in (a), and versus voltage in (b). Note that the active region temperature $T_{active} \sim T + (30 - 50)$ K, as deduced from the values of $T_{max,pul}$ and $T_{max,cw}$ for this laser.

from the resonance condition, which might possibly explain this behavior. This is also evident from the I-V and $\mathcal{R}$-V plots in Fig. 3-24(b), which become broader in bias space as temperature increases. It is less likely for increased dephasing to be the primary cause of this broadening, since that should also cause a reduction in the current flow, which is clearly not observed.

Further improvement in temperature performance of FL-C designs?

After multiple iterations in the FL-C series of designs, it may seem as if there has been a saturation in their temperature performance (which is so far the best temperature performance from any terahertz QCL design). However, it is believed that this design scheme has more to offer. As discussed in section 3.3.3, the higher temperature performance for the FL178C-M7, FL179C-M9-2, and FL178C-M10 lasers might be more due to issues other than those related to the design itself. It is perceived that thickening of the collector barrier to reduce the collector anticrossing might reduce the rate of increase of parasitic current with temperature (as noticed from Fig. 3-24).

Moreover, making the injection somewhat "downwards-diagonal" as in FL175C and FL178C-M1 (Fig. 3-9) to make the $1' - 5$ injection stronger than the $2' - 5$ injection, thinning the intra-injector barrier to make available a larger range in voltage for lasing, and thinning the injector barrier to make injection more coherent might help in obtaining a larger $J_{\max}$, while keeping the parasitic channels in check due to a thicker collector barrier. These suggested changes are somewhat against the paradigm established in the designs developed after FL178C-M1, whereby the injection was made more "upwards-diagonal" and the intra-injector barrier was thickened to limit the lower level parasitic current channels. This was needed since the collector barrier was always kept relatively thin. However, as was argued in section 2.5.2 (Fig. 2-19b), the peak gain is not affected by the collector anticrossing critically even though the lower state lifetime is. It then becomes a question of whether a smaller lower state lifetime, or instead a more transparent injection mechanism but with a longer lower state lifetime will obtain a better temperature performance. In that regard, smaller values for the collector anticrossing are worth trying. The FL177C-M8 design (Fig. 3-11, Table 3.1) was a step in this direction with its collector barrier ~ 1 ML thicker than the other three best performing designs (FL178C-M7, FL179C-M9-2, FL178C-M10); however, its temperature performance was a worse due to a inexplicably low value of the peak current density (~ 500 A/cm^2). Hence, a constructive conclusion cannot be drawn from its experimental results.

3.4 Towards an active region design to better the temperature performance of the FL-C series designs

This section describes many different types of terahertz QCL designs that have been tried in this group to obtain a better performing design as compared to the FL design of section 3.3. Several designs that are significantly different from each other have achieved lasing, which in part is due to an improved understanding of the transport

properties of phonon-depopulated QCLs obtained by studying the operation of the FL-C series lasers. This is not a mean feat considering the fact that only a handful of different types of designs that have been developed elsewhere are known to exist, despite continued efforts around the world to develop terahertz QCLs. However, none of the designs presented in this section have bettered the temperature performance of the FL-C series designs, even though some of these designs are believed to hold promise with future design iterations. Nevertheless, every new design is a step towards better understanding of the operating principles of this class of devices, as will become apparent in the following discussions.

3.4.1 Designs with two-phonon depopulation: FTP series

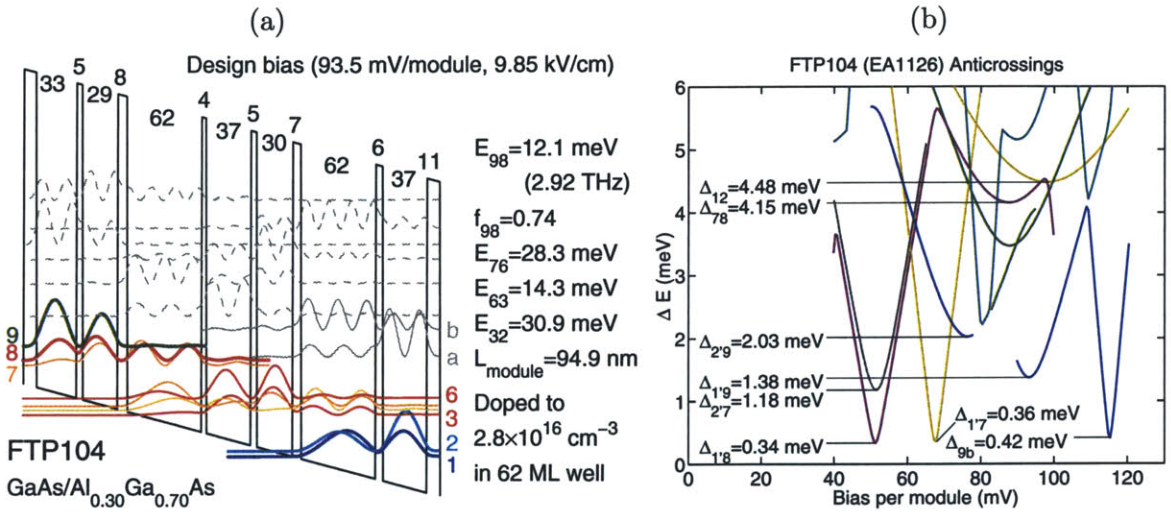
It was recognized early in the development of mid-infrared (IR) QCLs that thermal backfilling of the lower radiative state by carriers from the heavily populated injector states could effectively lengthen the lower state lifetime and degrade population inversion. This can happen due to excitation of the carriers from the injector states by reabsorption of non-equilibrium LO-phonons (the hot-phonon effect) [108, 84]. The injector regions were therefore typically designed with large energy separations ($\Delta \geqslant 100$ meV) between the lower lasing level and the ground state in the injector [149]. More recently, quantum-cascade active regions with doubly resonant LO-phonon depopulation schemes have been used to demonstrate room-temperature cw operation [21, 43], and are by-far the are the best performing mid-IR QCL designs.

In terahertz QCLs the energy levels have to be much more closely spaced than in mid-IR QCLs. In that case, thermal backfilling can also arise due to other scattering mechanisms such as electron-electron scattering, and electron-impurity scattering. While this is a bigger problem for the miniband extraction based designs such as those discussed in section 1.2.3, this type of backfilling is possibly reduced in the single-phonon depopulated designs such as the FL designs presented in section 3.3.1 due to a relative large thermal barrier between the injector and the lower radiative states ($\Delta \sim 36$ meV $\sim k_B \cdot 420$ K). This is one of the main reasons why phonon-depopulated terahertz QCLs have demonstrated a better temperature performance.

This section discusses terahertz QCLs with two-phonon depopulation schemes to investigate whether thermal backfilling due to the presence of non-equilibrium LO-phonons limits the QCL performance. These designs are categorized as the FTP series designs ("F"–FL, "TP"–two-phonon) because of their similarity with the FL design in terms of the resonant-phonon depopulation scheme, and because of a two-phonon depopulation mechanism. Three such designs were tried, namely FTP, FTP-B, and FTP-C, and all three of them achieved lasing. Whereas FTP and FTP-C lased in the first try, FTP-B lased only in the second design iteration.

The design details and experimental results for the first of such designs, named as FTP104 because of 104 repeated modules in its MBE growth, are shown in Fig. 3-25 [176]. This is the first terahertz QCL that has been developed in the $GaAs/Al_{0.3}Ga_{0.7}As$ material system. The greater Al percentage in the barriers makes the barrier height larger (~ 0.27 eV), which allows for a bigger voltage drop per module that is needed for the two-phonon depopulation scheme. For this design, the LO-phonon depopulation scattering events take place in two separate wide wells, which are connected by a four-state miniband with a bandwidth of approximately 14 meV. The thickness of the barriers in this miniband were chosen in an attempt to provide fast transport between the wide wells, but thick enough to suppress parasitic coupling between the upper radiative level 9 and levels a and b that are close in energy ($\Delta_{9a} \sim 0.25$ meV, $\Delta_{9b} \sim 0.42$ meV). In terms of energy level design and applied field ($\sim$ 10 kV/cm at design bias), this design is very similar to the four-well FL design of section 3.3.1, except with three extra wells per module added to allow a second phonon depopulation event.

Fig. 3-25(c) shows pulsed and cw measurement results from a MM (Cu-Cu) ridge laser. In pulsed mode the threshold current density at 10 K is $J_{\text{th}} = 170$ A/cm^2 and $T_{\text{max,pul}} \sim 140$ K. The threshold current density in this design, along with one-well injector designs to be described in section 3.5.2, are the lowest of all resonant-phonon terahertz QCLs developed so far. However, the low value for this design may not necessarily be due to intrinsically better depopulation or more gain, but rather a better suppression of lower level parasitic current channels due to the misalignment



(c) FTP104 (Cu-Cu), 50 μm $\times$ 1.06 mm

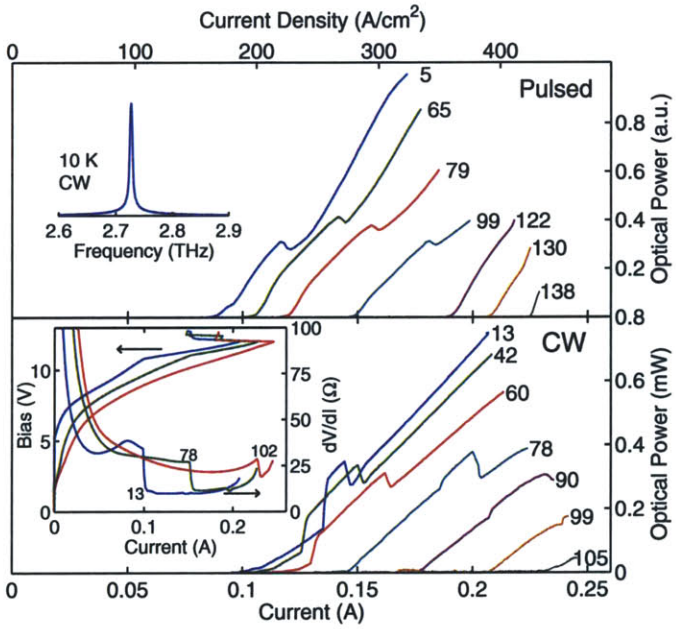


Figure 3-25: Design details and experimental results for FTP104. (a) One module calculation at the design bias. Levels 9 and 8 are the upper and the lower radiative levels, respectively. Level 8 is depopulated by a resonant-phonon scheme down into the $3-6$ miniband, from where there is another LO-phonon depopulation event down into the injector levels 1 and 2. (b) Anticrossing plots for a two-module calculation. (c) Experimental results from a MM Fabry-Pérot ridge laser.

of energy levels in the four-level miniband at low bias, which thus limits current transport. At higher temperatures, the Fermi distribution of the electrons becomes broader and transport through the miniband becomes more efficient, which pushes up the parasitic current, and consequently the J_{th}. This is also likely to be the reason why a strong increase in J_{max} is observed for this design with temperature as compared to the FL-C designs of section 3.3.1, for which J_{max} remains approximately unchanged up to the maximum lasing temperatures.

The second in the series of FTP designs were the FTP-B designs. The second iteration of the FTP-B design, named as FTP160B-M1, is shown in Fig. 3-26(a). The operating bias field for this design is almost twice as high as the FTP104 design ($\sim$ 20 kV/cm) because of the absence of the intermediate miniband, whereby, the two LO-phonon depopulation relaxation events take place vertically within a region of several quantum wells (as determined from the spatial extent of level 3 in the module). This is done to remove any bottleneck to current transport that may be caused by that miniband. The depopulation region is composed of three wells with very thin barriers, which push up any low-lying parasitic energy levels to higher energies. Moreover, by adjusting the well and barrier widths in the triplet of the depopulation wells, the energy separation E_{43} and E_{32} could independently be adjusted close to the desired value of $\hbar\omega_{\text{LO}}$. This design features a four-well active region that helps obtain a large value for the oscillator strength (~ 0.82) by spatially extending the radiative levels. For injection, a doublet of levels is used in this design similar to that in the FL designs of section 3.3.1.

Note that lasing was not observed in a preceding design named as FTP159B, which is similar to FTP160B-M1 except with two additional wells in the injector region to have a quadruplet of injector levels. It is speculated that the injection mechanism in that design was not efficient since the electron population gets divided amongst the injector levels as the number of levels are increased. A yet another design named as FTP174B-M2, which succeeded FTP160B-M1 and was similar to FTP160B-M1 except with a two-well active region, did not achieve lasing either. This was possibly due to a small value of the oscillator strength (~ 0.57), which arises because the active

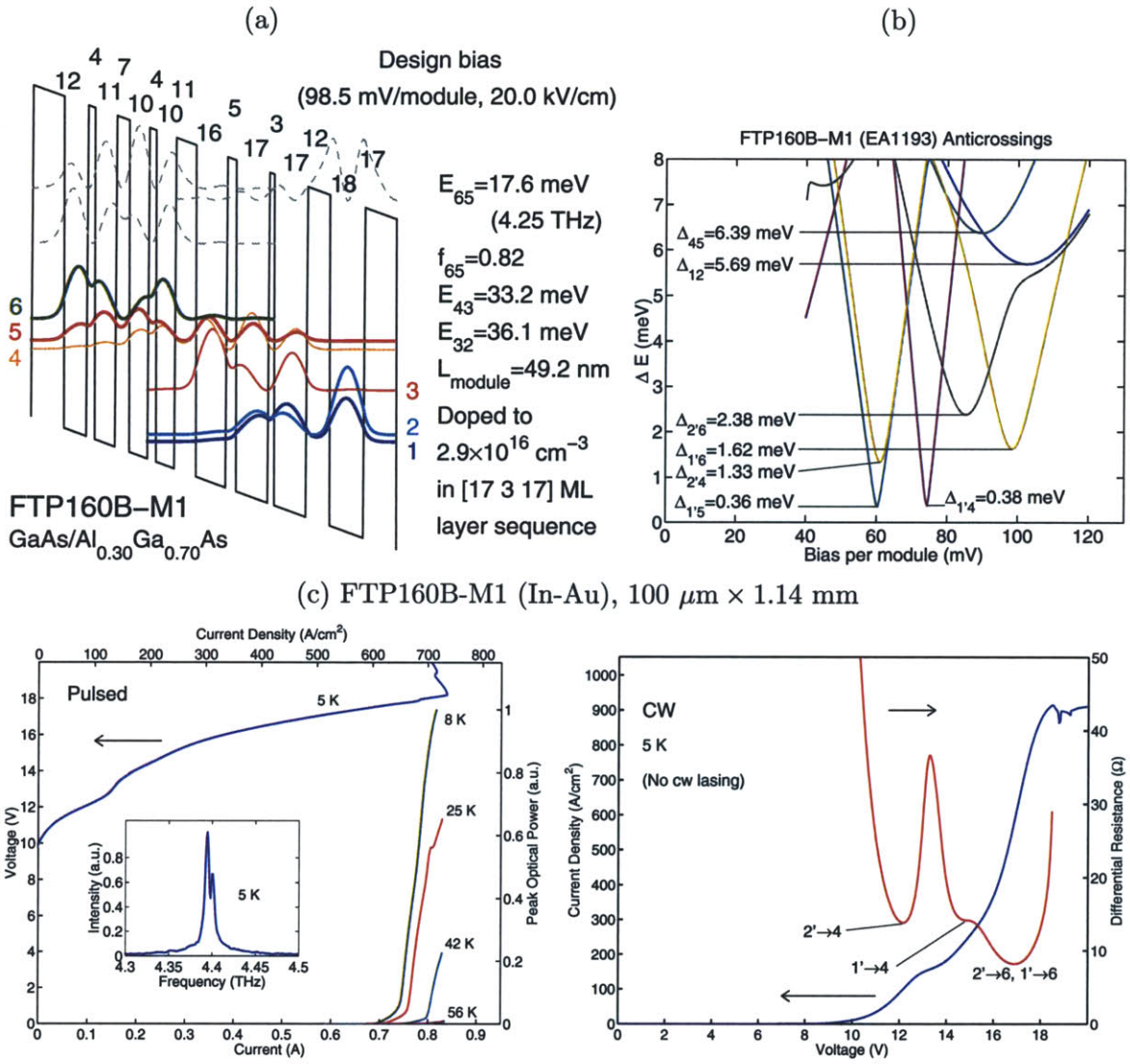


Figure 3-26: Design details and experimental results for FTP160B-M1. (a) One module calculation at the design bias. Levels 6 and 5 are the upper and the lower radiative levels, respectively. Level 5 is depopulated by a resonant-phonon scheme down into level 3, which itself is depopulated by another LO-phonon scattering event down into the injector levels 1 and 2. (b) Anticrossing plots for a two-module calculation. (c) Experimental results from a MM Fabry-Pérot ridge laser.

region wells in the two-well active region design had to be made narrow (15 ML and 17 ML as compared to $\sim$ 25 ML and $\sim$ 30 ML in the FL-C designs) to push the energy levels higher as is required for the two-phonon depopulation scheme.

Experimental results from a 100 μm wide FTP160B-M1 MM (In-Au) ridge laser are shown in Fig. 3-26(c). This device had a very small lasing range in current, and consequently a low value for the maximum temperature in pulsed operation ($T_{\mathrm{max,pul}} \sim$ 57 K). Lasing in cw mode was not observed, possibly due to the fact that the active region temperature in a 100 μm wide ridge could be higher than the heat-sink temperature by at least 100 K even if a conservative value of 5 K/W is used for the thermal resistance of the active region [94]. This is because of the high electrical power dissipated in the device ($\sim$ 18 W close to peak bias) due to a large voltage drop per module required for the two-phonon depopulation scheme. For the absence of any additional parasitic levels, and for the values of various anticrossings and the radiative oscillator strength, the FTP160B-M1 design looks at least as good as the FL-C designs of section 3.3.1. Also, from the measured differential resistance plot shown in Fig. 3-26(d), it seems as though the design could be biased well past the lower level parasitic current channels to the designs bias. The poor performance of this design is then more likely due to a large linewidth that could arise due to inhomogeneous broadening. The radiative energy E_{65} in this design depends critically on the thickness of the barriers in the active region, which in this case are as narrow as 4 ML. Since growth fluctuations from module to module can induce a large fractional change in the barrier widths for such narrow barriers, the radiative levels 6 and 5 are perceived to be affected by inhomogeneous broadening much more than the FL-C designs in this case.

The last design in the FTP series of designs was named as FTP185C and is shown in Fig. 3-27(a) [176]. This design is similar to the FTP160B-M1 design except that the wide well for phonon depopulation is a single well without any further sub-divisions, and a two-well active region is used instead of a four-well active region (since for this design, an oscillator strength as large as $\sim$ 0.69 could be obtained even with two wells in the active region unlike in the FTP-B design where more number of wells had to

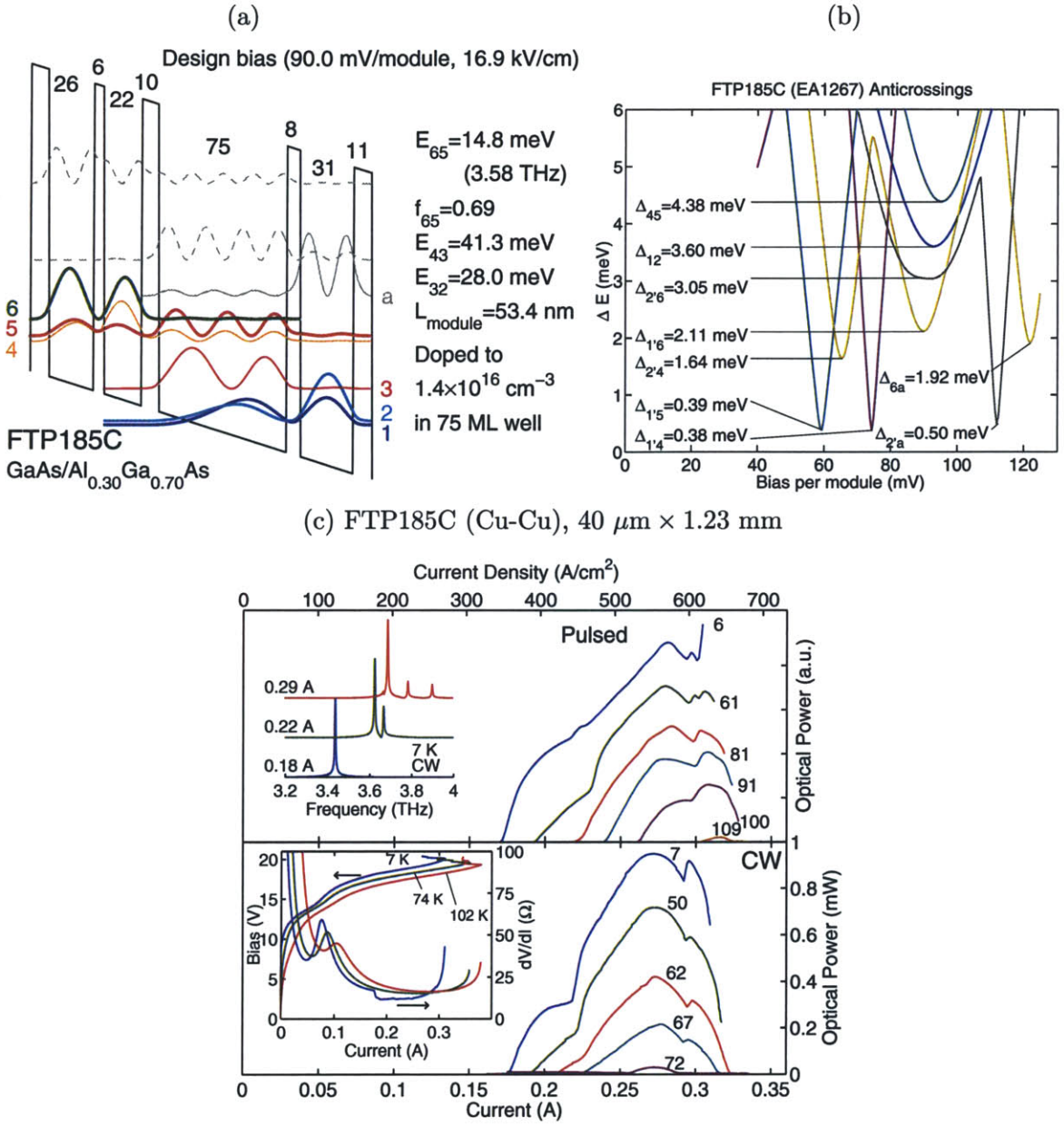


Figure 3-27: Design details and experimental results for FTP185C. (a) One module calculation at the design bias. Levels 6 and 5 are the upper and the lower radiative levels, respectively. Level 5 is depopulated by a resonant-phonon scheme down into level 3, which itself is depopulated by another LO-phonon scattering event down into the injector levels 1 and 2. In concept, this design is similar to the FTP160B-M1 design. (b) Anticrossing plots for a two-module calculation. (c) Experimental results from a MM Fabry-Pérot ridge laser.

be used). Because of the intrinsic energy level spacings in a single quantum well, the subband separations in the wide quantum well are not exactly equal, but their sum is kept close to $2\hbar\omega_{\mathrm{LO}}$. Note that due to the absence of the thin barriers, a parasitic level a appears in this design, which is close in energy to the upper radiative level 6. For the calculation shown in the figure, which does not include corrections in the level energies due to non-parabolicity (that can be empirically included by taking an energy-dependent effective mass [121, 97, 147]), $E_{6a} \sim 21$ meV. However, including the corrections due to non-parabolicity, which mainly affect the levels with energies high above the bottom of the quantum wells (such as the level a), E_{6a} becomes ~ 12 meV, whereas all the other relevant energy separations and values of energy anticrossings remain approximately the same. Consequently, the Δ_{6a} anticrossing moves much closer in bias space to the $\Delta_{1'6}$ anticrossing (the design bias). As will be discussed below, this has important implications on the performance of the laser.

Fig. 3-27(c) shows the experimental results from a FTP185C MM (Cu-Cu) ridge laser. In pulsed operation, $J_{\mathrm{th}} \sim 350$ A/cm^2 at 5 K and $T_{\mathrm{max,pul}} = 109$ K. Despite a relative high peak voltage (~ 20 V) this device lased up to ~ 72 K in cw operation, which is due to the improved thermal properties of the Cu-Cu bonding interface. Examination of the cw transport characteristics in this device reveals that the parasitic current channel indicated by the minimum in the dV/dI-I at $I \sim 0.05$ A barely increases in current between 7 K and 74 K, by which point lasing already ceases. Hence, an additional strong temperature degradation mechanism must be active, which is believed to be due to a large value of the parasitic anticrossing between levels 6 and a ($\Delta_{6a} \sim 1.9$ meV). The lifetime of the upper radiative level 6 in this design is small even at 7 K, as is apparent from a low value of $\Delta\mathcal{R}_{\mathrm{th}}/\mathcal{R}_{\mathrm{th}}$ (~ 0.32), which depends critically on the upper state lifetime (section 2.4.2). As the temperature increases, the $6-a$ coupling will be enhanced further due to the reasons mentioned in section 3.3.4, which explains the relatively smaller $T_{\mathrm{max,pul}}$ for this design. The roll-off in the L-I curves that occurs before the NDR for this laser is also possibly due to this parasitic coupling since level 6 moves closer in energy to level a as the applied voltage increases.

The two-phonon depopulated designs discussed in this section show no particular

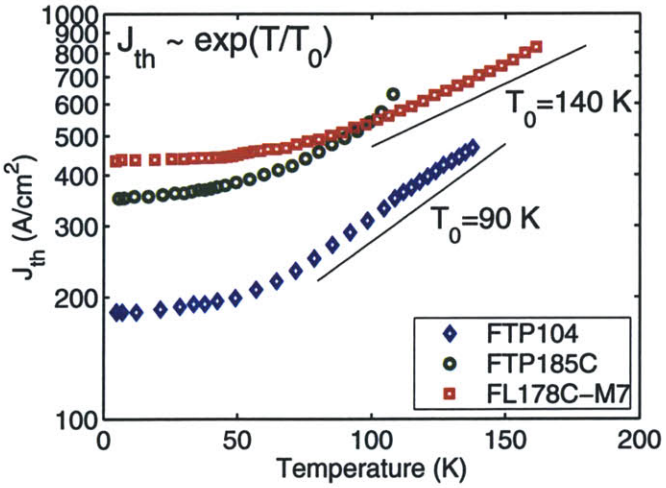


Figure 3-28: Semi-logarithmic plot of the J_{th}-T variation for the FTP104, FTP185C, and FL178C-M7 lasing devices fabricated with MM (Cu-Cu) waveguides. A phenomenological fit of the high-T data points to the expression $J_{\text{th}} \propto \exp(T/T_0)$ is also indicated. The data is for pulsed mode operation, when $T_{\text{active}} \sim T$, where T is the heat-sink temperature.

temperature improvement compared to the single-phonon depopulated FL-C series designs, which typically lased up to $130 - 170$ K in pulsed mode. Whereas the FTP160B-M1 and FTP185C designs are perceived to have additional problems that might be the cause for their worse performance, the FTP104 design should have performed better compared to the FL-C designs should thermal backfilling due to reabsorption of hot-phonons be the primary temperature degradation mechanism in these lasers. The highly doped top contact layer was not removed for the FTP104 results shown in Fig. 3-25(c); moreover, scope for further improvement in this design exists since it was only the first try, unlike the FL-C designs that went through many iterations. Regardless of this fact, FTP104 does not appear to be a better design as seen from the J_{th}-T variation in Fig. 3-28, since it has a lower T_0 when compared to the FL178C-M7 design, and consequently, has a smaller $T_{\text{max,pul}}$ despite the fact that both designs have a similar dynamic range in lasing current ($J_{\text{max}}/J_{\text{th}} \sim 2$).

Table 3.2 summarizes the experimental results from FTP series of devices. Note that FTP104-M1 was designed to be a minor modification of FTP104 by making the injector barrier thinner by 1 ML and by thickening one of middle the barriers in the module to limit the lower level parasitic current channels. These changes were

FTP series summary						
Device	$J_{\mathrm{th,5K}}$ ($\mathrm{A/cm^2}$)	$J_{\mathrm{max,5K}}$ ($\mathrm{A/cm^2}$)	$T_{\mathrm{max,pul}}$ (K)	$T_{\mathrm{max,cw}}$ (K)	ν (THz)	Waveguide
FTP104	170–210	330–380	140	110	2.65–2.80	MM (Cu-Cu)
FTP104-M1	210–240	400–420	128	97	2.35–2.70	MM (Cu-Cu)
FTP159B	No lasing	240-260	– –	– –	– –	MM (In-Au)
FTP160B-M1	650	720	57	– –	4.38–4.47	MM (In-Au)
FTP174B-M2	No lasing	820	– –	– –	– –	MM (Cu-Cu)
FTP185C	340–360	600-630	109	72	3.40–3.90	MM (Cu-Cu)

Table 3.2: Summary of experimental results from FTP series of terahertz QCLs. Each entry in the table corresponds to a different MBE wafer.

meant to increase the dynamic range $J_{\mathrm{max}}/J_{\mathrm{th}}$. However, the wafer was overgrown by 3.6%, which made all the layers including the barriers thicker and thus nullified those changes. Consequently, no performance improvement were observed as compared to FTP104 (which was grown approximately to the same thickness as designed).

3.4.2 Designs with direct-phonon depopulation and a diagonal radiative transition: DSL series

As was discussed in section 3.3.4, the upper state lifetime due to LO-phonon emission $\tau_{ul,\mathrm{LO}}$ depends critically on the electron temperature T_{e} in the upper state u due to thermally activated LO-phonon scattering. As $T_{\mathrm{e}} \uparrow$, $\tau_{ul,\mathrm{LO}} \rightarrow \tau^{\mathrm{hot}}_{ul,\mathrm{LO}}$ (equation 3.7), where $\tau^{\mathrm{hot}}_{ul,\mathrm{LO}}$ depends critically on the spatial overlap of the upper and lower radiative states. A more diagonal radiative transition has a larger value for $\tau^{\mathrm{hot}}_{ul,\mathrm{LO}}$ and thus can maintain population inversion up to higher operating temperatures. For example, for the FL-C designs in section 3.3.1 (Fig. 3-7) that have a more vertical radiative transition, $\tau^{\mathrm{hot}}_{ul,\mathrm{LO}} \sim 0.5$ ps, whereas for the simple 3-level QCL design in Fig. 3-3(a) that has a diagonal radiative transition, $\tau^{\mathrm{hot}}_{ul,\mathrm{LO}} \sim 0.75$ ps. The DSL series designs are similar in concept to the 3-level QCL design of Fig. 3-3(a) but with additional wells in both the injector as well as the active region to improve the design characteristics. As in the 3-level design, the lower radiative state is directly depopulated via LO-phonon emission as opposed to the resonant-phonon depopulation scheme for the FL designs,

DSL series designs		
	2-well injector region	3-well injector region
1-well active region	DSL228G-M1 • DSL222G-M2	DSL190G
2-well active region	• DSL209F-M1 DSL209F-M2 • DSL209F-M3 DTP113*	DSL176F
3-well active region	DSL210E • DSL203E-M1	
4-well active region	DSL184 DSL201D DSL207D-M1	
5-well active region	DSL201C	

Table 3.3: Number of quantum-wells in the active and the injector regions for the various DSL series terahertz QCL designs that were experimentally realized. The bulleted designs achieved lasing.
*DTP113 is a design with a two-phonon depopulation scheme similar to that of FTP104 (Fig. 3-25), whereas all others are designed with single-phonon depopulation.

and the radiative transition is diagonal (i. e. involving two states that are primarily localized in two different quantum-wells). Consequently, for the DSL designs, τ_l is lowered and τ_{ul} is increased, both of which aid in obtaining a large population inversion. However, a diagonal interwell transition affords a smaller value of the radiative oscillator strength as opposed to a vertical intrawell transition, and also can lead to greater inhomogeneous broadening of the gain linewidth as was discussed for the M100 design in section 3.1. Therefore, it is challenging to obtain lasing in such a design. But once a robust DSL laser with a low-temperature dynamic range $J_{\max} - J_{\text{th}}$ rivaling that of the best FL-C lasers can be obtained, a DSL design should provide a more robust temperature performance.

The "D" in "DSL" stands for diagonal radiative transition, whereas "SL" signifies a superlattice based active region. The original DSL designs were designed with multiple wells in the active region within a QCL module (thus forming a sub-superlattice within each module of the overall QCL superlattice) to obtain a large radiative oscillator strength; however, this acronym has since lost its meaning because the most successful DSL series lasers are the ones with fewest number of wells in the

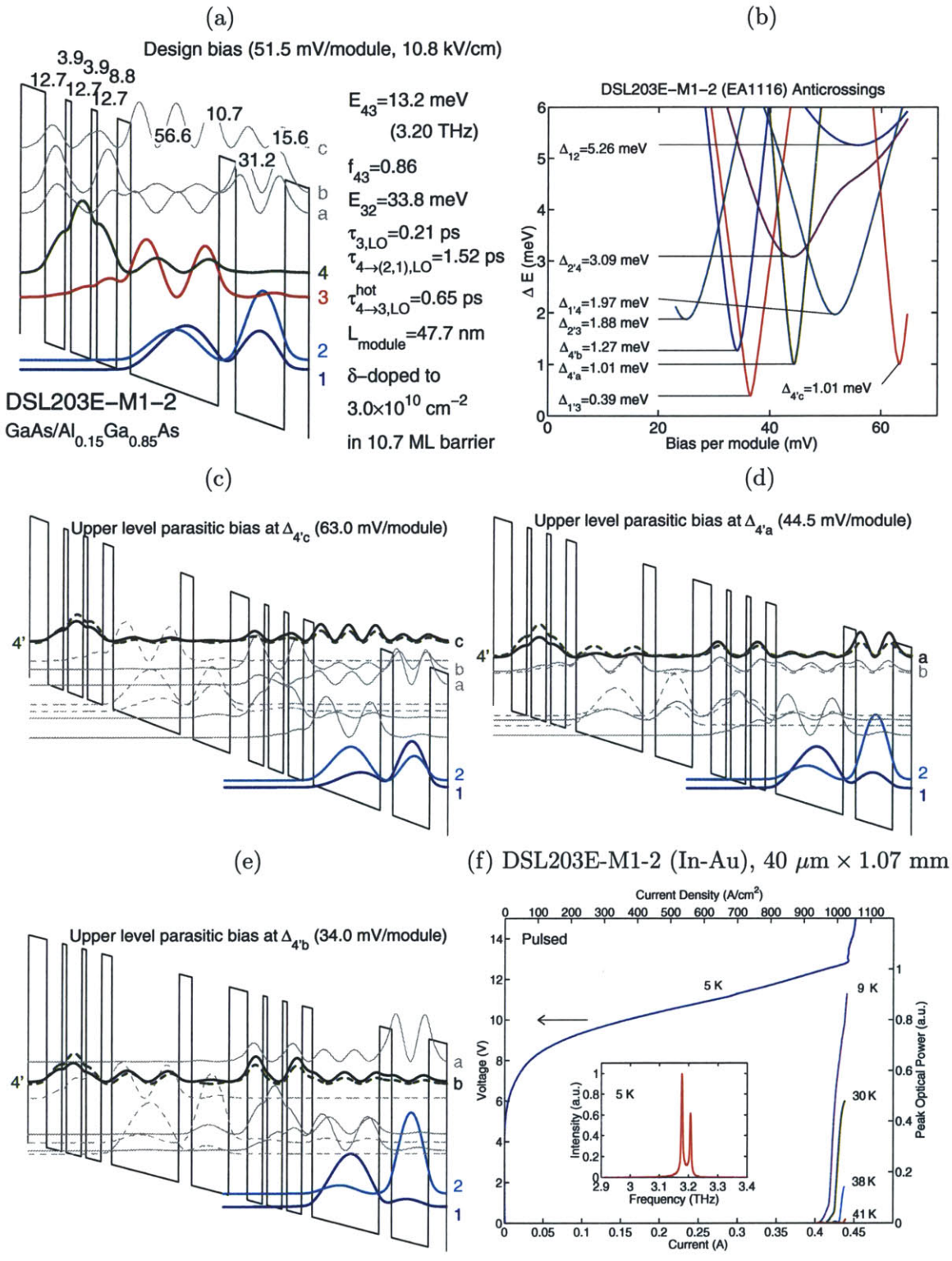


Figure 3-29: Design details and experimental results for DSL203E-M1-2. (a) One module conduction band diagram at the design bias. The radiative transition is from $4 \rightarrow 3$. (b) Anticrossing plots for a two module calculation. (c),(d),(e) Two module conduction band diagrams at various upper level parasitic anticrossings. (e) Pulsed measurement results from a MM (In-Au) ridge laser. No cw lasing was observed.

active region, the reasons for which will become apparent in the following discussions. Table 3.3 lists the names of different DSL designs that were experimentally realized. Only a selected few of these designs, predominantly the lasing ones, will be described in the following paragraphs. The designs to be discussed are chosen so as to cover the underlying operating principles of all the DSL designs.

The design details and experimental results for DSL203E-M1-2 (the second MBE growth for the DSL203E-M1 design), which was the first design in the DSL series to achieve lasing, are shown in Fig. 3-29. This design features a 3-well active region and a 2-well (2-level) injector. Due to multiple wells in the active region, the upper radiative level 4 is spatially extended across the module yielding a relatively large radiative oscillator strength ($f_{43} \sim 0.86$) for a diagonal transition. Notice, however, that very thin barriers have to be used in the active region, which might cause greater inhomogeneous level broadening of level 4 due to growth fluctuations. This design has three high energy parasitic levels a, b, and c as seen from Fig. 3-29(a). All of these parasitic levels have relatively large parasitic anticrossings with the upper radiative level $4'$ in the adjacent module as determined from the anticrossing plots of Fig. 3-29(b). These parasitic anticrossings are analogous to the ones described in section 3.2.1 and thus have a detrimental effect on QCL operation. This can be seen from the experimental results shown in Fig. 3-29(f), whereby, lasing was observed for only a small region in the I-V and the overall performance of these lasers was poor.

The second design to achieve lasing in the DSL series was DSL209F-M1, which featured a two-well active region and a two-well injector. The next modification of that design, named as DSL209F-M2, had a 1-ML thicker collector barrier as compared to DSL209F-M1, and failed to lase. Note that, unlike in the FL designs, the collector barrier in the DSL designs also influences the radiative energy separation, which reduces for a thicker barrier due to a smaller anticrossing splitting between the radiative levels. Additionally, the radiative oscillator strength also decreases for a thicker barrier due to a reduced spatial overlap between those levels. In the next design modification, named as DSL209F-M3, the collector barrier thickness was restored to original, whereas, predominantly, the intra-injector barrier was made thicker by

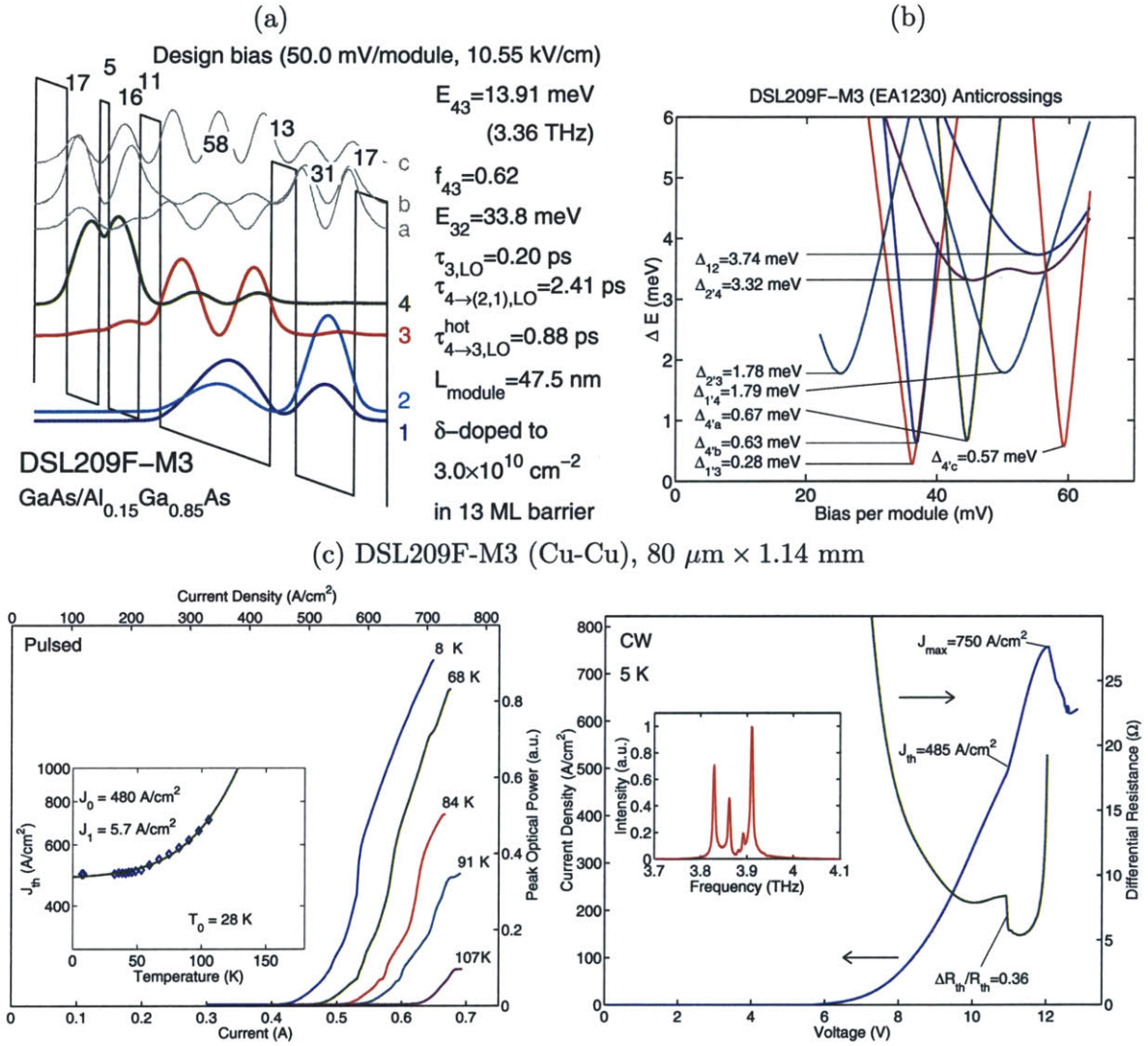


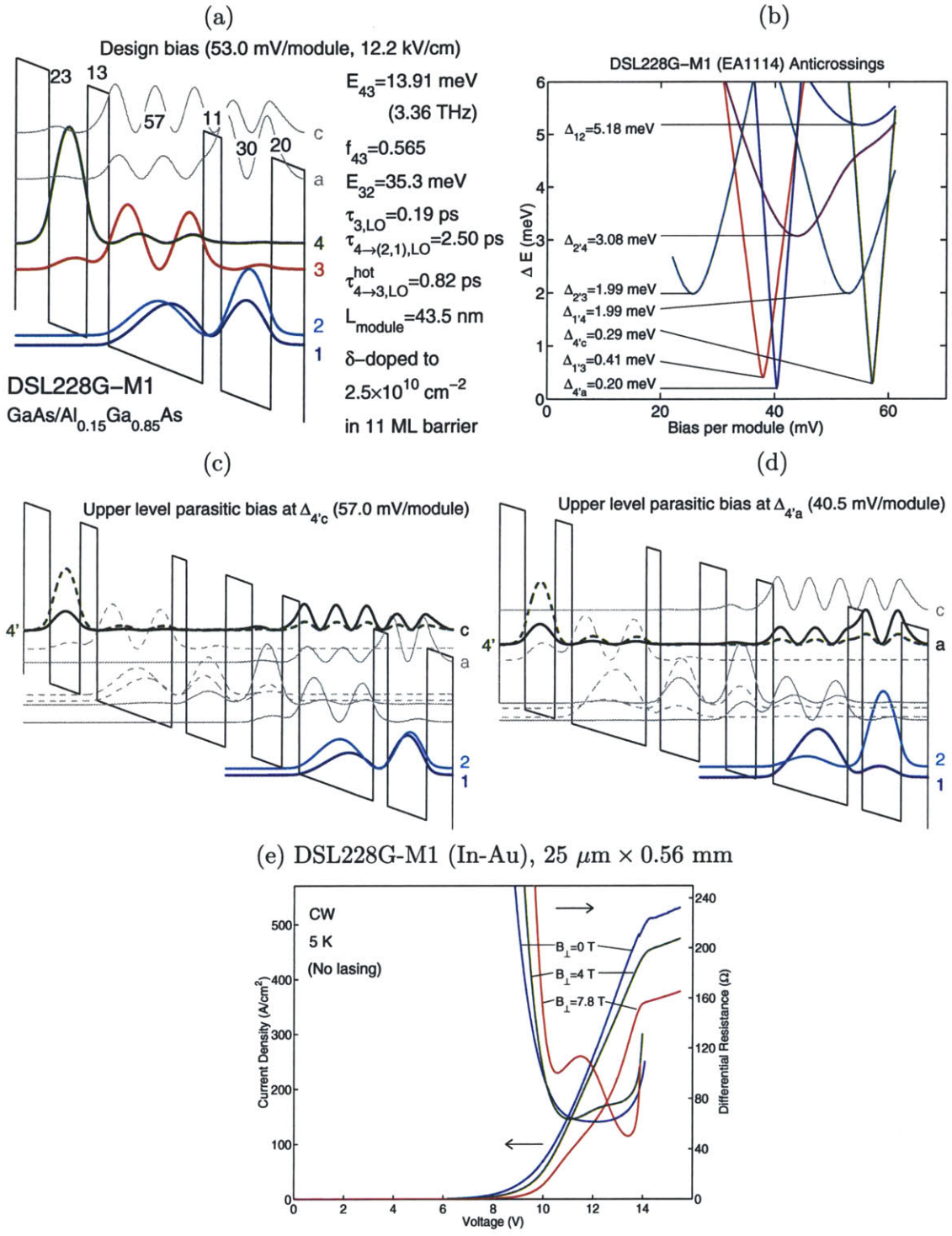
Figure 3-30: Design details and experimental results for DSL209F-M3. (a) One module conduction band diagram at the design bias. The radiative transition is from $4 \rightarrow 3$. (b) Anticrossing plots for a two module calculation. (c) Pulsed and cw measurement results from a MM (Cu-Cu) ridge laser. For this device, cw L-Is were not measured and thus $T_{\max,\mathrm{cw}}$ is not available.

2-ML. DSL209F-M3 performed only marginally better as compared to DSL209F-M1 as can be noted from the results summary in Table 3.4.

Fig. 3-30 shows the design details and experimental results for DSL209F-M3. This design has a significantly smaller oscillator strength ($f_{43} \sim 0.62$) as compared to that for DSL203E-M1 ($f_{43} \sim 0.86$) due to one fewer wells in the active region. However, since the wavefunctions are less spatially extended, the barriers in DSL209F-M2 are thicker for similar set of design parameters. This reduces the upper level parasitic anticrossings in this design as compared to those in DSL203E-M1. Consequently, DSL209F-M3 lasers performed much better than the DSL203E-M1-2 lasers, as shown in Fig. 3-30(c). However, when compared to the FL-C designs with similar values of $J_{\max} - J_{\text{th}}$ (Tables 3.1 and 3.4), the performance of DSL209F-M3 lasers was worse. This is attributed to the aforementioned parasitic current channels due to levels a, b, and c, respectively, which, despite being less prominent than in DSL203E-M1, are still significant as can be noted from the corresponding anticrossing values in Fig. 3-30(b).

The gain linewidth for DSL designs is expected to be larger as compared to that for the FL-C designs due to enhanced lifetime broadening of the direct-phonon depopulated lower radiative state, and also due to a inhomogeneously broadened interwell radiative transition. Therefore, the discontinuity in the differential resistance at threshold $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}}$, which is proportional to the population inversion needed to meet gain threshold (section 2.4.2), is also expected to be higher. However, a low-temperature value of $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}} = 0.36$ was measured for a DSL209F-M3 laser, which is smaller than the typical values obtained for the FL-C lasers. This confirms that the upper state lifetime in this design is small as predicted by equation (2.95).

The aforementioned upper level parasitic couplings can be significantly weakened by reducing the number of wells in the active region to just one. This was done for a subsequent design, named as DSL228G-M1, the design details and experimental results for which are shown in Fig. 3-31. DSL228G-M1 features a two-well injector and a one-well active region. The upper parasitic level b in the DSL203E-M1 and DSL3209F-M1 designs was due to the second excited state in the quantum-wells that constituted their active region. For DSL228G-M1, this level is not a confined state



(e) DSL228G-M1 (In-Au), 25 μm $\times$ 0.56 mm

Figure 3-31: Design details and experimental results for DSL228G-M1. (a) One module conduction band diagram at the design bias. The radiative transition is from $4 \rightarrow 3$. (b) Anticrossing plots for a two module calculation. (c),(d) Two module conduction band diagrams at the dominant parasitic anticrossings. (e) CW I-Vs and $\mathcal{R}$-Vs from a MM ridge laser immersed in liquid Helium in a superconducting magnetic dewar, measured at different values of the magnetic field $B_\perp$ applied perpendicular to the layers of the MQWs. No lasing was observed for any of the DSL228G-M1 devices.

anymore since its energy is pushed over the top of the barriers due to a narrower active region. Consequently, the parasitic levels a and c are not drawn towards the upper radiative level $4'$ in the adjacent module, like they are for DSL203E-M1 and DSL209F-M3 due to a resonant interaction with the level b. This can be seen from the parasitic anticrossing conduction band diagrams by comparing Figs. 3-31(c,d) for the DSL228G-M1 design to the Figs. 3-29(c,d) for the DSL203E-M1 design. Consequently, the parasitic anticrossings $\Delta_{4'a}$ and $\Delta_{4'c}$ for DSL228G-M1 are much lower and well within the acceptable range, and therefore, not likely to cause degradation in the lifetime of upper radiative level 4 (section 3.2.1, Fig. 3-4).

Despite the absence of any strong parasitic current channels, DSL228G-M1 did not achieve lasing, which is likely due to two reasons. First, the radiative oscillator strength in DSL228G-M1 is low ($\sim$ 0.56) compared to the previous two successful DSL designs due to its narrower active region. Second, the $2' - 4$ injector anticrossing occurs very close in bias space to the $1' - 3$ parasitic anticrossing, which can enhance current flow at that bias ($\sim$ 40 mV/module) and thus cause an early NDR before enough population inversion is established. This can be seen from the experimentally measured I-Vs and $\mathcal{R}$-Vs for a MM ridge laser at 5 K as shown in Fig. 3-31(e). For this particular device, the measurements were performed in a magnetic dewar for various values of the magnetic field $B_\perp$ applied perpendicular to the layers of the MQWs. As was explained in section 3.3.1 (Fig. 3-14), application of $B_\perp$ limits intersubband scattering thereby decreasing the lifetime broadening of the levels, and makes resonant-tunneling more specular due to reduced dephasing. Consequently, the valleys in $\mathcal{R}$-Vs due to various current channels are enhanced. It is speculated that the lower bias valley in the $\mathcal{R}$-Vs of Fig. 3-31(e) is due to $2' \rightarrow 3$ resonant-tunneling, and the higher bias valley is due to $2' \rightarrow 4$ resonant-tunneling, beyond which the device gets biased into the NDR region.

The DSL228G-M1 design was modified to DSL222G-M2 in the next iteration. The design details and experimental results for DSL222G-M2 are shown in Fig. 3-32. This design was primarily modified in two ways. First, the intra-injector barrier was thickened significantly (Δ_{12} decreased from 5.2 meV in DSL228G-M1 to 3.3 meV in

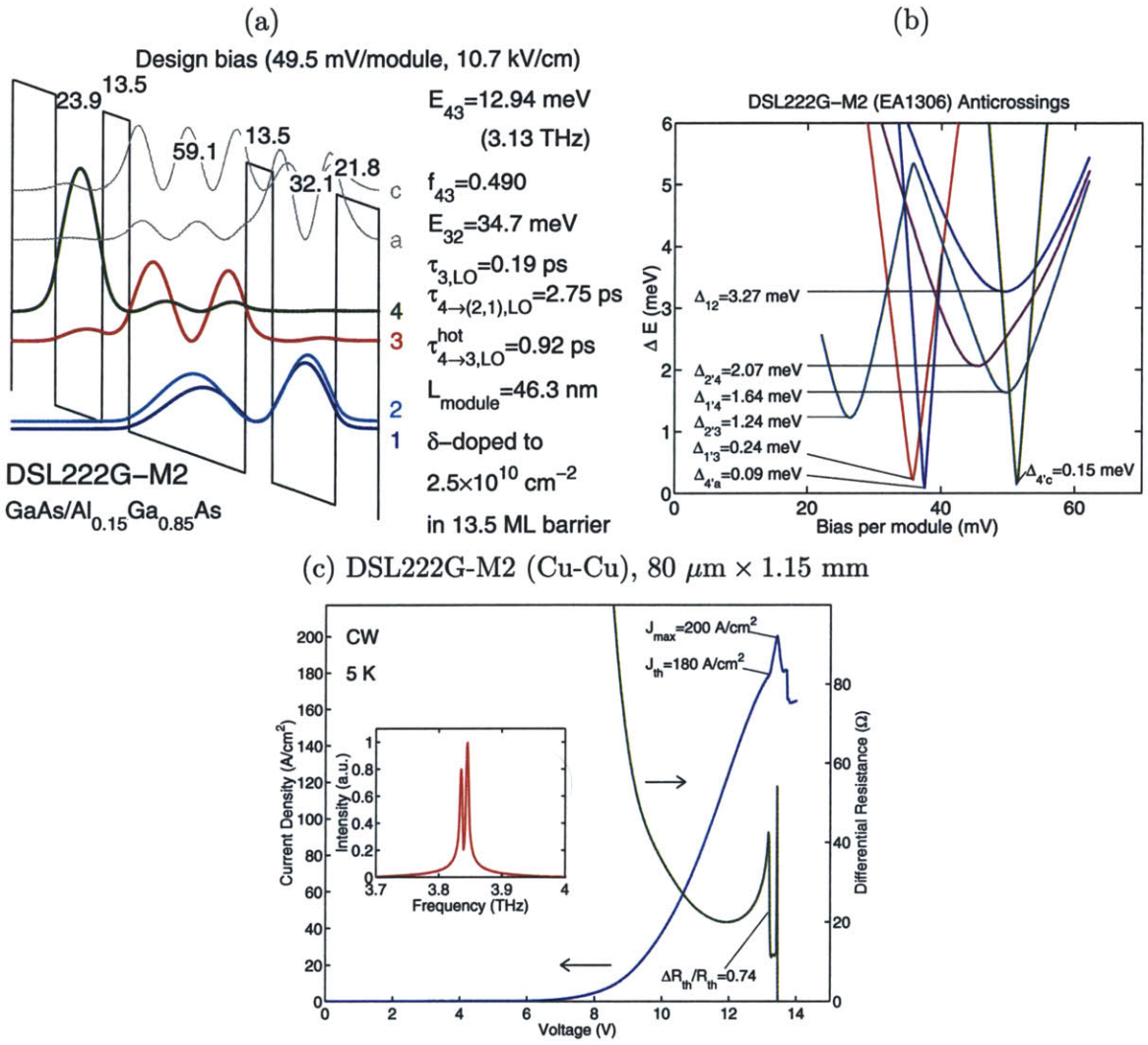


Figure 3-32: Design details and experimental results for DSL222G-M2. (a) One module conduction band diagram at the design bias. The radiative transition is from $4 \rightarrow 3$. (b) Anticrossing plots for a two module calculation. (c) CW I-V and $\mathcal{R}$-V measured at 5 K from a MM (Cu-Cu) ridge laser. This device lased up to $\sim$ 27 K in cw operation. No lasing was observed in pulsed operation due to a very small lasing range in current, which may not be accessible by short pulses. This is because a small overshoot typically exists at the beginning of a pulse, which can push the device into the NDR region at the beginning of the pulse itself when the current in the pulse reaches close to the lasing threshold.

DSL222G-M2), which brought $1'-4$ and $2'-4$ anticrossings closer to each other in bias space, and also reduced all the parasitic anticrossings. Second, the $1-2$ anticrossing was brought closer to the design bias (i. e. the $1'-4$ anticrossing) in bias space by changing the relative width of the two injector wells so that the current flow at the design bias is maximized. This can be seen from a comparison of the anticrossing plots in Figs. 3-31(b) and 3-32(b). With these changes, lasing was achieved for the DSL222G-M2 design as shown for a MM ridge device in Fig. 3-32(c). The relatively low threshold current density in this device ($J_{\mathrm{th,5K}} \sim 180$ A/cm^2) is due to a significant reduction in both the lower and upper level parasitic current channels. The fact that any prominent valleys are absent in the measured $\mathcal{R}$-V indicates negligible lower level parasitic couplings. A large value of $\Delta\mathcal{R}_{\mathrm{th}}/\mathcal{R}_{\mathrm{th}}$ in DSL222G-M2 (~ 0.74) as compared to that in DSL209F-M3 (~ 0.36) indicates a relatively long lifetime of the upper radiative state in DSL222G-M2, which is suggestive of negligible upper level parasitic couplings. It may be noted that a value of $\Delta\mathcal{R}_{\mathrm{th}}/\mathcal{R}_{\mathrm{th}} \sim 0.74$ is believed to be the maximum such value recorded for any terahertz QCL, which is possibly due to a broad gain linewidth in this laser that requires a large population inversion to obtain lasing threshold (equation 2.95).

The fact that such a low value of $J_{\mathrm{th,5K}}$ is obtained for DSL222G-M2, even though it has a small radiative oscillator strength ($f_{43} \sim 0.49$) and perceivably a large gain linewidth, bodes well for the operating principle of this design. However, two facts about its experimental results are not clearly understood. First, the peak current density $J_{\mathrm{max}} \sim 200$ A/cm^2 for this laser is much lower than expected even if it is assumed that the NDR happens prematurely right after the $2'-4$ anticrossing. As a consequence, the lasing range in DSL222G-M2 is extremely small ($J_{\mathrm{max}} - J_{\mathrm{th}} \sim 20$ A/cm^2), which leads to its poor temperature performance ($T_{\mathrm{max,cw}} \sim 27$ K). In contrast, the FL176C-M4-2 lasers that had a similar injector design as in DSL222G-M2 (Fig. 3-10) showed typical values of $J_{\mathrm{max}} \sim 450$ A/cm^2 even though they were believed to have a premature NDR right after the $2'-5$ anticrossing (Fig. 3-15). Second, the observed lasing frequency for DSL222G-M2 (~ 3.85 THz) is much greater than that estimated from the band diagram (~ 3.35 THz). This behavior was also observed

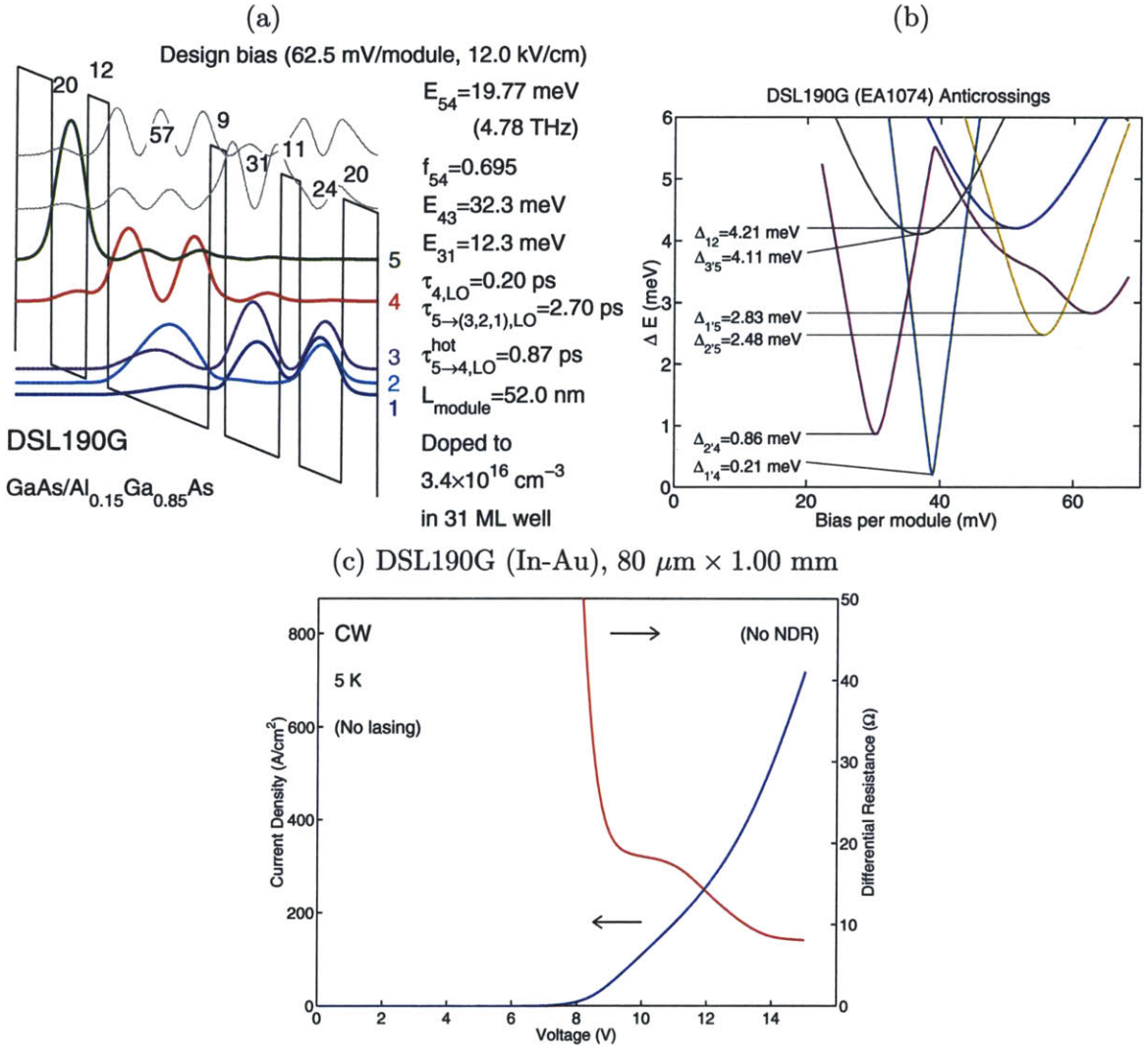


Figure 3-33: Design details and experimental results for DSL190G. (a) One module conduction band diagram at the design bias. The radiative transition is from $5 \rightarrow 4$. (b) Anticrossing plots for a two module calculation. (c) CW I-V and $\mathcal{R}$-V measured at 5 K from a MM (In-Au) ridge laser. No lasing was observed either in pulsed or cw operation for any of the DSL190G devices that were tested. Also, no NDR region was observed in these measurements, the cause of which is not well understood.

for DSL209F-M3, but not for DSL203E-M1-2. For the FL-C and FTP series lasers, the experimentally measured values are typically well within 10% of simulated values. The cause for this discrepancy is not yet clear. A self-consistent Schrödinger-Poisson calculation for the band-diagram yielded negligible band-bending and the calculated value of E_{43} remained approximately the same. It is possible that this is due to some unknown effect caused by δ-doping in the DSL series wafers. A continuous doping in the wide injector well, like in the FL-C devices, should probably be tried in a future design iteration of the DSL222G-M2 design. Also the intra-injector barrier can be thinned, but not as thin as that for DSL228G-M1, since the parasitic current channels are not large for the DSL222G-M2 design. Also, to maximize the lasing range in current, the injector doublet may not need to be designed as upwards diagonal as in DSL228G-M1.

DSL series summary						
Device	$J_{th,5K}$ (A/cm^2)	$J_{max,5K}$ (A/cm^2)	$T_{max,pul}$ (K)	$T_{max,cw}$ (K)	ν (THz)	Waveguide
DSL203E-M1	In NDR	1150	– –	– –	3.2	MM (In-Au)
DSL203E-M1-2	970	1030	41	– –	3.2	MM (In-Au)
DSL209F-M1	450–475	650–675	88	28	3.80–3.85	MM (In-Au)
DSL209F-M2	No lasing	675	– –	– –	– –	MM (In-Au)
DSL209F-M3	450	700	105	58	3.80–3.90	MM (Cu-Cu)
DSL228G-M1	No lasing	500–600	– –	– –	– –	MM (In-Au)
DSL222G-M2	180	200	– –	27	3.85	MM (Cu-Cu)

Table 3.4: Summary of experimental results from DSL series of terahertz QCLs. Only the most relevant designs are included. Each entry in the table corresponds to a different MBE wafer.

It may be noted that a three-well injector, one-well active region design, named as DSL190G, has also been tried but no lasing was observed. The design details and experimental results for DSL190G are shown in Fig. 3-33. From a design perspective, the DSL190G design is similar to the two-well injector designs DSL228G-M1 and DSL222G-M2, but with the lower-bias parasitic anticrossings even smaller due to an additional well in the injector. Also it does not have any dominant upper level parasitic channels. However, no lasing was obtained for the DSL190G lasers. Moreover, no NDR was observed in the I-Vs, which is rather inexplicable but for an error in

the MBE growth. A slight variation of this design with a somewhat less downwards diagonal injector alignment can also be possibly tried in a future a iteration. Since a three-well injector region affords smaller parasitic couplings, the radiative oscillator strength can be further increased for such a design by thinning the collector barrier.

3.4.3 Modified FL design with a four-well active region: FL-L series

The oscillator strength of the radiative transition in the FL design of Fig. 3-7 can be enhanced considerably by increasing the number of wells in the active region, which spatially extends the radiative levels over a greater length. The first of a design with a four-well active region and a two-well injector was named as FL148L and is described in chapter 6 of Williams [166]. Although no lasing was obtained from that design, a super-linear L-I and narrow linewidth ($\Delta\nu_{\mathrm{FWHM}} \sim 0.42$ THz) electroluminescence was measured, which was indicative of gain in the device. After two design iterations, lasing was achieved in FL163L-M2, the designs details and experimental results of which are shown in Fig. 3-34. FL163L-M2 improved upon FL148L by making the barriers in the active region as well as the collector barrier thinner, which made both the injector and collector anticrossings larger.

As seen from Fig. 3-34(a), the oscillator strength in FL163L-M2 ($f_{54} \sim 1.24$) is much larger than for the FL-C designs ($f_{54} \sim 0.9$). However, this design requires very thin barriers in the active region, which possibly leads to greater inhomogeneous broadening of the radiative levels. Moreover, additional upper parasitic levels (levels 7 and 9) are introduced in the band diagram, which leads to an increase in the parasitic current channels of the type discussed in section 3.2.1. Hence, the operating differences between FL163L-M2 and the FL-C designs are similar to those between DSL203E-M1 and DSL222G-M2 that were discussed in section 3.4.2. For example, the parasitic anticrossing $\Delta_{5'8} \sim 0.5$ meV for FL163L-M2 is relatively large when compared to the analogous value of $\Delta_{5'7} \sim 0.17$ meV for the FL178C-M7 design, where the level numberings are according to Fig. 3-34(a), and Fig. 3-36, respectively.

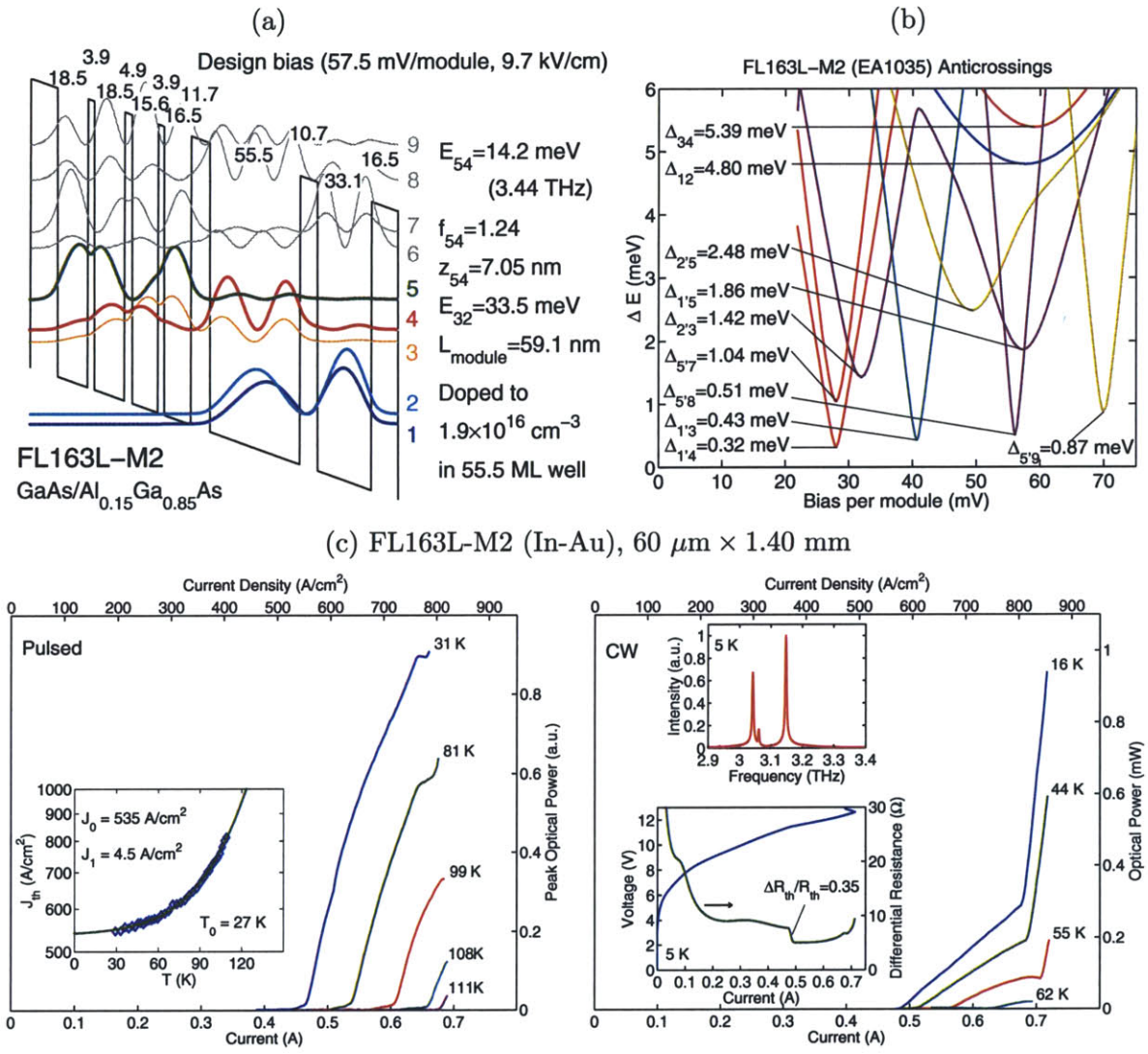


Figure 3-34: Design details and experimental results for FL163L-M2. (a) One module conduction band diagram at the design bias. The radiative transition is from $5 \rightarrow 4$. This design is similar to the FL design in Fig. 3-7, except that it has a four-well active region even though the number of relevant levels remain the same. (b) Anticrossing plots for a two module calculation. (c) Pulsed and cw measurement results from a MM (In-Au) Fabry-Pérot ridge laser.

A value of $T_{\text{max,pul}} \sim 111$ K was measured for one of the FL163L-M2 lasers, as shown in Fig. 3-34(c). But the high oscillator strength advantage in FL163L-M2 is seemingly offset by its aforementioned negative features, leading to a worse performance as compared to that of the FL-C designs. A value of $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}} \sim 0.35$, which is lower than the typical value of $\Delta\mathcal{R}_{\text{th}}/\mathcal{R}_{\text{th}} \gtrsim 0.5$ observed for the FL-C lasers is indicative of a small upper state lifetime in the FL163L-M2 laser.

3.4.4 Modified FL design with a three-well injector region: FL-I series

A multiple-well injector region can prevent current flow in the QCL prior to the design bias since the injector levels are misaligned with respect to each other at low bias. This can minimize the lower level parasitic current channels, thereby potentially increasing the lasing range in current. This forms the basis for the FL-I series designs, which have a three-well injector region and correspondingly three injector levels. The first successful laser from this series of designs was named as FL148I-M1 and is discussed in chapter 7 of Williams [166]. Whereas a value of $T_{\text{max,pul}} \sim 78$ K was obtained for FL148I-M1, the performance has since been improved in a subsequent design named as FL157I-M4 with $T_{\text{max,pul}} \sim 109$ K, as shown for a MM ridge laser in Fig. 3-35(c). For FL157I-M4, the injector levels 1 and 2 are in anticrossing at the design bias (as specified by the $1' - 6$ anticrossing), whereas for FL148I-M1, the injector levels 2 and 3 are in anticrossing at the design bias. This could be the reason why FL157I-M4 performed better, since each of the levels 1 and 2 are stipulated to have a greater population than the level 3 due to their lower energy, and hence, an injector doublet consisting of levels 1 and 2 is likely to produce more population inversion.

Even after multiple iterations in the FL-I series, the best performing FL157I-M4 design is far worse in performance as compared to the FL-C designs. This is despite the fact that the two dominant lower level parasitic anticrossings due to the $1' - 3$ and $2' - 3$ couplings are markedly reduced for the FL-I designs. One possible reason for its worse performance could be the additional optical loss in this structure due

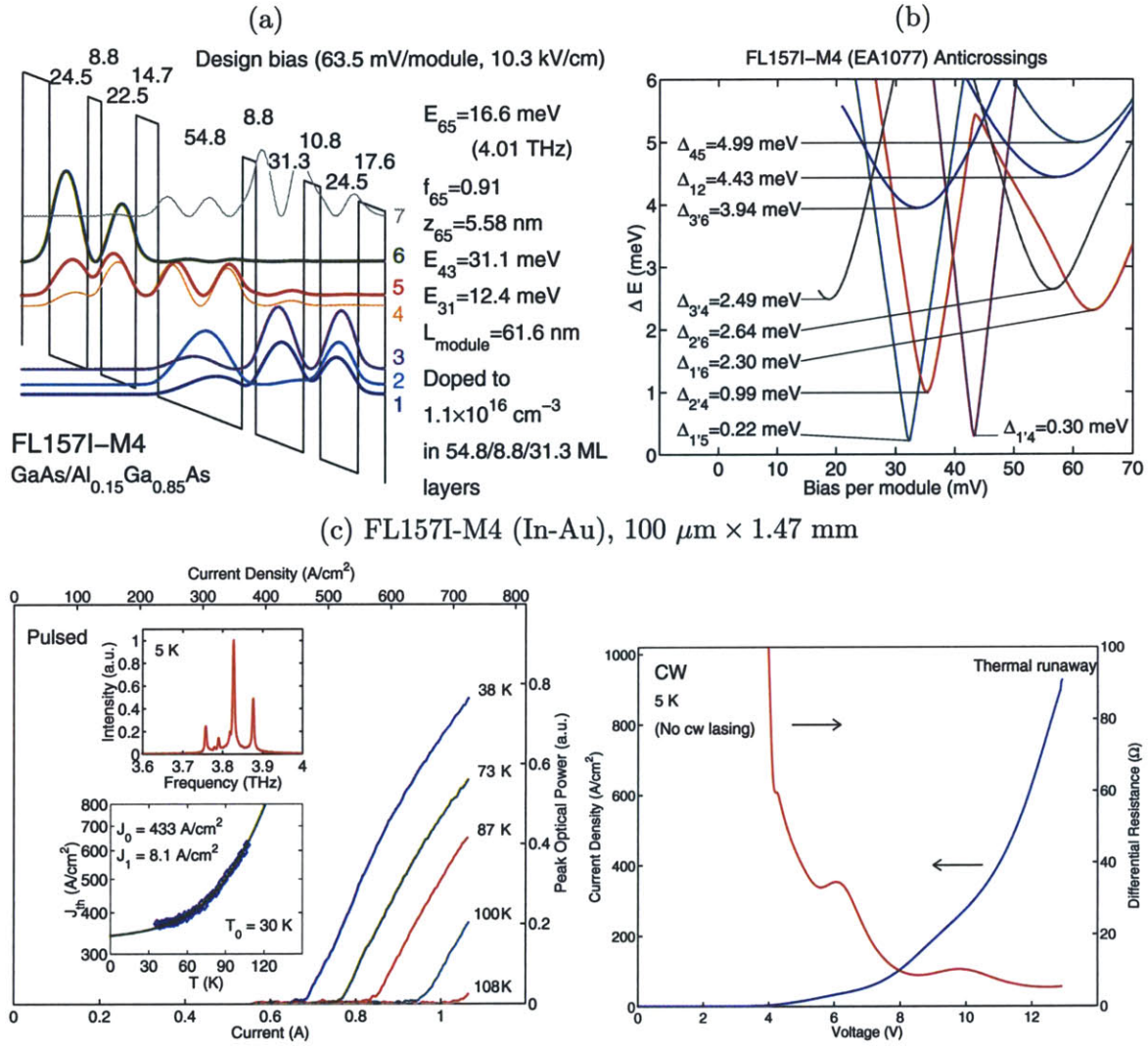


Figure 3-35: Design details and experimental results for FL157I-M4. (a) One module conduction band diagram at the design bias. The radiative transition is from $6 \rightarrow 5$. This design is similar to the FL design in Fig. 3-7, except that it has three wells in the active region, and thus it has three injector levels instead of two. (b) Anticrossing plots for a two module calculation. (c) Pulsed and cw measurement results from a MM (In-Au) Fabry-Pérot ridge laser. No cw lasing was observed for this device due to a poor metal-metal (In-Au) bonding interface. The device was destroyed due to over heating in cw operation before it could be biased up to the NDR.

to the $1 \rightarrow 3$ intersubband absorption, since the value of $E_{31} \sim 12.4$ meV is close to the radiative energy ($\hbar\omega \approx 16$ meV). Moreover, since the injection into the upper state 6 still happens via resonant-tunneling rather than intersubband scattering, the additional injector level 3 has a lesser role to play at the design bias; rather, it causes a smaller population of electrons to be available for injection through the $1-2$ doublet. Comparing the pulsed L-Is for FL157I-M4 in Fig. 3-35(c) to those for the FL-C designs in, for example, Fig. 3-20, the value of the characteristic temperature $T_0 \sim 30$ K for FL157I-M4 is significantly smaller than the typical values of $T_0 \sim 40-60$ K for the FL-C designs. This is likely due to a large $3' \rightarrow 4$ parasitic coupling ($\Delta_{3'4} \sim 2.5$ meV) in FL157I-M4, which can possibly cause the low-bias parasitic current to increase more rapidly with temperature due to a possible increase in the population of the injector level $3'$ at higher temperatures. In that case, the threshold current densities are *artificially* enhanced, rather than caused by a reduction in the population inversion or an increase in the optical losses.

3.4.5 FL designs in GaAs/$Al_{0.3}Ga_{0.7}As$ material system: FL-P and SFL series

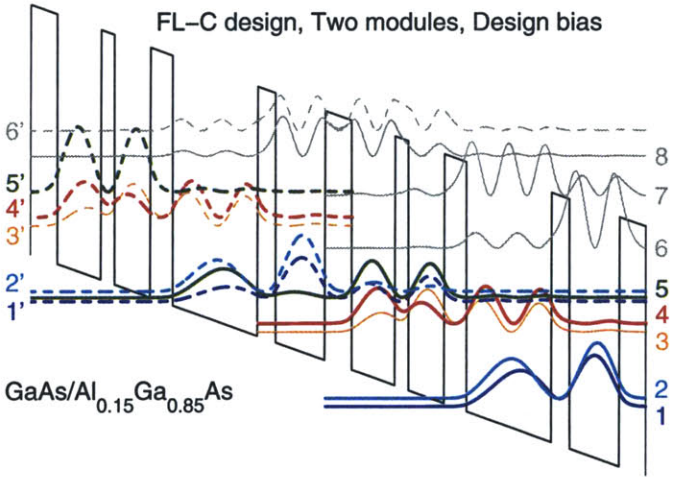


Figure 3-36: Design bias conduction band diagram for a typical FL-C series design of section 3.3.1 (two module calculation).

Figure 3-36 shows the design bias conduction band diagram for a typical FL-C

series design. Due to a relatively small barrier height in the $GaAs/Al_{0.15}Ga_{0.85}As$ material system (~ 0.135 meV), it is possible for the electrons in the upper radiative level 5′ to tunnel into the continuum over the top of the potential barriers as they gain enough thermal energy at higher temperatures, thus leading to a reduction in the upper level lifetime with increasing temperature. Also, scattering into the low-lying upper parasitic levels is also a possibility, for example, $E_{85'} \sim 19$ meV in Fig. 3-36 for the FL178C-M7 design of Fig. 3-10. This section discusses two different types of FL designs that were implemented in the $GaAs/Al_{0.3}Ga_{0.7}As$ material system to investigate the role of barrier-height and the upper parasitic levels on QCL operation.

FL184P-M1

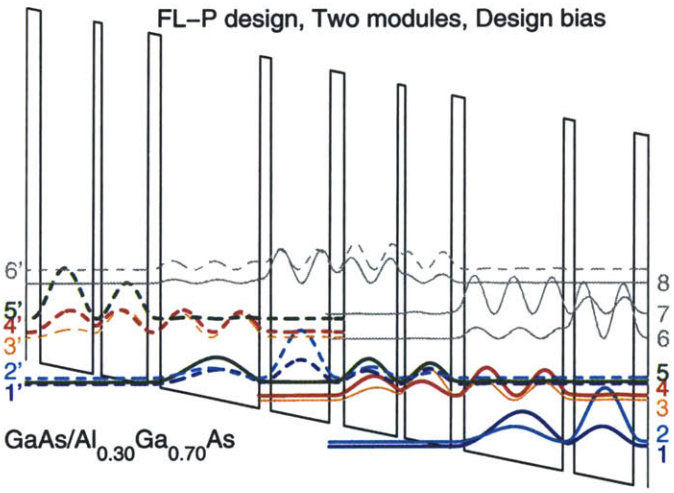


Figure 3-37: Design bias conduction band diagram for the FL-P design (two module calculation). This is similar to the FL-C design of Fig. 3-36, except for the $GaAs/Al_{0.3}Ga_{0.7}As$ material system that has a barrier height twice as much as in $GaAs/Al_{0.15}Ga_{0.85}As$. Levels with energies higher than that of level 6′ are not shown.

Figure 3-37 shows the design bias conduction band diagram for the FL-P design, which is similar to the FL-C design of Figs. 3-7 and 3-36, except that it is in the $GaAs/Al_{0.3}Ga_{0.7}As$ material system. This material system offers two distinct advantages. First, the possibility of electrons tunneling to the continuum states is much reduced due to a higher potential barriers (barrier height ~ 0.27 meV). Second, the separation between the upper parasitic level 8 and the upper radiative level 5′ is in-

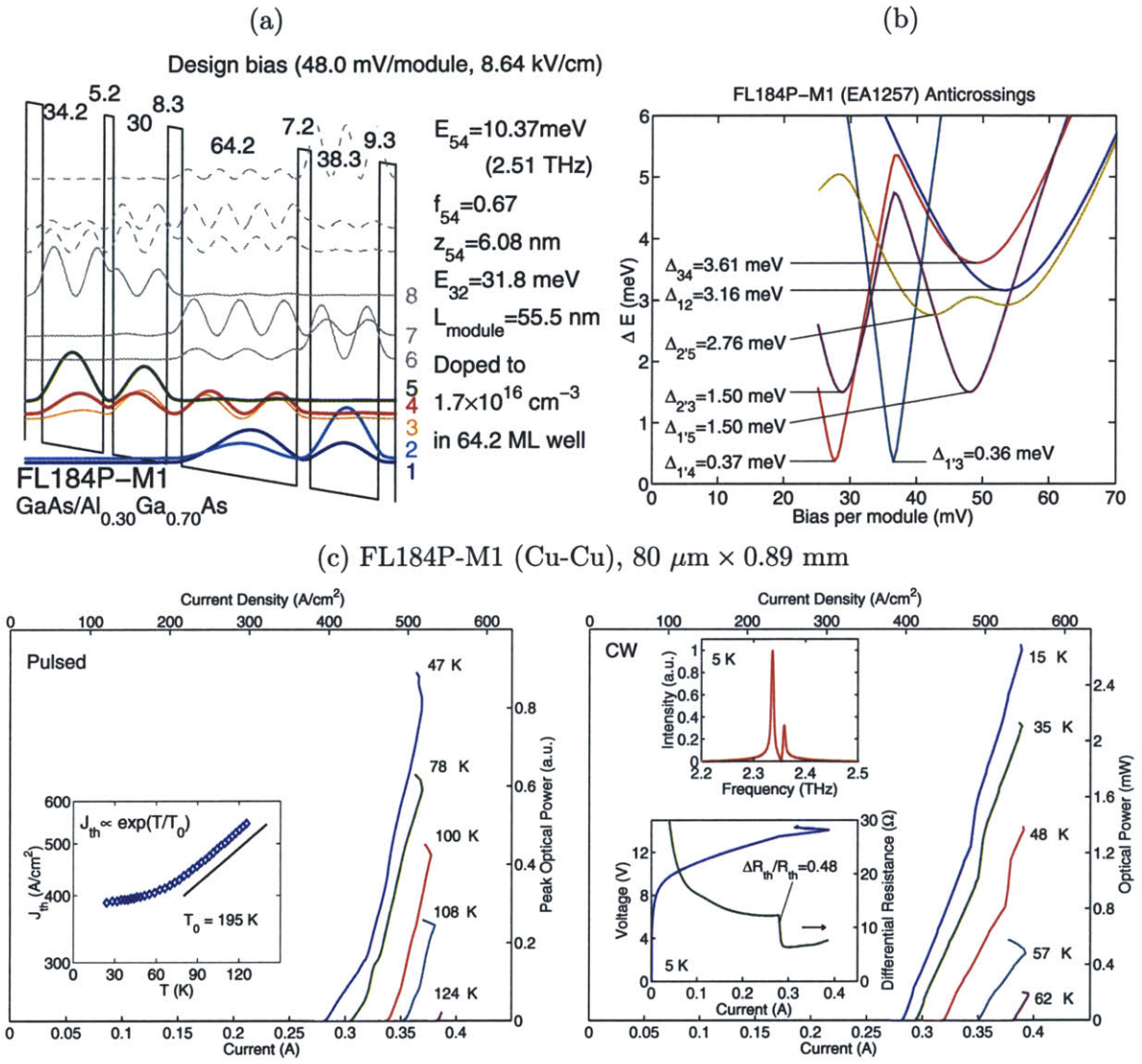


Figure 3-38: Design details and experimental results for FL184P-M1. (a) One module conduction band diagram at the design bias. The radiative transition is from $5 \rightarrow 4$. This design is similar to the one in Fig. 3-7, except that it is in the $GaAs/Al_{0.3}Ga_{0.7}As$ material system. (b) Anticrossing plots for a two module calculation. (c) Pulsed and cw measurement results from a MM (Cu-Cu) Fabry-Pérot ridge laser. Note that the lasing frequency for this design ($\sim$ 2.35 THz) is considerably lower than for the FL-C series designs in section 3.3.1 ($\sim$ 3 THz).

creased ($E_{85'} \sim 27$ meV for the FL184P-M1 design of Fig. 3-38), since for the FL-C designs, the upper parasitic levels are pushed lower in energy as their energies get closer to the top of the barriers. However, the higher barriers in $GaAs/Al_{0.3}Ga_{0.7}As$ also have a disadvantage. Due to the reduced tunneling probability through such barriers, they have to be designed approximately twice as thin as compared to that for the $GaAs/Al_{0.15}Ga_{0.85}As$ designs, which makes the energy levels more prone to inhomogeneous linewidth broadening due to thickness variations during MBE growth.

Figure 3-38 shows the design details and experimental results for FL184P-M1, which was the second design to be tried in the FL-P series. The first design, named as FL184P, lased only in the NDR possibly because of processing problems since the etch-stop layer (which is discussed in chapter 4) was not grown by mistake for the FL184P wafer. A FL184P-M1 device lased up to $\sim$ 125 K in pulsed operation as shown in Fig. 3-38(c). However, this temperature performance cannot be directly compared to that of the FL-C series QCLs because the lasing frequency for FL184P-M1 is lower ($\sim$ 2.35 THz). At these low frequencies, the lower level parasitic current channels become more severe for the FL designs thus limiting their temperature performance, as will become apparent in section 3.5.1. Hence, it is more appropriate to compare the FL184P-M1 design to the FL175M-M3 design (to be discussed in section 3.5.1), which was implemented in the $GaAs/Al_{0.15}Ga_{0.85}As$ material system and lased at similar frequencies ($\sim$ 2.2 THz). Comparing the $T_0 \sim 195$ K obtained for the FL184P-M1 design (Fig. 3-38c), to the $T_0 \sim 210$ K for the FL175M-M3 design (Fig. 3-21), the temperature performance obtained for the FL-P design ($GaAs/Al_{0.3}Ga_{0.7}As$) does not seem to be significantly different than that for the FL-M designs ($GaAs/Al_{0.15}Ga_{0.85}As$).

SFL228

To push the upper parasitic levels higher up in energy, a thin barrier could be introduced in the wider of the two injector wells of the FL design, while other design parameters could be kept the same by appropriately adjusting the width of the other wells and barriers. Such a design has to be designed in a material system with a greater barrier height than that in $GaAs/Al_{0.15}Ga_{0.85}As$, since introduction of a thin barrier

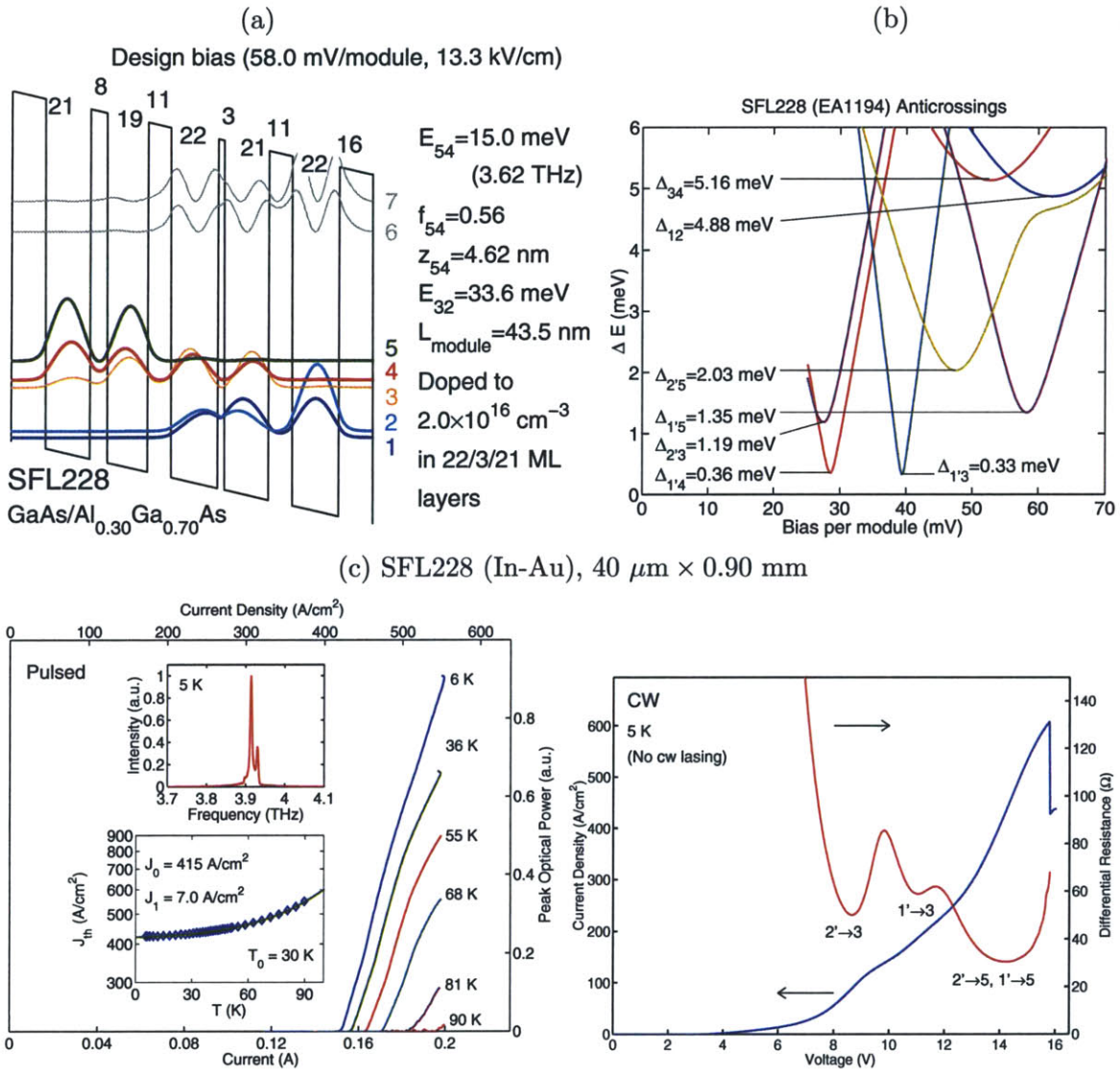


Figure 3-39: Design details and experimental results for SFL228. (a) One module conduction band diagram at the design bias. The radiative transition is from $5 \rightarrow 4$. This design is similar to FL184P-M1 in Fig. 3-38, except that the wide injector well is split into two by a 3 ML wide barrier, which pushes the upper parasitic levels higher up in energy. (b) Anticrossing plots for a two module calculation. (c) Pulsed and cw measurement results from a MM (In-Au) Fabry-Pérot ridge laser. No cw lasing was observed for this device, which is most likely due to a poor metal-metal (In-Au) bonding interface that can cause over heating during cw operation.

in the injector well pushes up all the energy levels in the module, which could not be afforded in the FL-C designs of section 3.3.1. SFL228, which is such a "split-well" FL design for the GaAs/$Al_{0.3}Ga_{0.7}As$ material system, is shown in Fig. 3-39. Due to the introduction of a thin (3 ML thick) barrier in the wider injector well, the upper parasitic levels 6 and 7 are pushed up much higher in energy ($E_{65} \sim 100$ meV). However, note that the radiative dipole matrix element in this design is considerably reduced ($z_{54} \sim 4.6$ nm) when compared to that in the FL-C designs ($z_{54} \sim 6$ nm), which directly translates into a smaller oscillator strength. This is because the quantum-wells in SFL228 have to be designed narrower than those in the FL-C designs such that the subbands are energetically located higher in the wells to allow for the introduction of a "split-off" barrier that is not very thin. Even the chosen value of 3 ML can cause significant inhomogeneous broadening of the injector levels 1 and 2 due thickness fluctuations of the order of a ML that can occur during the MBE growth from one module to another. For a yet thicker split-off barrier, the radiative oscillator strength will get reduced even further.

SFL228 did achieve lasing, as shown by the experimental results from a MM Fabry-Pérot ridge laser in Fig. 3-39(c). However, the temperature performance of this design does not suggest any improvement over the FL184P-M1 or the FL-C series designs. The thin split-off barrier, which determines the anticrossing state of the injector doublet consisting of levels 1 and 2, is the most likely cause of the worse performance in this design since it possibly reduces the specularity of the injection due to increased dephasing for the resonant-tunneling injector transport.

3.5 Low frequency terahertz QCLs

As was mentioned in section 1.1, the solid-state electronic sources of radiation are limited to operation below the frequency of approximately 1 THz. Due to several applications that exist for coherent sources in the frequency range of $1-2$ THz, almost immediately after the invention of terahertz QCLs, efforts were under way to extend their frequency of operation below 2 THz (wavelengths longer than 150 μm). However,

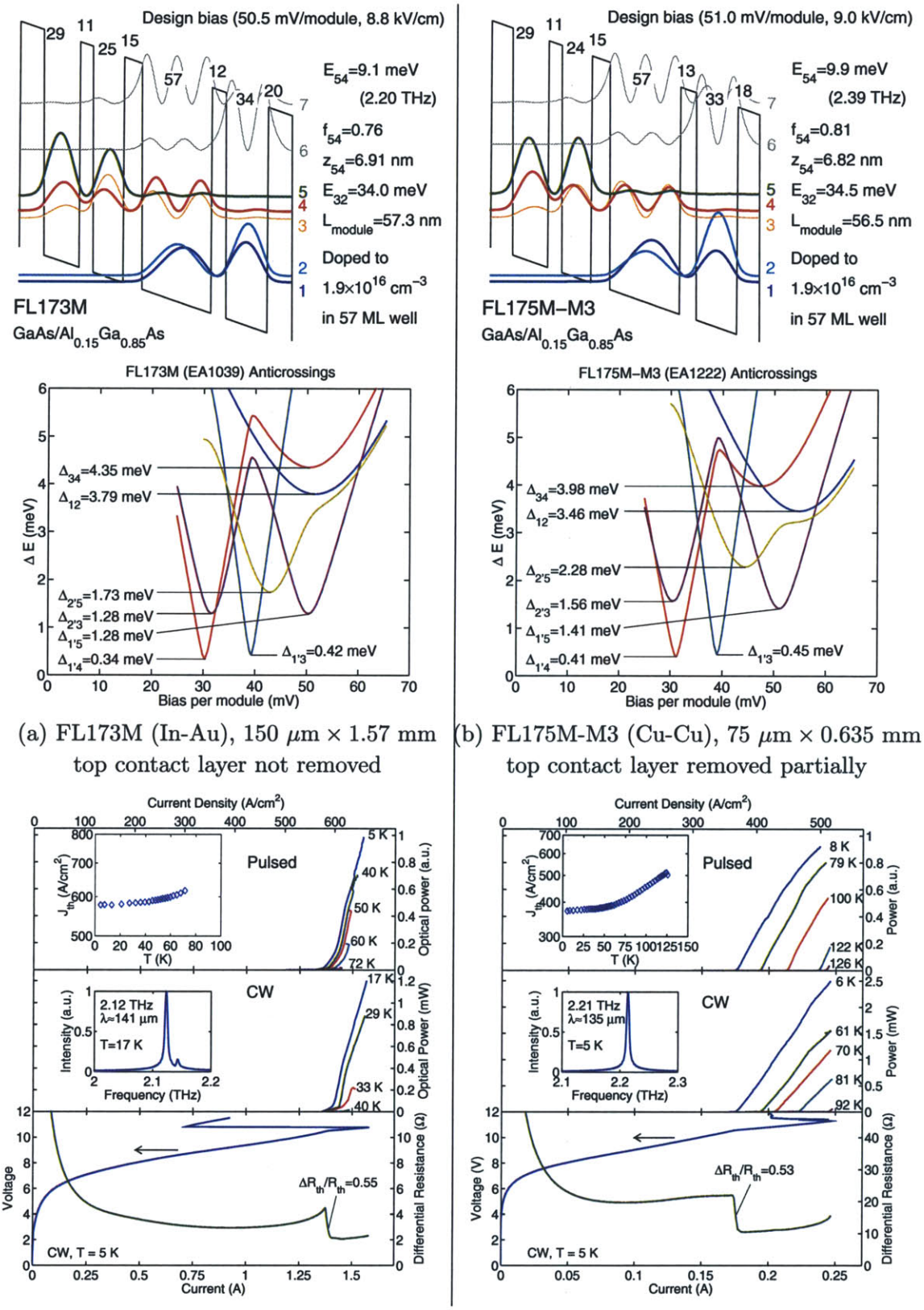
the design of QCLs for such low frequencies becomes increasingly challenging since the intersubband energy separations become extremely small, essentially of the order of the broadened linewidth of the subbands itself (1 THz corresponds to $\hbar\omega \sim 4$ meV). In that case, it becomes difficult to achieve the selective injection and removal of carriers necessary to obtain an intersubband population inversion. Furthermore, for subbands closely spaced in energy, intersubband optical absorption may significantly increase the optical losses inside the injector region. This makes the design of the injector region more challenging since most of the electrons reside in the injector, and for multiple level injectors, the subband spacings in the injector typically correspond to frequencies on the order of 1 THz. For waveguiding, the semi-insulating surface-plasmon waveguides become exceedingly inefficient in confining the optical mode for $\nu < 2$ THz [88]; however, to the contrary, the metal-metal waveguides only become more favorable (chapter 4). This section describes terahertz QCLs based on the FL design, which has a two-well injector region, operating at $\nu \sim 2$ THz. For operation below 2 THz, a different design based on a one-well injector region is described, whereby lasing down to $\nu \sim 1.6$ THz ($\lambda \sim 190$ μm) is demonstrated.

3.5.1 Low frequency FL designs: FL-M series

The radiative barrier in the FL design of Fig. 3-7 can be thickened to reduce the anticrossing energy between the radiative levels 5 and 4, and thus the radiative energy E_{54} at the design bias[1]. The radiative barrier for the FL-C designs in section 3.3.1 was ~ 9 ML thick, which yielded $E_{54} \sim 12$ meV ($\nu \sim 3$ THz). The FL-M series of designs are the low frequency analogs of the FL-C designs with a thicker radiative barrier. By thickening the radiative barrier to 11 ML, the radiative energy is reduced to a value of $E_{54} \sim 9 - 10$ meV ($\nu \sim 2.2 - 2.4$ THz).

The design details and experimental results from the first of the FL-M designs, named as FL173M, are shown in the left column of Fig. 3-40. This design achieved

[1]It may be noted that the width of the two wells in the active region also affects E_{54}, although for most of the FL designs reported in this chapter the width of the active regions wells has been kept the same.



(a) FL173M (In-Au), 150 μm × 1.57 mm top contact layer not removed

(b) FL175M-M3 (Cu-Cu), 75 μm × 0.635 mm top contact layer removed partially

Figure 3-40: Design details and experimental results for the FL173M (left column) and FL175M-M3 (right column) low-frequency terahertz QCLs. The substrate thickness was $\sim$ 350 μm for the FL173M laser, and $\sim$ 150 μm for the FL175M-M3 laser.

lasing at $\nu \sim 2.1$ THz [172], a low frequency record for terahertz QCL that stood for almost two years since its first demonstration, which is testimony to the challenges associated with the low frequency designs. However, the FL173M lasers were not robust. Out of many different devices tested, only a small fraction lased; the others had an early NDR with peak current densities in the range of $500 - 600$ A/cm^2. For the low-frequency FL designs, the $2' - 5$ injector anticrossing happens very close to the $1' - 3$ parasitic anticrossing in bias space. Consequently, the current flow in the structure close to the $\Delta_{1'3}$ bias becomes large due to simultaneous conduction by the $1' - 3$ and $2' - 5$ current channels, and it becomes more challenging to avoid the occurrence of an NDR before the required population inversion is achieved. The problem of the early NDR cannot simply be remedied by lowering the $1' - 3$ coupling. For the FL design that has a doublet of injector levels, the relative location of $\Delta_{2'5}$ and $\Delta_{1'5}$ with respect to $\Delta_{1'3}$ in the bias space must be carefully designed due to the closely spaced energy levels for a low-frequency design. Moreover, the design is made more challenging because there always exists a small uncertainty in the percentage of Al in the barrier material, which determines the height of the barriers and thus critically affects the bias location of various anticrossings with respect to each other.

For the lasing FL173M devices, the lasing range in current was small, as shown in Fig. 3-40(a) for a device that had a value of $J_{\mathrm{max,5K}} - J_{\mathrm{th,5K}} \sim 70$ A/cm^2, and thus a relatively small value of $T_{\mathrm{max,pul}} \sim 72$ K. A subsequent modification of this design, named as FL175M-M3, lased at $\nu \sim 2.2$ THz and obtained a much more robust temperature performance with $T_{\mathrm{max,pul}} \sim 128$ K, as shown in the right column of Fig. 3-40. Moreover, all the devices that were measured from the FL175M-M3 wafer lased robustly. The differences between the FL173M and the FL175M-M3 designs are rather marginal as deduced from their anticrossing plots Fig. 3-40. Hence, the principal cause for the radically improved performance of FL175M-M3 is not believed to be design related. For the devices processed from the FL175M-M3 wafer, the highly doped top-contact layer was partially etched for $\sim$ 10 sec in a $NH_4OH : H_2O_2 : H_2O$ 10:6:480 solution (etch rate of $200 - 300$ nm/min) with an objective of cleaning up the top surface after the removal of the etch-stop layer. This

wafer had a top contact layer thickness of 100 nm, and thus it is likely that a significant portion of that layer was removed since the etch rate can be faster in the initial stages of the etch. As was discussed in section 3.3.3, this could very well be the reason why a higher value of $J_{\mathrm{max,5K}} - J_{\mathrm{th,5K}} \sim 140$ A/cm^2, and therefore $T_{\mathrm{max,pul}} \sim 128$ K, was obtained for the FL175M-M3 lasers when compared to the FL173M lasers ($J_{\mathrm{max,5K}} - J_{\mathrm{th,5K}} \sim 70$ A/cm^2, $T_{\mathrm{max,pul}} \sim 72$ K).

Terahertz QCLs based on a FL design for $\nu < 2$ THz?

A modification of the FL173M design with a 12 ML radiative barrier ($E_{54} \sim 8$ meV, $\nu_{54} \sim 1.9$ THz), named as FL173N, failed to achieve lasing and the measured devices had I-Vs similar to that for the non-lasing FL173M devices with peak current densities in the range of $500 - 600$ A/cm^2. This might lead to a conclusion that FL designs may not operate at frequencies below 2 THz. However, the two following completely different design schemes, which are yet to be tried, might lead to operation at even lower frequencies. In the first scheme, a downwards-diagonal injector design can be tried such that the Δ_{12} anticrossing happens at a bias lower than the design bias. For some particular alignment, the $2' - 5$ coupling can be made smaller as compared to the $1' - 5$ coupling such that $\Delta_{2'5} \lesssim \Delta_{1'5}$. For such a design, the current should continue to increase beyond the $\Delta_{2'5}$ bias, and hence the occurrence of an early NDR can possibly be prevented. However, the values of the parasitic anticrossings $\Delta_{1'3}$ and $\Delta_{2'3}$ increase for a downwards diagonal design. This can be countered by thickening the collector barrier at the cost of a reduced the collector anticrossing Δ_{34}, which may not affect the performance of the laser severely, if lasing can indeed be obtained. In the second scheme, the intra-injector barrier could be made much thicker such that Δ_{12} is reduced down to a value as low as ~ 2 meV. The basis for such a design is discussed further in section 3.5.2. In this case the $2' - 5$ and $1' - 5$ anticrossings happen closer to each other in the bias space, and can be well separated from the $1' - 3$ parasitic anticrossing. For such a design, the collector barrier need not be made much thicker since a thicker intra-injector barrier also reduces the parasitic $1' - 3$ coupling. It may be noted that for designs with small collector or intra-injector anticrossings,

an accurate knowledge of the height of the potential barriers becomes important, since the barrier height sensitively affects the alignment of various anticrossings with respect to each other in bias space. Hence, the calibration of the fraction of Al in the $Al_xGa_{1-x}As$ barriers is of utmost importance for the low frequency QCL designs.

3.5.2 The one-well injector designs for $\nu < 2$ THz ($\lambda > 150$ μm) operation: OWI series

The FL designs, which have a two-well injector region, primarily face two challenges for operation at frequencies below 2 THz. The first problem is that of poor injection selectivity into the upper radiative level since the energy separation between the injector levels ($E_{12} \sim 4$ meV), and also the linewidth broadening of the individual levels ($\gtrsim 4$ meV) become similar to the radiative energy separation ($E_{54} \lesssim 8$ meV for $\nu < 2$ THz). A doublet of injector levels also causes an additional problem as explained in section 3.5.1, whereby the current flow below the design bias can get enhanced leading to an early NDR, thus preventing device operation at the design bias. The second potential problem is due to the increased optical losses in the structure at lower frequencies. This is caused by photon absorption in the injector region due to the $1 \rightarrow 2$ intersubband transition as the photon energy becomes comparable to E_{21}. Since the injector levels 1 and 2 are anticrossed at the design bias, the dipole matrix element z_{12} is large (for example $z_{12} \sim 7.5$ nm for FL175M-M3 in Fig. 3-40 as compared to the radiative value $z_{54} \sim 6.8$ nm). The problem of intersubband absorption becomes more severe as the number of levels increase in the injector. The chirped superlattice based designs (section 1.2.3) are one such example, for which QCL operation down to a frequency of 1.94 THz has been demonstrated [180].

For a two-well injector, the anticrossing energy Δ_{12} for the doublet of injector levels can be reduced to improve the injection selectivity and minimize the possibility of intersubband optical absorption from $1 \rightarrow 2$. This scheme has very recently been realized in a terahertz QCL that operated down to a frequency of 1.6 THz [161]. That design has a two-well injector region and a miniband based depopulation scheme. The

key enabling mechanism for that design is perceived to be the "tight" injector doublet such that $E_{12} \sim 1.6$ meV at the design bias. Such low values of the intra-injector (Δ_{12}) anticrossings have not yet been tried for the FL designs, a scheme that seems to hold good promise for realizing QCLs based on the FL design at frequencies below 2 THz as was discussed in section 3.5.1.

In this section, a terahertz QCL design based on a one-well injector is presented. Based on that design, a QCL operating at $\nu = 1.59$ THz ($\lambda = 188.5$ μm) is demonstrated. The following paragraphs discuss the operating principles for this design and the sequence of design iterations that led to this result.

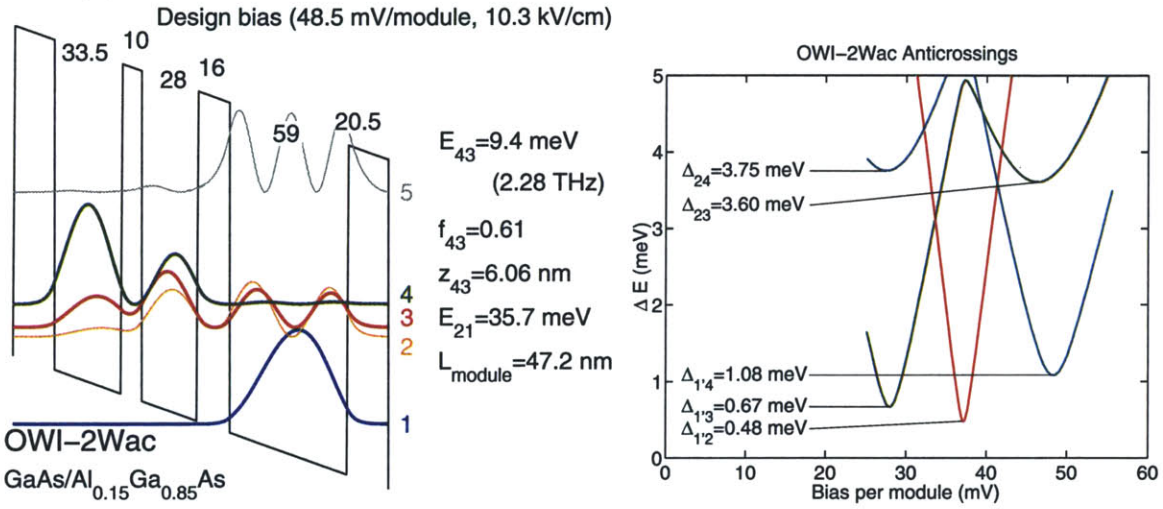
The intersubband dipole selection rule in QCLs (section 2.1.1, equation 2.25) requires the electric-field to be polarized perpendicular to the quantum-well layers (parallel to the growth direction) and hence precludes it from coupling to the in-plane electron motion *directly*. Since most of the carriers reside in the injector region, the lack of additional subbands in a one-well injector will prevent reabsorption of terahertz photons due to intra-injector intersubband transitions[1]. The one-well injector design has an additional advantage in that it provides the best injection selectivity that there could be since it has only one subband for injection.

The first of the modified[2] one-well injector (OWI) designs was aimed for operation at $\nu \sim 2.2$ THz. Fig. 3-41 shows two different designs that were considered, with two, and three wells in the active region, respectively, and both with a resonant-phonon depopulation scheme (section 2.5.2) similar to that in the FL designs (section 3.3.1). Both the designs in Fig. 3-41 are designed for similar values of the radiative energy E_{43}, the radiative oscillator strength f_{43}, the injector anticrossing $\Delta_{1'4}$, and the collector anticrossing Δ_{23}. With all other relevant parameters being equal, the design with a

[1]Note that this is contrast to very low frequency QCLs (operating down to $\nu = 1.39$ THz), which have been demonstrated by using strong magnetic fields applied perpendicular to the layers. A perpendicular magnetic field increases the modal gain by increasing the lifetime of the upper radiative state, and decreases the free carrier absorption in the structure by in-plane localization of carriers and subsequent quenching of intersubband scattering channels [138, 139].

[2]This section will focus mostly on the series of the one-well injector designs that were tried since June, 2005 onwards. A previous series of one-well injector designs were tried in April–June, 2003 but none of them had achieved lasing; the design details and experimental results from two of those early designs are shown in Fig. 3-42.

(a) A terahertz QCL design with a one-well injector, and a two-well active region



(b) A terahertz QCL design with a one-well injector, and a three-well active region

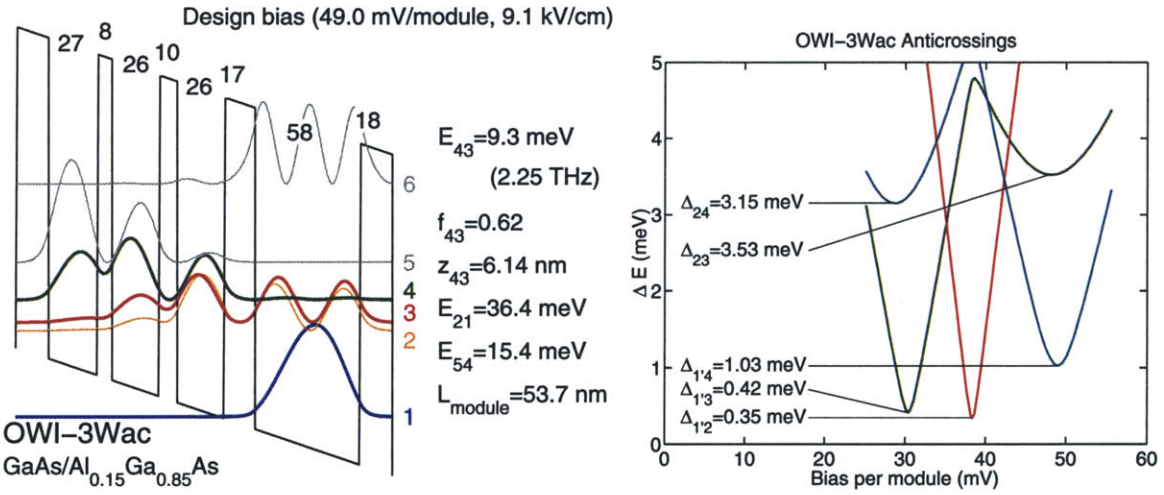
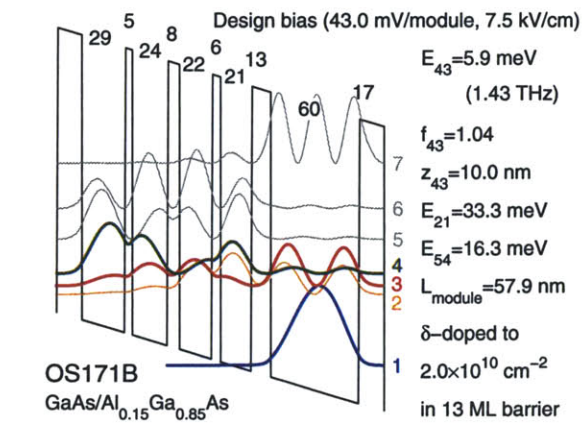
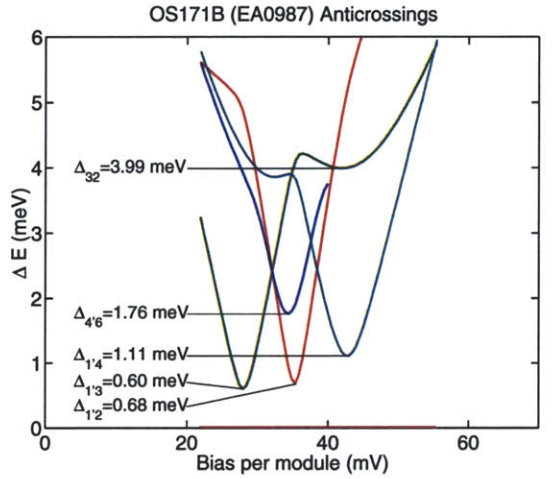
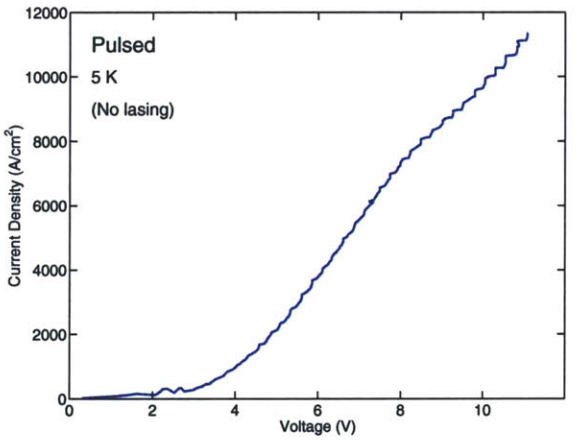


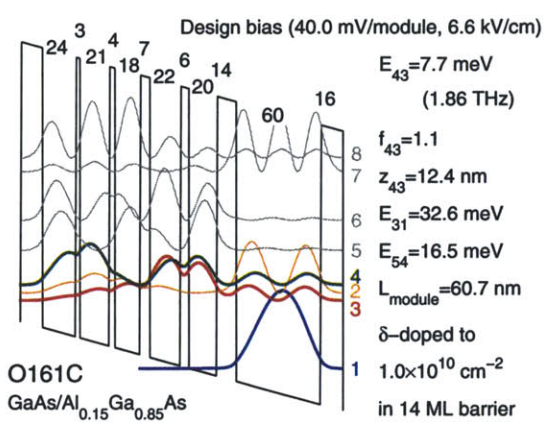
Figure 3-41: Terahertz QCL designs with a one-well injector, and (a) a two-well active region (OWI-2Wac), or (b) a three-well active region (OWI-3Wac), respectively. In both the designs the radiative transition is from $4 \rightarrow 3$. Level 3 is anticrossed with level 2, whereby it is depopulated by the resonant-phonon scheme (section 2.5.2). As compared to the doublet of levels for injection in the FL design (Fig. 3-7), the OWI design has only a single injector level (i. e. level 1). Both the designs in (a) and (b) are designed for similar values of the radiative energy E_{43}, the radiative oscillator strength f_{43}, the injector anticrossing $\Delta_{1'4}$, and the collector anticrossing Δ_{23}. That, however, results in different values for their lower level parasitic anticrossings $\Delta_{1'2}$ and $\Delta_{1'3}$.

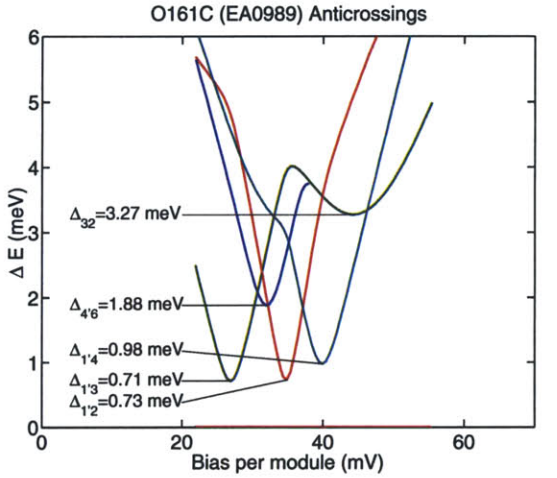




(a) OS171B (In-Au), 150 μm × 150 μm







(b) O161C (In-Au), 200 μm × 1.14 mm

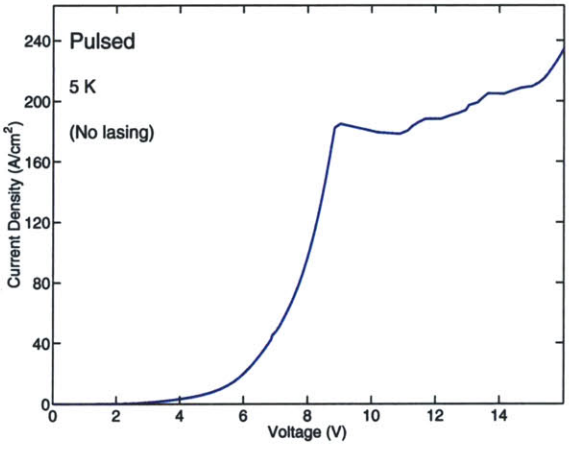


Figure 3-42: Design details and experimental results for OS171B and O161C. No lasing was observed from either of these designs. The unusually high current density for OS171B is most likely due to a doping error during the growth. It may be noted that both OS171B and O161C were designed much earlier (April–June, 2003) than all the other one-well injector designs discussed in this section (June, 2005 onwards). In the earlier stages the role of the parasitic current channels was not well understood. Hence, the $1' \rightarrow 2$ parasitic couplings in both of these designs are relatively large (as characterized by the anticrossing energy $\Delta_{1'2}$) as compared to the values typically desired ($\Delta_{1'2} \lesssim 0.4$ meV).

three-well active region (OWI-3Wac) has a considerably smaller value of the lower level parasitic anticrossing $\Delta_{1'2} = 0.35$ meV as compared to a value $\Delta_{1'2} = 0.48$ meV in the design with a two-well active region (OWI-2Wac). The $1' - 2$ coupling is the cause of the dominant parasitic current channel in this design (section 3.2.2). The current flow due to this parasitic current channel at the $1' - 2$ resonance is approximately $\propto \Delta_{1'2}^2$ for the typical values of the lower state lifetime ($\tau_{2,\mathrm{LO}} \sim 0.2$ ps) and the pure dephasing time ($T_2^* \sim 0.33$ ps [29]) as specified by equation (3.6) and plotted in Fig. 3-6. Consequently, OWI-3Wac was finally selected in preference to OWI-2ac. Note that the additional well in the active region of OWI-3ac introduces a higher energy level 5 above the upper radiative level 4 such that $E_{54} \sim 15$ meV. Even though the level 5 does not cause any additional parasitic current channels in the structure, it is likely to share some of the electron population with level 4 as the electrons become hot, thus reducing gain at higher temperatures. In theory, designs with more than three wells in the active region could further reduce the $1'-2$ parasitic coupling but those designs introduce additional undesirable higher energy levels in the active region, which are potential source of parasitic current channels and make the design more complex. Two of such (non-lasing) designs are shown in Figs. 3-42. The relatively large $4' \rightarrow 6$ coupling (of the type discussed in section 3.2.1) can be observed from the anticrossing plots, which happens due to the introduction of the upper parasitic level 6 in these designs.

An important parameter for a QCL design is the radiative oscillator strength (equation 2.34). The OWI-3Wac design in Fig. 3-41(b) was designed with a value of $f_{\mathrm{rad}} = 0.62$, which is much smaller than that for the FL-C designs (for example, $f_{\mathrm{rad}} = 0.92$ for the FL178C-M7 design in Fig. 3-10). This decision was based on the understanding that the gain could potentially be sacrificed in exchange of a better $\Delta_{1'4}/\Delta_{1'2}$ ratio, so as to lower the value of the parasitic current relative to the design bias current to avoid the problem of an early NDR. If OWI-3Wac is re-designed for a value of $f_{\mathrm{rad}} = 0.92$ while keeping all the other relevant parameters approximately the same, the value of the parasitic anticrossing $\Delta_{1'2}$ increases to 0.47 meV, as shown in Fig. 3-43. This will make the $1' \rightarrow 2$ parasitic current approximately a factor

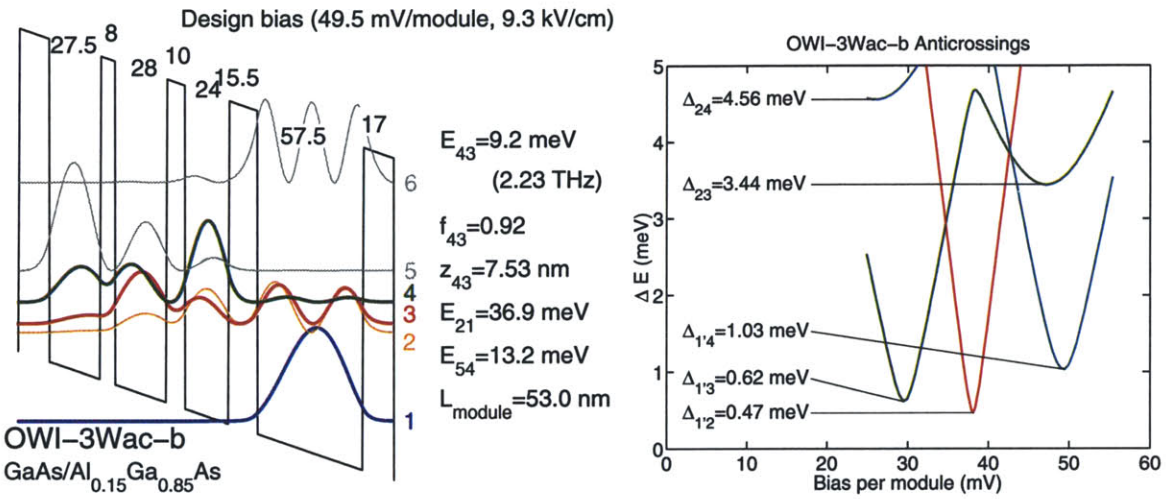


Figure 3-43: Modified version of the OWI-3Wac design (Fig. 3-41b) with a larger value of the radiative oscillator strength f_{43}. The values of the radiative energy E_{43}, the injector anticrossing $\Delta_{1'4}$, and the collector anticrossing Δ_{23} are kept similar to those for OWI-3Wac. A higher value of the oscillator strength causes the lower level parasitic anticrossings $\Delta_{1'2}$ and $\Delta_{1'3}$ to increase as compared to OWI-3Wac.

of $(0.47/0.35)^2 \sim 1.8$ larger as compared to that in the original OWI-3Wac design. The particular value of $f_{\text{rad}} = 0.62$ was chosen for the OWI-3Wac design to keep its radiative dipole matrix element $z_{\text{rad}} \sim 6.1$ nm similar to that for the FL-C designs (for example, $z_{\text{rad}} \sim 6.2$ nm for FL178C-M7). While there is no strict theoretical basis for this choice, the fact that the diagonality of a radiative transition is better represented by the dipole matrix element rather than the oscillator strength, the OWI-3Wac design should not have any larger inhomogeneous broadening of the gain linewidth as compared to that for the FL-C lasers despite having a considerably smaller f_{rad}. A qualitative reasoning for the dependence of the inhomogeneous broadening on the diagonality of the radiative transition is provided in section 3.1.

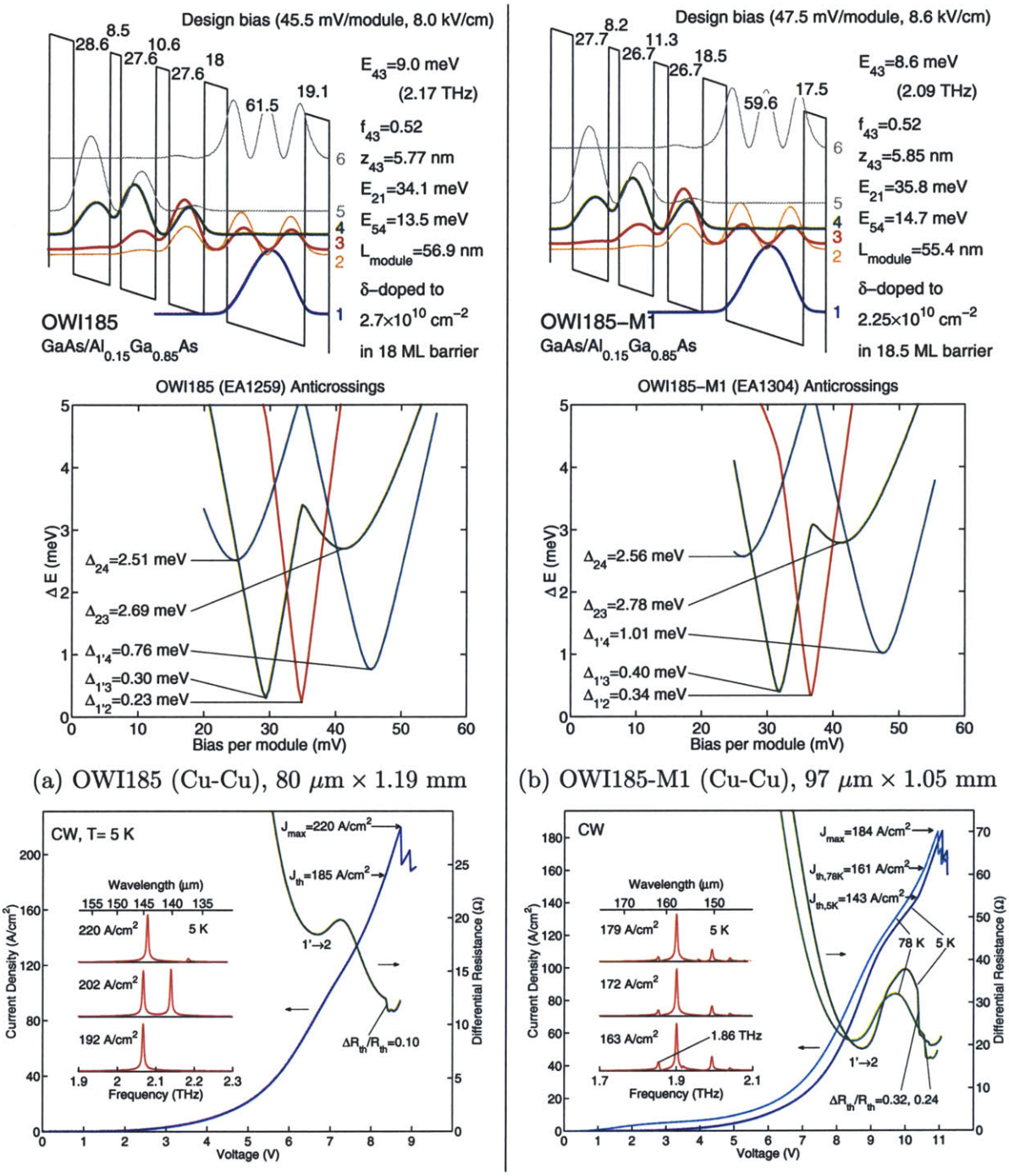
The OWI-3Wac design of Fig. 3-41(b) was the first one-well injector design that was grown and tested. The MBE grown wafer, named as OWI185, was overgrown in thickness by approximately 6% as compared to the design, which is a relatively large deviation for a terahertz QCL design. For the closely spaced energy levels in a terahertz QCL any variation in the energy level alignments, which depend critically on the width of the wells and the barriers, can significantly alter the design characteristics. Nevertheless, the OWI185 devices were able to achieve lasing ($\nu \sim 2.1$ THz). The

design details for the structure as it was grown, and the CW I-V, $\mathcal{R}$-V, and spectra from a MM (Cu-Cu) Fabry-Pérot ridge laser are shown in the left-column of Fig. 3-44. In spite of a relative small lasing range in current ($J_{\mathrm{max,5K}} - J_{\mathrm{th,5K}} \sim 35$ A/cm^2, $J_{\mathrm{max,5K}}/J_{\mathrm{th,5K}} \sim 1.19$) all the OWI185 devices that were tested made a laser.

The second design in the OWI-series was named as OWI185-M1, which was designed for a slightly lower frequency operation as compared to OWI185. Its design details, and the CW I-Vs, $\mathcal{R}$-Vs, and spectra from a MM (Cu-Cu) ridge laser are shown in the second column of Fig. 3-44. The OWI185-M1 lasers had a slightly better dynamic range in the lasing current ($J_{\mathrm{max,5K}} - J_{\mathrm{th,5K}} \sim 40$ A/cm^2, $J_{\mathrm{max,5K}}/J_{\mathrm{th,5K}} \sim 1.29$) as compared to that in OWI185, in spite of the fact that the operating frequency in OWI185-M1 was lower ($\nu \sim 1.9$ THz). The slightly better characteristics for the OWI185-M1 lasers are not clearly understood, since the anticrossing plots and design parameters for both the devices look similar. The worse performance in OWI185 could possibly be due to a non-uniform growth rate that is likely to have happened since its growth rate was significantly off-target. The lower *absolute* values of current densities in OWI185-M1 are most likely due to its smaller overall doping, which is approximately 85% of that in OWI185.

Pulsed and cw L-I characteristics measured from the same OWI185-M1 laser as in Fig. 3-44(b) are plotted in Fig. 3-45. For this device $J_{\mathrm{th,5K}} \sim 140$ A/cm^2 and $J_{\mathrm{max,5K}} \sim 180$ A/cm^2 in both pulsed and cw operation. Fabry-Pérot cavities with several different dimensions (varying widths and lengths) were tested and were found to have similar values of $J_{\mathrm{th,5K}}$. The insensitive dependence of $J_{\mathrm{th,5K}}$ on the cavity dimensions strongly indicates that $J_{\mathrm{th,5K}}$ is determined by the $1' \rightarrow 2$ parasitic current channel rather than the optical losses in the cavity. A relatively small J_{th}-T variation manifests this fact. Thus, notwithstanding a very small lasing range in current, a reasonably good value of $T_{\mathrm{max,pul}} \sim 110$ K was observed for this laser.

The fact that $J_{\mathrm{th,5K}}$ for the OWI185-M1 lasers is the lowest measured value from any resonant-phonon QCL attests to the low optical losses in the one-well injector designs. It is difficult to accurately determine the waveguide losses (α_{w}) for these lasers. However, an order of magnitude estimate can be obtained as follows. Due



(a) OWI185 (Cu-Cu), 80 μm × 1.19 mm

(b) OWI185-M1 (Cu-Cu), 97 μm × 1.05 mm

Figure 3-44: Design details and experimental results for the OWI185 (left column) and OWI185-M1 (right column) low-frequency terahertz QCLs. The 0.1 μm thick highly doped top-contact layer was partially etched for $\sim$ 15 sec in a $NH_4OH : H_2O_2 : H_2O$ 10:6:480 solution (etch-rate: $200 - 300$ nm/min), similar to that for the FL175M-M3 lasers (section 3.5.1).

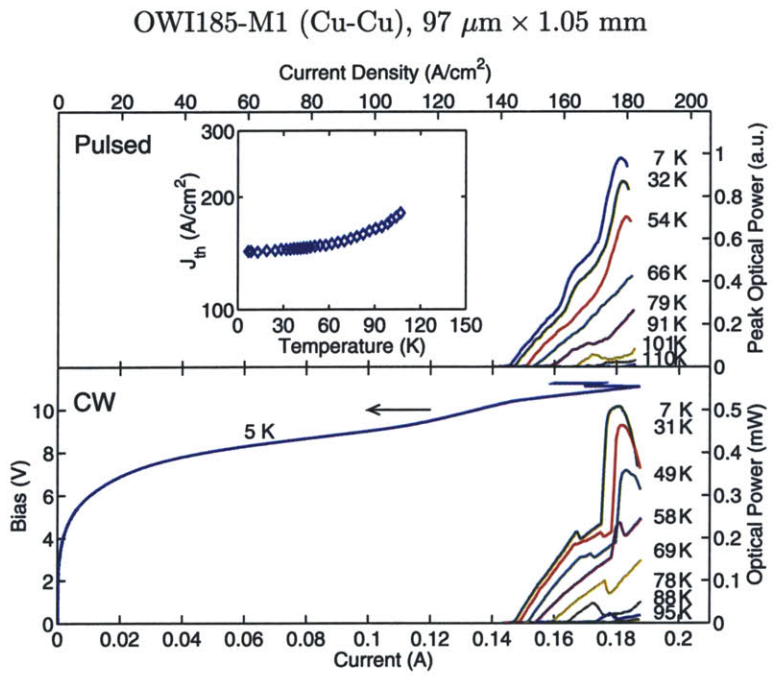
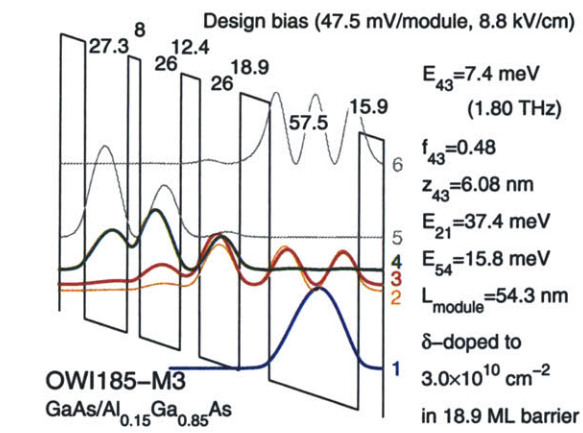
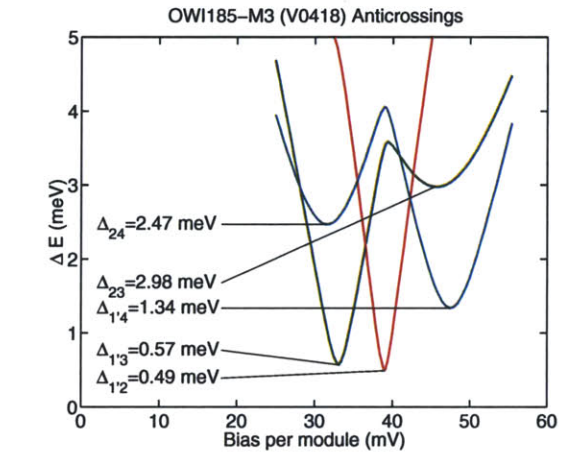


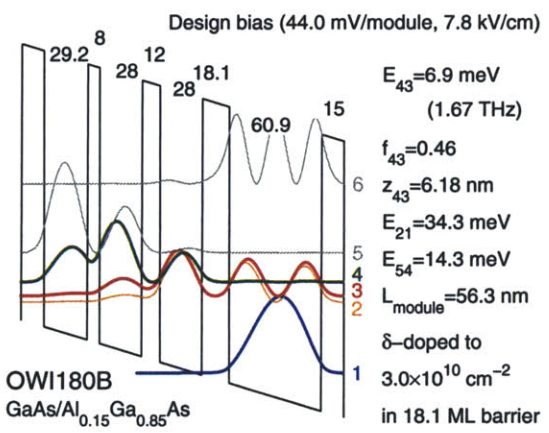
Figure 3-45: Pulsed and cw L-I characteristics from a OWI185-M1 laser operating at $\nu \sim 1.9$ THz. CW I-Vs and spectra for the same device are shown in Fig. 3-44(b).

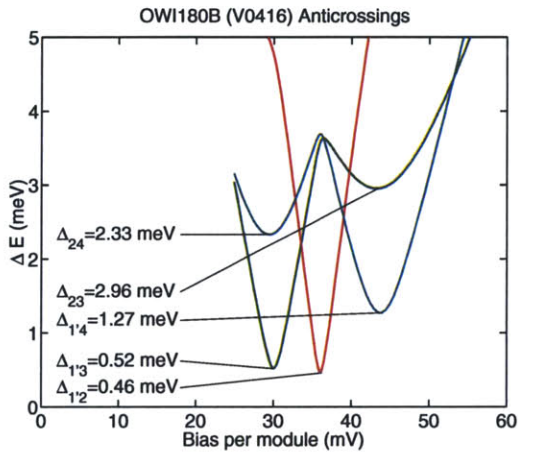
to sub-wavelength dimensions of the waveguide aperture in the growth direction, the mirror losses ($\alpha_{\rm m}$) in the Fabry-Pérot ridge waveguides are expected to be low [88]. A finite-element simulation for an infinite-width waveguide yielded $\alpha_{\rm m} \sim 1.3$ cm^{-1} for a 1 mm long cavity (both facets open), corresponding to a mirror reflectivity of $\sim 88\%$. Ridge lasers with cavity lengths of ~ 0.5 mm, 1 mm, and 2 mm, and the same width were tested, and all had a value of $J_{\rm max,5K} - J_{\rm th,5K} \sim 40$ A/cm^2. However, the longer devices in general had more cw output power. For a constant value of the lasing range in current densities, the peak power output from a ridge laser of length L is $\propto 1/(\alpha_{\rm w} + \alpha_{\rm m})$, where $\alpha_{\rm m} \propto 1/L$. The increase of cw power in going from a length of 0.5 mm to that of 1 mm was typically $50\% - 100\%$, whereas further increase in the length led to diminishing gains. These observations suggest that $\alpha_{\rm w}$ is of the order of $\alpha_{{\rm m},L=1\ {\rm mm}} \sim 1$ cm^{-1} for these lasers. A Drude model calculation yielded surface losses of ~ 5 cm^{-1} in the metal claddings, and $\lesssim 0.5$ cm^{-1} in the thin doped contact layers at the top and bottom of the active region. This suggests that losses in the active region itself are small for this design, as is expected from a one-well injector scheme.



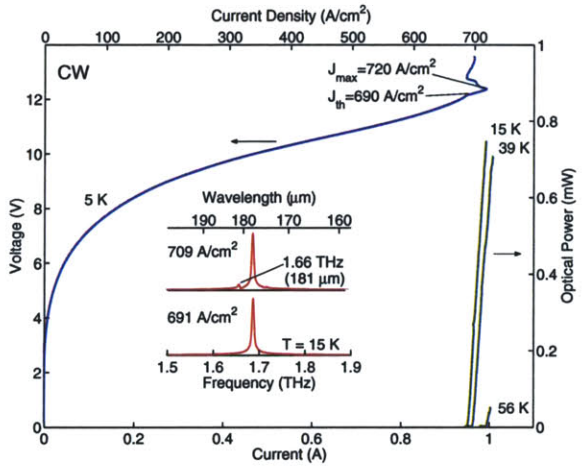


(a) OWI185-M3 (Cu-Cu), 100 μm × 1.38 mm





(b) OWI180B (Cu-Cu), 76 μm × 1.04 mm



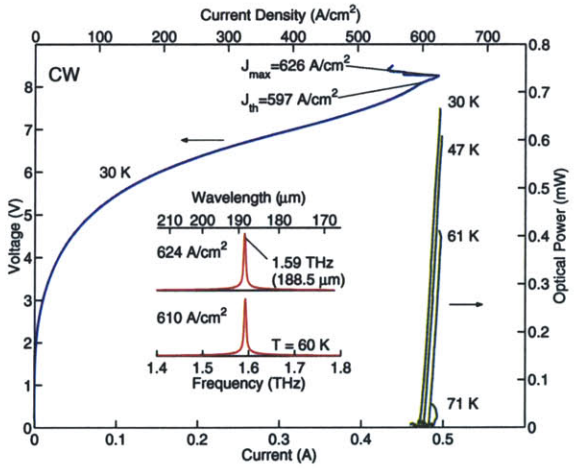


Figure 3-46: Design details and experimental results for the OWI185-M3 (left column) and OWI180B (right column) low-frequency terahertz QCLs. The 0.1 μm thick highly doped top-contact layer was completely removed for both these devices. The value of $\nu \sim 1.59$ THz for the OWI180B laser is presently the second lowest frequency for any solid-state laser that operates without the assistance of a magnetic field. Without any corrections in the detected power, the single-facet slope-efficiency for the OWI185-M3 laser is ~ 18 mW/A (~ 2.6 photons/electron), and for the OWI180B laser it is ~ 28.5 mW/A (~ 4.3 photons/electron).

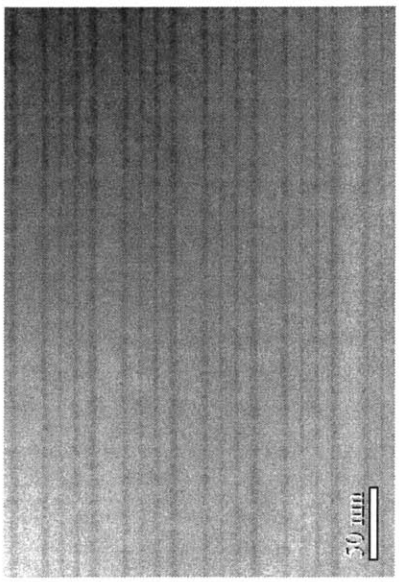


Figure 3-47: Scanning tunneling microscope (STEM) image of the OWI180B wafer showing top few modules of the active region. The dark layers correspond to $Al_{0.15}Ga_{0.85}As$. (Image courtesy of Dr. Z. R. Wasilewski and Dr. X. Wu, National Research Council, Canada.)

The ratio of an injector anticrossing to a low-bias parasitic anticrossing (for example, $\Delta_{1'4}/\Delta_{1'2}$ in the OWI-3Wac design) decreases as the radiative frequency is lowered in a particular QCL design. As the applied bias is increased in a QCL structure, the subbands are located at higher energies relative to the bottom of the quantum-wells, and hence they couple more strongly to the subbands in the adjacent wells. For a high frequency design, the bias at which the lower level parasitic anticrossings happen is considerably smaller compared to the design bias, which confines the subbands deeper in the quantum-wells at the parasitic bias. Therefore, the interwell coupling of the subbands is weakened at the parasitic bias as compared to that at the design bias. This is the primary reason for the absence of the low-bias parasitic current channels in mid-infrared QCLs. However, these parasitic channels are an inherent feature of the resonant-phonon terahertz QCLs that artificially enhance the threshold current densities at low temperatures, as has been discussed throughout this chapter. The enhancement of these low-bias parasitic current channels in the lower frequency designs becomes the foremost impediment to obtain lasing at low frequencies.

Figure 3-46 shows the design details and experimental results for OWI185-M3

and OWI180B, two subsequent iterations of the OWI-3Wac design for even lower-frequency operation. As compared to OWI185-M1, the radiative oscillator strengths in both these designs were lowered; however, in spite of that, their injector anticrossing to the parasitic anticrossing ratios ($\Delta_{1'4}/\Delta_{1'2}$) were further reduced due to the reasons explained in the previous paragraph. This is manifested into an extremely small lasing range in current ($J_{\mathrm{max,5K}} - J_{\mathrm{th,5K}} \sim 30$ A/cm^2) for both these lasers. Nevertheless, it is worthwhile to note that these wafers achieved lasing in their first attempt. Both of these wafers were grown in a different MBE machine than the one that was used to grow the rest of the devices discussed in this chapter. Hence, these designs will possibly need to be readjusted for an optimum performance, for differences as small as 0.5% in the Al concentration of the $Al_xGa_{1-x}As$ barriers can change the design characteristics significantly. The operation of the OWI180B QCLs at a frequency of 1.59 THz ($\lambda = 188.5$ μm) is presently the second lowest reported value for any solid-state laser that operates without the assistance of a magnetic field (the lowest frequency of operation has very recently been obtained by a different research group, although the exact details are currently unavailable). Taking cue from these results, it is reasonable to predict a more robust operation at even lower frequencies in the subsequent design iterations of the OWI-3Wac design.

3.6 High power terahertz QCLs

For some applications such as providing a local oscillator signal for a cryogenic heterodyne receiver, optical power levels of few milli-Watts are often sufficient. Whereas for other applications that involve significant material penetration, propagation through atmosphere, or illumination of a large area while still maintaining a good local signal-to-noise ratio, such as in imaging or some of the spectroscopic applications, large amount of optical power is the topmost priority. For terahertz QCLs, the best temperature performance has been obtained in the metal-metal (MM) waveguides [170], the properties of which will be discussed in greater detail in chapter 4. Most of the experimental results presented up to this section were therefore from MM QCLs. However,

terahertz QCLs with the highest power have been developed in the semi-insulating surface-plasmon (SISP) waveguides [89] that were discussed briefly in section 1.2.3. This section first describes the key differences between the two types of the waveguides, and then focuses mainly on the experimental results. Detailed electromagnetic modeling and analyses of the two types of waveguides can be found in Ref. [88].

3.6.1 Metal-metal waveguides versus semi-insulating surface-plasmon waveguides

The following expressions can be written for a laser operating above threshold:

$$\Gamma g_{\mathrm{mat}}(\omega) = \alpha_{\mathrm{w}} + \alpha_{\mathrm{m}} \tag{3.10}$$

$$\frac{dP_{\mathrm{out}}}{dI} \propto \hbar\omega \frac{\alpha_{\mathrm{m}}}{\alpha_{\mathrm{w}} + \alpha_{\mathrm{m}}} \tag{3.11}$$

Equation (3.10) is the well known condition that the gain must be equal to the losses during laser operation. This condition was previously derived using QCL rate equations in section 2.3.2. In this expression, Γ is the fraction of the mode that propagates in the active medium (also known as the mode confinement factor) and g_{mat} is the *material* gain of the active medium as was derived in section 2.1.2 (equation 2.32) and also in section 2.2 (equation 2.45), therefore, $\Gamma g_{\mathrm{mat}}(\omega)$ becomes the *modal* gain. Also, α_{w} is the modal propagation loss in the waveguide, and α_{m} is the out-coupling (mirror) loss in the resonator cavity. For a current I flowing in the device, the slope efficiency dP_{out}/dI can be written as in equation (3.11) following the derivation in section 2.3.3 (equation 2.62), where P_{out} is the power out-coupled from the cavity and $\hbar\omega$ is the photon energy for the lasing mode. In a Fabry-Pérot ridge cavity of length L, the expression for the slope efficiency can be rewritten as

$$\frac{dP_{\mathrm{out}}}{dI} \propto \hbar\omega \frac{\alpha_{\mathrm{m,f}}}{\alpha_{\mathrm{w}} + \alpha_{\mathrm{m,f}} + \alpha_{\mathrm{m,r}}} \tag{3.12}$$

$$\alpha_{\mathrm{m,f/r}} = \frac{1}{2L} \log_{\mathrm{e}}(R_{\mathrm{f/r}}) \tag{3.13}$$

where $R_{f/r}$ is reflectivity of mode from the front/rear facet of the cavity, $\alpha_{m,f/r}$ is the mirror loss due to the corresponding facet, and the output power P_{out} is assumed to be collected from the front facet only.

To bring about the key differences between MM and SISP waveguides, the mode shapes and the propagating mode parameters calculated at two different frequencies for typical waveguide geometries are shown in Fig. 3-48. As can be noticed, the MM waveguides confine the mode in sub-wavelength dimensions in the vertical direction, by virtue of a *double-plasmon* mode that is bound to both the bottom and the top metal cladding. This leads to a large mismatch in the mode shapes of a mode that can propagate in free-space and the mode inside the waveguide. Consequently, the mirror reflectivity in MM waveguides becomes very high ($R \sim 0.7 - 0.9$ [88]), as opposed to that expected for plane wave reflection at a normal air/GaAs boundary ($R_{norm} = \left(\frac{n_{GaAs}-1}{n_{GaAs}+1}\right)^2 \sim 0.32$). In contrast, even though the mode in a SISP waveguide is also a double-plasmon mode, which is bound to the top metal and the thin n^+ GaAs layer at the bottom, it is not confined within sub-wavelength dimensions. Hence, the mirror reflectivity in SISP waveguides is approximately the same as $R_{norm} \sim 0.32$ [88].

As can be noticed from the calculated values in Fig. 3-48, MM waveguides have a smaller value of the waveguide losses α_w/Γ in comparison to the SISP waveguides. It may be noted that the semi-insulating GaAs substrate is likely to contribute additional losses for SISP waveguides that have not been included in Fig. 3-48. These losses can occur due to impurity absorption at terahertz frequencies caused by shallow donors and acceptors in the nominally semi-insulating substrate. Also, $\alpha_m/\Gamma \sim 0.5 - 2$ cm^{-1} per facet for a 1 mm long cavity in MM waveguides (calculated for $R \sim 0.9 - 0.7$), as compared to $\alpha_m/\Gamma \sim 15 - 30$ cm^{-1} per facet for a 1 mm long cavity in SISP waveguides (calculated for $\Gamma \sim 0.4 - 0.2$). Hence, even for relatively long cavity lengths, the mirror losses in SISP waveguides are an order of magnitude greater than those in MM waveguides.

The aforementioned arguments indubitably suggest that the overall losses in SISP waveguides are significantly higher as compared to those in MM waveguides. Hence, QCLs realized in MM waveguides show much improved temperature performance.

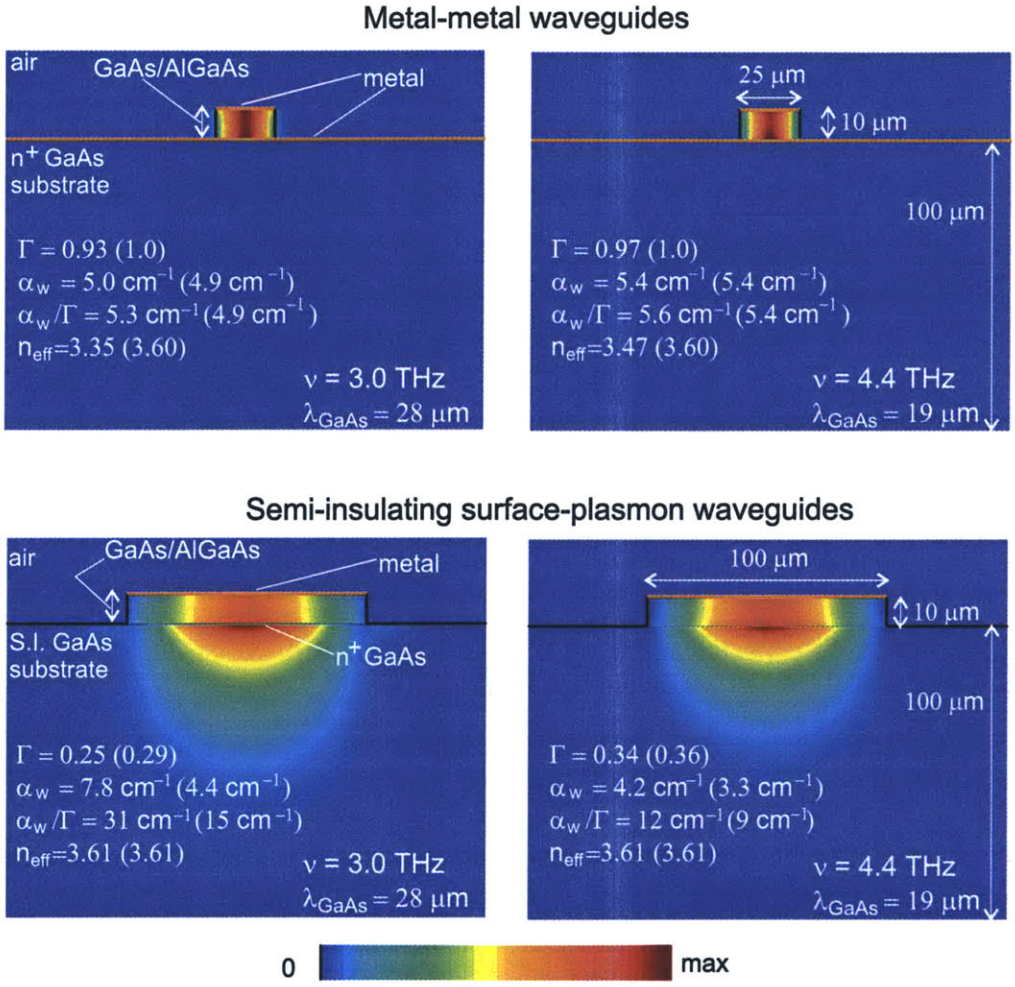


Figure 3-48: Two-dimensional electromagnetic mode calculations for the fundamental lateral modes in the MM and SISP waveguides (schematic shown in Fig. 2-3). The calculations are done within a Drude model [16] at $\nu = 3.0$ THz, and $\nu = 4.4$ THz, respectively. For SISP waveguides, the 3 THz design (FL178C-M10) has a 0.7 μm thick n^+ GaAs layer doped to 2×10^{18} cm^{-3}, whereas the 4.4 THz design (FL183R) has a 0.4 μm thick n^+ GaAs layer doped to 3×10^{18} cm^{-3}. The total energy density across the crosssection of the waveguide for a mode propagating perpendicular to the plane of the figure is plotted. A Drude scattering time of $\tau = 50$ fs and a doping density of 5.9×10^{22} cm^{-3} is used for metal (Au) [4], and a $\tau = 0.1$ ps is used for the n^+ GaAs layer [170, 88]. The refractive index of GaAs is taken to be $n_{\text{GaAs}} = 3.6$. n_{eff} is the effective index of the propagating mode. The values in parentheses are the results for infinitely wide waveguides. The highly-doped top contact layer is not included in these calculations. For these calculations the active region is taken to be lossless, and so is the semi-insulating GaAs substrate for the SISP waveguides. These calculations are performed using a finite-element solver [2].

This will be shown distinctly through experimental results in section 3.6.2 by comparing results from QCLs with the same active region and implemented in both MM and SISP waveguides. In contrast, the factor $\frac{\alpha_{m,f}}{\alpha_w+\alpha_{m,f}+\alpha_{m,r}} \lesssim 0.1$ for MM waveguides as compared to the values of $\frac{\alpha_{m,f}}{\alpha_w+\alpha_{m,f}+\alpha_{m,r}} > 0.2 - 0.3$ that are possible for SISP waveguides. This suggests that the slope efficiency for MM QCLs is expected to be considerably smaller than that for SISP QCLs. This has indeed been observed experimentally. The highest measured values of the differential quantum efficiency ($\equiv \frac{|e|}{\hbar\omega}\frac{dP_{out}}{dI}$, section 2.3.3) for the resonant-phonon terahertz QCLs have been ~ 4.5 photons/electron for MM waveguides (Figs. 3-17 and 3-46) and ~ 16.5 photons/electron for SISP waveguides (Fig. 3-50) calculated for power as measured at the detector. Hence, notwithstanding a smaller lasing range in current for the SISP waveguides due to a higher value of J_{th}, significantly greater amount of power (> 100 mW) can be obtained from the SISP lasers as compared to that obtained from the MM lasers ($\lesssim 10$ mW, without the use of an output coupling lens), as will be shown next.

3.6.2 High frequency FL designs for high power output: FL-R series

It serves well to design the resonant-phonon terahertz QCLs for operation at relatively higher frequencies ($\nu \sim 4-5$ THz) to obtain the highest power output. This is because of the following reasons:

- A higher frequency design has smaller low-bias parasitic couplings for reasons explained in section 3.5.2. Consequently, $J_{th,5K}$ is lower, and hence the lasing range in current $J_{max,5K} - J_{th,5K}$ is higher. This directly leads to higher optical power output for a given slope efficiency[1].

- The SISP waveguides are able to confine a higher frequency mode better. Consequently, a thinner bottom n^+ GaAs layer can be used in the waveguide design.

[1]Note that a higher value of $J_{max,5K} - J_{th,5K}$ does not imply higher $T_{max,pul}$, since J_{th} increases more rapidly with temperature for QCLs operating at higher frequencies (section 3.3.4, Fig. 3-21).

Therefore, α_w is reduced in addition to an already high value of Γ, thereby significantly lowering the overall loss α_w/Γ in a higher frequency design (Fig. 3-48).

- The increased slope efficiency due to higher photon energy ($\frac{dP_{out}}{dI} \propto \hbar\omega$, equation 3.12) in a high frequency laser also increases the output power. Moreover, the voltage efficiency (section 2.3.3) is also improved in such a design, which causes a more efficient electrical to optical power conversion (i. e. a better wall-plug efficiency of the laser).

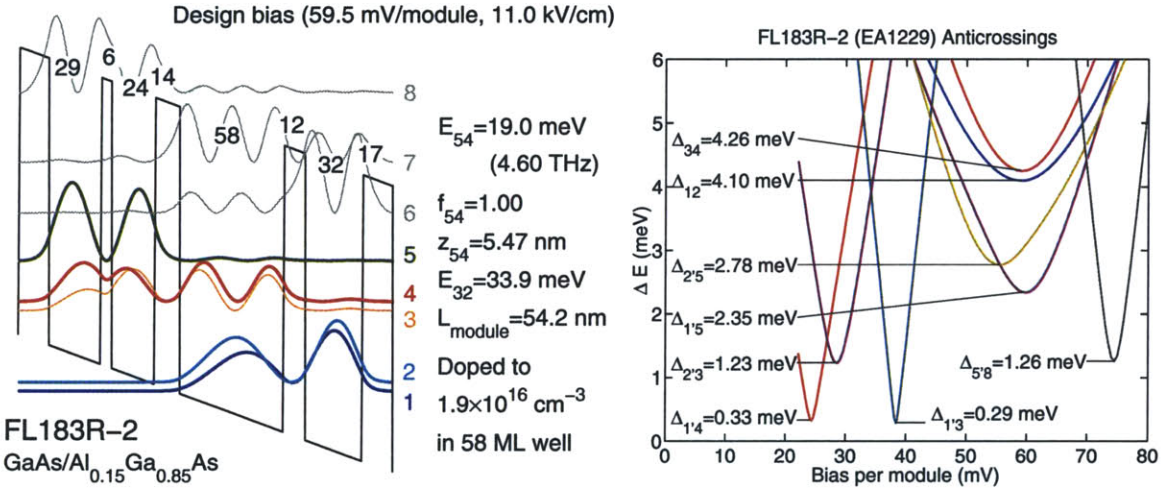


Figure 3-49: One module conduction band diagram at the design bias (left) and anticrossing plots for a two-module simulation (right) for FL183R-2, a high frequency FL design (the FL design is described in section 3.3.1).

The FL design of section 3.3.1 was modified for higher frequency ($\nu \sim 4.5$ THz) operation and the resulting design named as FL183R-2, the design details of which are shown in Fig. 3-49. The primary change in this design as compared to the FL-C series designs is a thinner (6 ML) radiative barrier, which increases the radiative anticrossing and hence the radiative energy E_{54} at the design bias. This wafer was processed into semi-insulating surface-plasmon waveguide ridge structures with a standard recipe [166], with the exception that the ridges were dry etched like for the metal-metal waveguides (chapter 4) and the lateral contacts were non-alloyed (Ti/Au 20/200 nm) without any extra thermal annealing step. The devices were lapped to a substrate thickness of ~ 170 μm to improve heat-sinking, ridges were cleaved, and an

$Al_2O_3/Ti/Au/Al_2O_3$ high-reflectivity (HR) coating was evaporated on the rear facet of a few selected devices. The HR coating was done in two steps to make the devices robust under cw operation, during which such a coating is often prone to failure due to electrical shorting. In the first step, only a part of the ridge near the cleaved facet was covered with Al_2O_3 by a uniform deposition (300 nm thick) with a shadow mask. In the second step, the wafer was turned over and a directional deposition was done at a slanted angle to coat the facets only with the $Al_2O_3/Ti/Au/Al_2O_3$ sequence (layer thicknesses: 200/20/200/100 nm).

FL183R-2 (SISP), 198 μm $\times$ 1.21 mm, substrate thickness $\sim$ 170 μm
rear-facet HR coated, dry-etched ridges

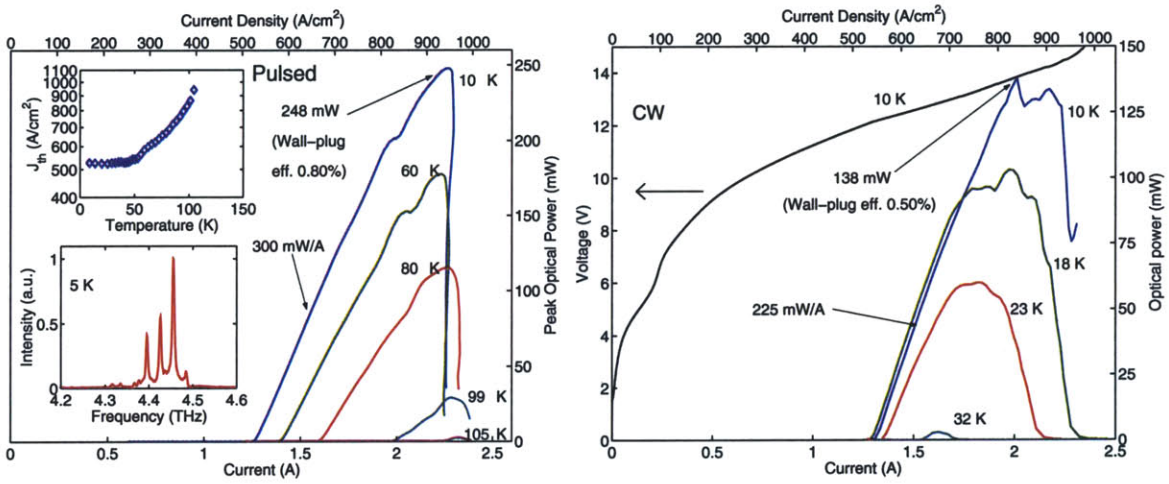


Figure 3-50: Experimental results from a FL183R-2 SISP waveguide laser that emitted the high reported optical power for any terahertz QCL in both pulsed and cw operation. In pulsed operation, the device was biased with pulse trains of 200 ns pulses repeated at 100 kHz, modulated by a 1 kHz square wave for an overall duty cycle of 1%, and the *L-I*s are measured with a room-temperature pyroelectric detector (Molectron model P4-42). The peak optical power values are as measured at a calibrated thermopile detector (ScienTech model AC2500H) without correcting for the collection efficiency (estimate unavailable) or the transmission losses ($\sim 85-90\%$) of a polypropylene dewar window [166]. The single-facet slope efficiency of 300 mW/A corresponds to a value of $\sim$ 16.5 photons/electron.

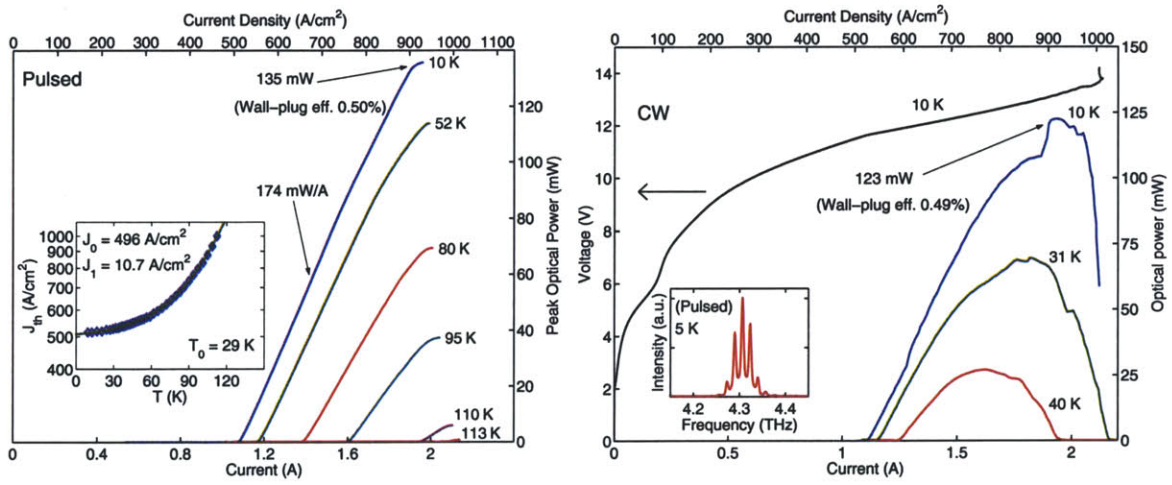
Pulsed and cw measurement results from a FL183R-2 Fabry-Pérot SISP ridge laser with the rear-facet HR coated are shown in Fig. 3-50. The cavity length for this device was kept relatively short (1.21 mm) to obtain a higher mirror loss ($\alpha_{m,f} \sim 4.7$ cm^{-1}) and hence a higher slope efficiency (equation 3.12). By choosing a relatively wide

(198 μm) waveguide, the confinement of the mode is improved and the waveguide losses are reduced for the SISP waveguides (Fig. 3-48). However, note that even though the peak power increases monotonously for longer cavities, the increase in peak power is diminishing as the cavities become longer since the slope efficiency decreases monotonously with cavity length. For the laser in Fig. 3-50 the cavity length was chosen keeping in consideration the maximum electrical current that could be supplied by the pulsed power supply ($I_{\max} \sim 2.5$ A at ~ 12 V using a 50 Ω to 12.5 Ω impedance transformer with an Avtech Electrosystems model AV-1011-B power supply). The values for both the slope efficiency and the peak optical power obtained for this device are presently the highest such reported values for any terahertz QCL in both pulsed and cw mode of operation.

A narrower cavity laser has better heat removal in the lateral (width) direction, and consequently its temperature performance in cw operation is improved. As compared to the device in Fig. 3-50, a somewhat more efficient cw performance was obtained from a narrower width (98 μm) device as shown in Fig. 3-51(a). For a narrower waveguide, the mode confinement factor Γ also reduces (Fig. 3-48), therefore, the device in Fig. 3-51(a) had to be made almost twice as long as the one in Fig. 3-50 to maintain a similar lasing range in current. The longer cavity device had a slightly better cw performance ($T_{\max,\mathrm{cw}} \sim 40$ K), which is significant for practical applications. Using a closed-cycle pulsed tube cryorefrigerator with a few watts of cooling power at 30 K (Cryomech PT60), a peak power of 50 mW was obtained in quasi-cw mode operation ($\sim 25\%$ duty-cycle) for this laser at ~ 33 K. That was sufficient to perform real-time terahertz imaging using a microbolometer focal-plane array as the detector [101], the first such reported imaging demonstration for any terahertz solid-state laser. This device has been cumulatively operated for more than 100 hours in the cryorefrigerator with over 25 thermal cycles without showing any signs of degradation.

An estimate for the waveguide loss α_{w} in the 98 μm wide FL183R-2 laser can be obtained by comparing the slope efficiency of the laser in Fig. 3-51(a), which has HR coating on its rear-facet, to another laser of similar dimensions but without any HR coating on its rear-facet. The measurement results from the latter device are shown

(a) FL183R-2 (SISP), 98 μm $\times$ 2.15 mm, substrate thickness $\sim$ 150 μm
rear-facet HR coated, dry-etched ridges



(b) FL183R-2 (SISP), 98 μm $\times$ 2.00 mm, substrate thickness $\sim$ 150 μm
no HR coating on the rear-facet, dry-etched ridges

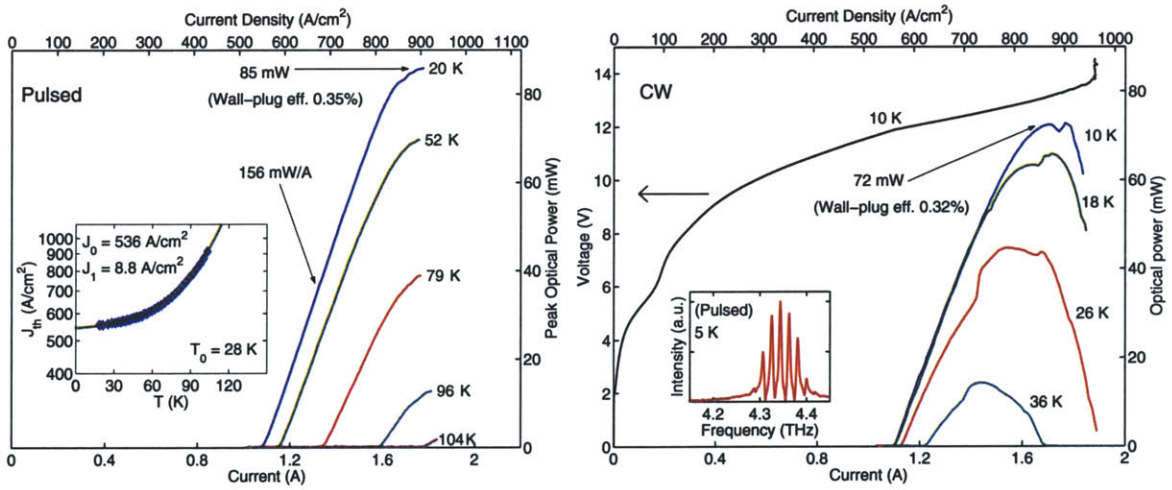


Figure 3-51: Experimental results from two high-power FL183R-2 SISP waveguide lasers, (a) with, and (b) without an HR coated rear-facet, respectively. Both these devices have similar dimensions, are from adjacent locations on the wafer, and were processed as part of the same die. These measurement results can be compared to determine the effect of HR coating the rear-facet on the operation of the laser. The laser with the HR coated rear-facet has been operated in quasi-cw mode ($\sim$ 25% duty-cycle) at $\sim$ 33 K in a cryorefrigerator, whereby a peak power of 50 mW was obtained that was sufficient to perform real-time terahertz imaging with a microbolometer focal-plane array detector [101].

in Fig. 3-51(b). Since both these devices were from the same die, the fluctuation of α_w from device to device is expected to be small. Unlike the MM QCLs, the optical power measurements from the SISP QCLs are more consistent and repeatable from device to device, since, for these waveguides, just a single type mode is excited in the direction along the width, and also their radiation pattern is much more directional as compared to that from the MM waveguides [73, 9]. Hence, the output power variation from device to device can be primarily attributed to the characteristics of the device itself, rather than fluctuations in the fraction of power collected at the detector. Using equation (3.12), a value of $\alpha_w \sim 12.7$ cm^{-1} is obtained for the lasers in Fig. 3-51, which is a factor of ~ 3 higher than that obtained from the Drude model calculation in Fig. 3-48. The extra loss is most likely due to losses in the semi-insulating substrate as well as in the active region, which were both assumed to be lossless in the calculation of Fig. 3-48; although an underestimation in the losses due to the top metal and/or the highly doped "plasmon" layer at the bottom cannot be entirely discounted. For a 2 mm long cavity, $\alpha_m \sim 2.2$ cm^{-1} per facet. For a value of $\Gamma = 0.34$ (Fig. 3-48), the threshold gain ($g_{th} = \frac{\alpha_w + \alpha_{m,f} + \alpha_{m,r}}{\Gamma}$) is estimated to be approximately 45 cm^{-1} and 54 cm^{-1} for the lasers in Figs. 3-51(a) and 3-51(b), respectively. In spite of a relatively large difference in the threshold gains, the values of $J_{th,5K}$ (510 A/cm^2, 550 A/cm^2) and $T_{max,pul}$ (105 K, 113 K) for these two lasers are not significantly different from each other[1]. One possible explanation for this behavior is the fact that J_{th} increases exponentially with temperature for the FL-series lasers (section 3.3.4), and hence, even though the difference in the overall losses for the two lasers is seemingly large, a rather small difference in their temperature performance is observed. It is also likely that the ratio of the threshold gains for the two devices in Fig. 3-51 approaches unity at higher temperatures ($\sim$ 100 K) since α_w should increase with temperature. It may also be noted that the low-temperature J_{th} for the FL designs is predominantly determined by the low-bias parasitic current channels,

[1] J_{max} for the laser in Fig. 3-51(a) is $\sim$ 60 A/cm^2 higher than that in Fig. 3-51(b) at high temperatures, which contributes partially to the higher $T_{max,pul}$ for the former. If their J_{max} values were to remain the same, the difference in the $T_{max,pul}$ values would have been even smaller. Note that such fluctuations in J_{max} from device to device are commonly observed, although their cause is not well understood.

FL183R-2 (Cu-Cu), 80 μm $\times$ 1.94 mm, wet-etched ridges
highly-doped top contact layer completely removed

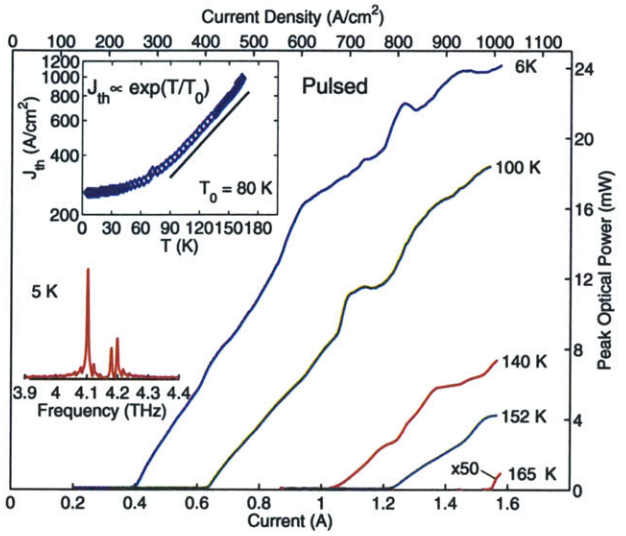


Figure 3-52: Experimental results from a FL183R-2 MM waveguide laser operating in pulsed mode. The optical power output from this laser is the highest reported value for any MM waveguide terahertz QCL. CW measurements were not performed for this device. In pulsed operation, $J_{\mathrm{th,5K}} \sim 255$ A/cm^2, $J_{\mathrm{max,5K}} \sim 1000$ A/cm^2, and $T_{\mathrm{max,pul}} \sim 165$ K. Without any corrections in the detected power, the single-facet wall-plug efficiency for this laser is $\sim 0.14\%$, and the slope-efficiency is ~ 28 mW/A (~ 1.6 photons/electron).

and hence is less representative of the threshold gain g_{th} in the device (section 3.3.1).

The FL183R-2 devices were also fabricated with the MM waveguides. Pulsed measurement results from one of the best performing MM Fabry-Pérot ridge lasers are shown in Fig. 3-52. In spite of a considerably small T_0 (section 3.3.4, Fig. 3-21) for the FL183R-2 lasers as compared to the FL-C series lasers, the temperature performance of this device is at par with the best of the FL-C series lasers (Fig. 3-16a, Fig. 3-18a). This device had a monotonously increasing power versus current relation up to the peak-bias (NDR) region even at temperatures close to the $T_{\mathrm{max,pul}}$ value, which is indicative of the fact that the peak current density is still the limiting factor for the temperature performance of this design. The peak optical power output from this laser ($\sim$ 24 mW) is a factor of $\sim$ 2 greater than that obtained from the best of the FL-C series lasers ($\sim$ 12 mW from a FL179C-M9-2 laser), and is the highest such reported value for any MM waveguide terahertz QCL. Both the high-power and

high-temperature performances of this laser can be attributed to a large value of $J_{\mathrm{max,5K}} - J_{\mathrm{th,5K}} \sim 750$ A/cm^2 for the FL183R-2 devices, wherein $J_{\mathrm{th,5K}}$ is lowered due to weaker low-bias parasitic current channels in this design because of its higher radiative energy as was discussed in the beginning of this section. This is the highest value of the lasing current range for any terahertz QCL discussed in this chapter. Also, the facet-reflectivities of the Fabry-Pérot MM waveguides are smaller for higher frequencies (chapter 4, Fig. 4-10b), which also causes a greater amount of optical power to be out-coupled from the FL183R-2 lasers. It may be mentioned that recently the peak power output from the device in Fig. 3-52 was increased by approximately 6 times to $\sim$ 145 mW by the use of a hyper-hemispherical Silicon lens abutted to the emitting facet [98]. A value of $T_{\mathrm{max,pul}} \sim 160$ K was obtained in that configuration, which is not significantly lower than the value $T_{\mathrm{max,pul}} \sim 165$ K obtained without the lens. It is likely that the increase in optical power with the lens is due in part to an improved collection of optical power by the lens, and also to a reduction in the mirror reflectivity of the facet in contact with the lens. More analysis on the effect of the lens is currently being pursued.

The relatively large difference in $J_{\mathrm{th,5K}}$ between the SISP and the MM waveguide implementations of the FL183R-2 QCLs cannot be simply explained due to lower losses in the MM waveguides (as calculated in Fig. 3-48). A reduction in the threshold gain g_{th} from approximately 54 cm^{-1} to 45 cm^{-1} for the SISP waveguide lasers in Fig. 3-51 reduced the $J_{\mathrm{th,5K}}$ by only 40 A/cm^2 to a value of 510 A/cm^2. This would imply that the optical losses for the MM waveguides need to be an order of magnitude lower than 50 cm^{-1}, otherwise a value of 255A/cm^2 for $J_{\mathrm{th,5K}}$ would seem disproportionately small. However, there is more to the low value of the threshold current densities in the MM waveguide lasers than just a lower value of the threshold gain. Fig. 3-53 shows cw I-Vs and differential-resistance versus voltage plots ($\mathcal{R}$-Vs) for both types of lasers. Whereas lasing starts only beyond the $2' \rightarrow 5$ injector anticrossing for the SISP device, it starts right after the $1' \rightarrow 3$ parasitic anticrossing for the MM device, which is when the $2' - 5$ level alignment starts to happen. This fact is also confirmed by the fact that the lasing frequencies for the MM device are lower,

(a) FL183R-2 (SISP), 98 μm $\times$ 2.15 mm
same device as the one in Fig. 3-51(a)

(b) FL183R-2 (MM), 80 μm $\times$ 1.10 mm
similar device to the one in Fig. 3-52

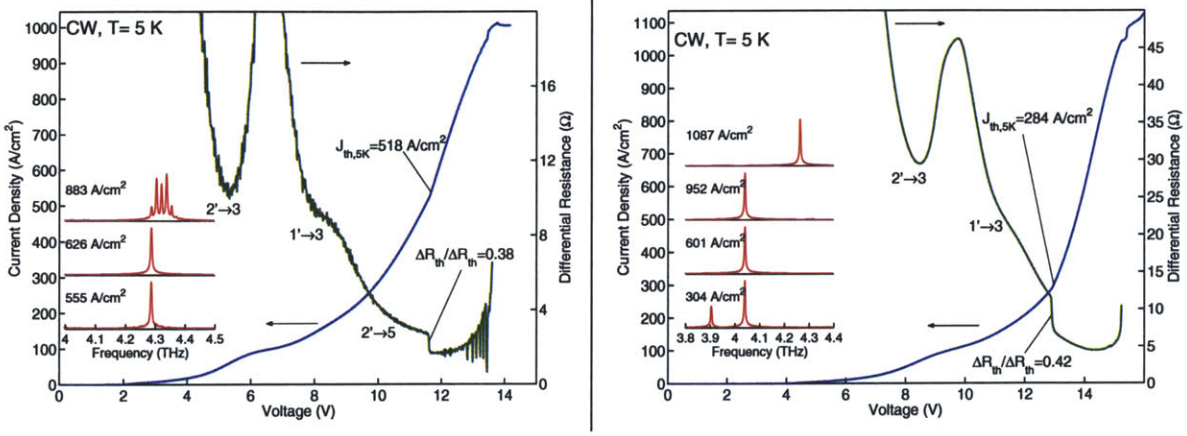


Figure 3-53: CW I-V, $\mathcal{R}$-V, and spectra from (a) a SISP waveguide FL183R-2 laser, and (b) a MM waveguide FL183R-2 laser to compare their transport characteristics. The $\mathcal{R}$-V in plot (a) is calculated by taking the derivative of the I-V as compared to the one in plot (b), which is measured experimentally as detailed in section 3.3.1. The following values were measured for the device in plot (b): $T_{\text{max,pul}} \sim 159$ K, $T_{\text{max,cw}} \sim 77$ K, and peak cw optical power output of ~ 5.3 mW at 5 K.

which is because the radiative energy E_{54} is smaller at the $\Delta_{2'5}$ bias as compared to the $\Delta_{1'5}$ bias due to the Stark effect. For QCLs based on the FL design, this difference in the spectral characteristics of the SISP and the MM waveguide lasers was observed in all the designs that have been realized with both types of the waveguides i. e. FL178C-M10, FL183R, and FL179R-M1 (except for FL178C-M1, which lased at much higher frequencies than expected for both types of waveguides, possibly due to high-field domain formation due to strong low-bias parasitic current channels [166]). Even though the current flowing through the structure might increase appreciably from the $\Delta_{2'5}$ bias to the $\Delta_{1'5}$ bias, the gain in the device is not likely to increase proportionately. As the predominant injection mechanism changes from $2' \rightarrow 5$ to $1' \rightarrow 5$, the gain linewidth in the active region is expected to increase without significantly increasing the peak gain. Consequently, it is possible that the value of g_{th} in the MM waveguides is not an order of magnitude smaller than 50 cm^{-1} at $\nu \sim 4$ THz as the $J_{\text{th,5K}}$ values might suggest. Hence, the experimental results in this section are not sufficient to provide a rough estimate for g_{th} in the MM waveguides, and any

FL178C-M10-2 (SISP), 148 μm $\times$ 2.10 mm, substrate thickness $\sim$ 300 μm
rear-facet HR coated, dry-etched ridges

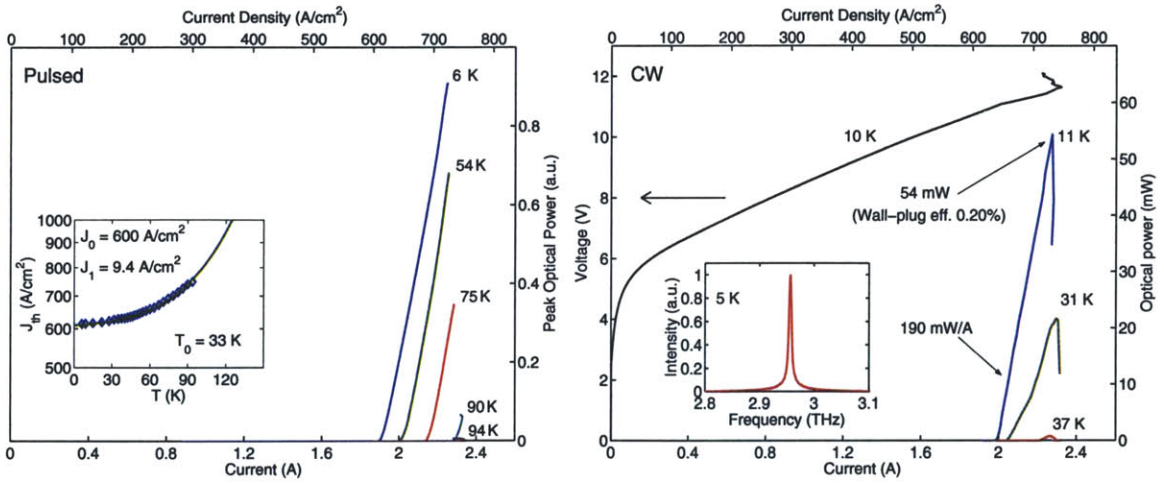


Figure 3-54: Experimental results from a FL178C-M10-2 (wafer EA1252) SISP waveguide laser. The corresponding results from FL178C-M10 (wafer EA1253) MM waveguide lasers are shown in Fig. 3-18. Even though no metal-metal devices were fabricated from the EAL1252 wafer, the wafers EA1252 and EA1253 were grown to within a thickness of $\sim$ 0.3% with respect to each other, hence their characteristics are expected to be very similar. For the pulsed measurement results in plot (a), a value of the peak optical power is not available.

value in the approximate range of $5 - 30$ cm^{-1} is possible at $\nu \sim 4$ THz[1].

When fabricated with the MM waveguides, the best of the FL183R-2 lasers had almost a similar temperature performance ($T_{\text{max,pul}} = 165$ K, Fig. 3-52) as for the best of the FL178C-M10 lasers ($T_{\text{max,pul}} = 169$ K, Fig. 3-18a). However, when fabricated with the SISP waveguides, the FL183R-2 devices performed considerably better. Experimental results from the best performing FL178C-M10-2 SISP waveguide laser are shown in Fig. 3-54. As can be noted, the value of $T_{\text{max,pul}} \sim 94$ K for this device is approximately $10 - 15$ K lower than that for the FL183R-2 devices. The worse temperature performance for this laser is attributed to the lower radiative frequency in this design ($\nu \sim 2.9$ THz). This is because the SISP waveguides have better mode confinement and lower losses at higher frequencies, as was described in the beginning of this section and is also apparent from the Drude model calculations of Fig. 3-48.

[1]Note that the Drude model calculations in Fig. 3-48 cannot be taken in an absolute measure due to the uncertainties in the values of the Drude parameters used and also due to the uncertainty in the contribution of the active region to the optical losses.

FL179R-M1: A QCL operating in the frequency range of $\nu \sim 4.7 - 5.0$ THz

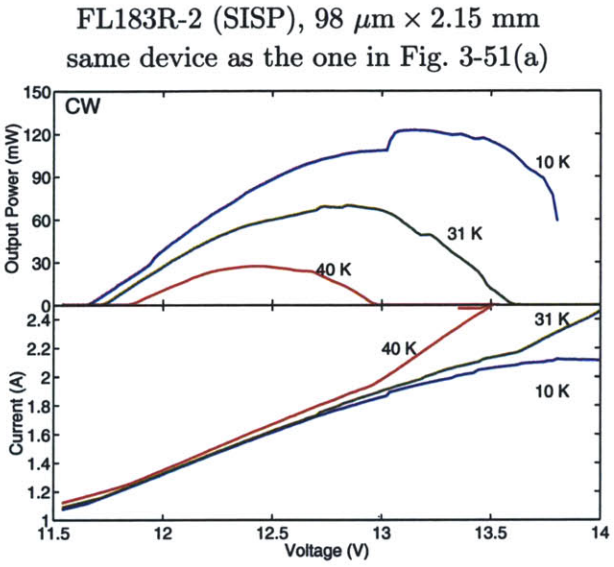


Figure 3-55: CW I-Vs and L-Vs for the FL183R-2 laser of Fig. 3-51(a) shown to highlight the thermal runaway phenomenon that quenches lasing in the device before the peak bias could be reached during cw operation.

The FL183R-2 QCLs had a problem due to the premature quenching of lasing before the peak bias could be reached during cw operation, which became worse as the operating temperatures increased. This problem surfaced particularly in the SISP waveguide lasers due to their bigger cavity dimensions and correspondingly a larger current requirement that lead to greater amount of electrical heating during operation. To highlight this problem, the cw I-Vs and L-Vs versus temperature for one such laser are shown in Fig. 3-55. As can be noticed, the maximum lasing bias in the L-Vs corresponds with an upward kink in the I-Vs at the same bias. This is perceived to be caused by a thermal runaway phenomenon due to which the current increases rapidly in the device, possibly due to the hot electrons leaking out energetically over the top of the potential barriers into the continuum. A strong upper level parasitic current channel of the type discussed in section 3.2.1 due to the $5' - 8$ inter-module coupling in the FL183R-2 design ($\Delta_{5'8} \sim 1.3$ meV, Fig. 3-49) is likely to be the cause of this effect. This problem is difficult to avoid in the higher frequency FL designs, due partly to the fact that the upper radiative state is energetically located close to

the top of the barriers as the radiative energy E_{54} is made higher, and also that the radiative barrier is thinner for such a design, which brings the upper parasitic level 8 closer in energy to the radiative level 5′ of the adjacent module (see Fig. 3-36 for a two-module conduction band diagram at design bias for a typical FL design).

A modified design, named as FL179R-M1, was grown in attempt to obtain higher frequency operation, nearer to the 4.7 THz neutral oxygen line (important for monitoring the cooling of interstellar dust), and the atmospheric transmission window at $4.8-4.95$ THz. The design details for FL179R-M1 are shown in Fig. 3-56. When compared to the FL183R-2 design, the radiative barrier in FL179R-M1 was thinned by 0.5 ML to increase the lasing frequency, and the injector barrier, the intra-injector barrier, and the collector barrier were all increased by 1 ML, respectively, in an attempt to suppress the aforementioned $5'-8$ parasitic coupling. This modification did prevent any runaway current phenomenon, but at the cost of a reduced $J_{\max}$ (~ 500 A/cm^2) as compared to that for the FL183R-2 lasers (~ 1000 A/cm^2) as shown in the experimental results from two different lasers in Fig. 3-58. Consequently, the peak optical power and the temperature performance of these lasers turned out to be worse then that for the FL183R-2 lasers. Nevertheless, since their lasing frequencies ($\nu \sim 4.7 - 5.0$ THz) lie close to a low-absorption atmospheric transmission window at $\nu \sim 4.8 - 4.95$ THz (Fig. 3-57), transmission mode imaging over a distance of 25 meters in air was possible with the FL179R-M1 devices [100]. The FL179R-M1 SISP waveguide devices typically lased at frequencies around 4.75 THz or 4.9 THz, depending on the lasing threshold of the cavity, which determines the electrical bias at which the device starts to lase. The device in Fig. 3-58(b) lased at a higher frequency ($\nu \sim 4.94$ THz) due to its slightly higher threshold current density as compared to the device in Fig. 3-58(b) ($\nu \sim 4.77$ THz). Also, the device in Fig. 3-58(b), which had better thermal properties due to it being narrower, emitted higher optical power of ~ 17 mW at the cryorefrigerator temperature of ~ 30 K as against a value of ~ 7 mW for the device in Fig. 3-58(a) when operated in a quasi-cw mode ($\sim 25\%$ duty-cycle). This is in spite of the fact that the narrower device had smaller amount of optical power at 5 K. Due to its favorable thermal and spectral properties, transmission-mode real-time imag-

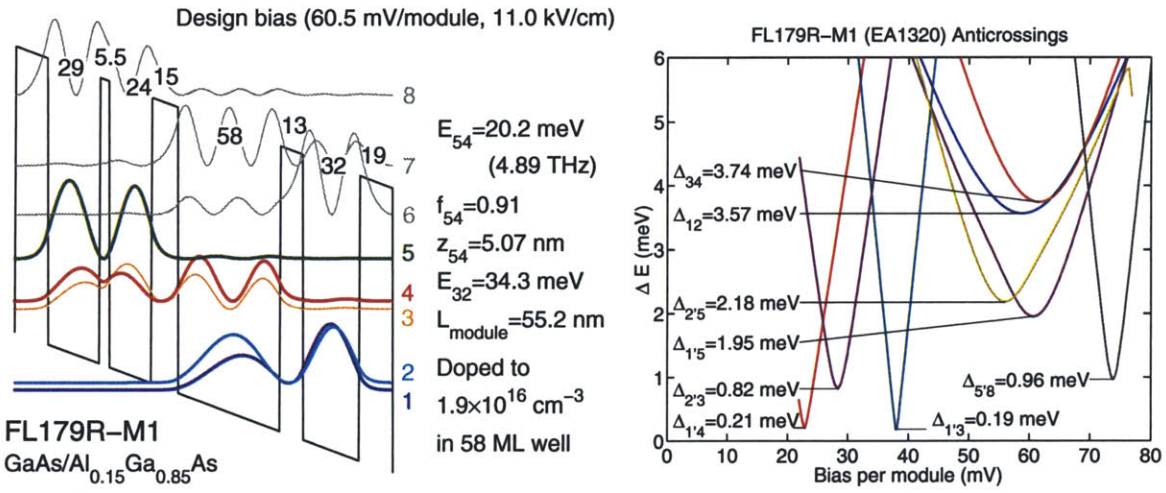


Figure 3-56: One module conduction band diagram at the design bias (left) and anticrossing plots for a two-module simulation (right) for FL179R-M1, the highest frequency design from amongst the FL-series of designs (section 3.3).

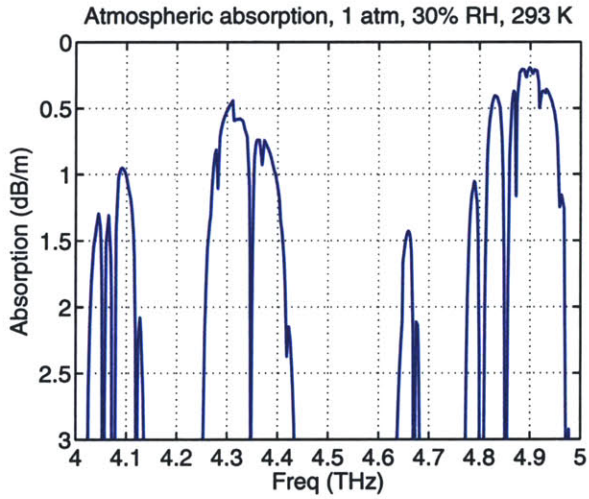


Figure 3-57: Atmospheric path loss in dB/m (mostly caused by absorption due to water vapor) calculated from HITRAN 2004 [3] at $T = 293$ K, 30% relative humidity, and 1 atm pressure for a continuous frequency range of $4 - 5$ THz.

(a) FL179R-M1 (SISP), 150 μm × 2.05 mm rear-facet HR coated, dry-etched ridges substrate thickness $\sim$ 170 μm

(b) FL179R-M1 (SISP), 100 μm × 1.97 mm rear-facet HR coated, dry-etched ridges substrate thickness $\sim$ 170 μm

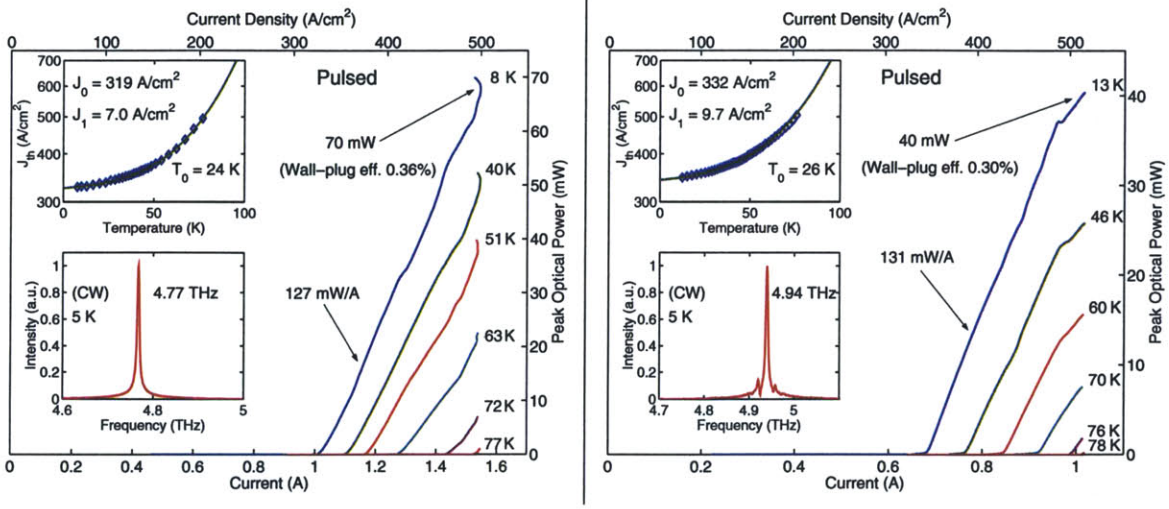


Figure 3-58: Experimental results from two different FL179R-M1 SISP waveguide lasers. The laser in plot (b) has been recently used to demonstrate real-time terahertz imaging at a standoff distance of 25 meters [100] due to its lasing frequency being in a low-absorption atmospheric transmission window that occurs from $4.8 - 4.95$ THz.

ing over a standoff distance of > 25 meters could be demonstrated with the laser in Fig. 3-58(b) [100]. It may be noted that FL179R-M1 wafer was also fabricated with the MM waveguides. The MM waveguide lasers (fabricated by completely removing the highly-doped top-contact layer) obtained values of $T_{\text{max,pul}} \sim 125 - 130$ K, which are considerably lower than that for the FL183R-2 MM waveguide lasers due to a smaller lasing range in current. However, it is likely that the performance of the FL179R-M1 lasers could be improved significantly in a subsequent design iteration by making the injection mechanism more transparent and by making the radiative transition more diagonal to reduce the aforementioned $5' - 8$ parasitic coupling.

3.7 Conclusions and summary

This chapter discussed many different aspects of the design and the operating principles of the active region of a terahertz QCL. A multitude of different designs were discussed, with each one of them aimed towards a slightly different functionality or to

better the performance of the similar designs that had been tried in the past. Apart from the common characteristic of an electron-longitudinal-optical (LO) phonon scattering based depopulation mechanism, other design characteristics were changed by varying degrees from one design to another, often yielding contrasting experimental results. These have cumulatively led to a better understanding of the electrical transport mechanisms in these devices, which have otherwise only been partially understood by the aid of complex theoretical models.

For the various possible applications of terahertz sources which were described in section 1.1, the most important requirement for terahertz QCLs is high temperature of operation such they become easily accessible for a wide range of operating environments. After the first terahertz QCL was demonstrated [89], the second major development happened when continuous-wave operation was demonstrated above the liquid-nitrogen temperature. This was achieved in a design based on LO phonon depopulation (Fig. 3-13a). The next, and a significantly more important milestone would be to obtain lasing at $T \approx 250$ K, a temperature at which operation with thermoelectric coolers is possible. This will greatly enhance the practical usability of such a laser source and also might lead to a much greater economic interest in the terahertz field in general, which will bode well for rapid research and development related to this otherwise unexplored region of the electromagnetic spectrum.

The first LO phonon depopulation based terahertz QCL, and the first from this research group, resulted from the so-called "FL" designs, which utilize a unique "resonant-phonon" based depopulation mechanism [168]. The operating principles of the FL design are described in great detail in section 3.3.1. Since its first development, more then ten variations of that design have been developed in the past four years, and significant performance improvements have be obtained. Currently, the modified versions of the first successful FL design hold the record for the highest pulsed and cw operating temperatures of 169 K (Fig. 3-18a) and 117 K (Fig. 3-16b), respectively. The temperature degradation mechanisms in these designs are discussed in detail in section 3.3.4. Several different variations of the FL design with some completely new designs were also implemented and many of them achieved lasing. These

designs were mainly aimed towards obtaining an improved temperature performance and are described in section 3.4. Whereas none of those designs bettered the temperature performance of the FL designs, the DSL designs described in section 3.4.2 currently hold the best promise for an alternative design. A value of $T_{\max,\text{pul}} \sim 200$ K for future design iterations of the FL design is perceivable; however, to reach close to a value of 250 K, a different design technique is deemed likely to be needed.

Another important goal for terahertz QCLs is to achieve lasing at a frequency as low as 1 THz ($\lambda \sim 300$ μm, $\hbar\omega \sim 4$ meV), to finally bridge the "terahertz gap" that exists between the electronic and optical sources of electromagnetic radiation. Besides a matter of scientific interest, the lower frequency lasers have some specific applications as were discussed in section 1.1. The primary challenges for QCL operation at such a low frequency are the increased intersubband absorption losses in the active region, and also reduced selectivity of injection into the upper radiative level, or depopulation from the lower radiative level, since the radiative energy becomes of the order of the collisional broadening of the subbands ($\sim$ few meVs). Towards this goal, a design based on a one-well injector scheme was developed, which significantly cuts down photon absorption-loss in the active region, and provides improved injection selectivity into the upper radiative level as was described in section 3.5.2. A terahertz QCL operating at 1.59 THz ($\lambda \sim 190$ μm) was demonstrated with such a design (Fig. 3-46), which is the second lowest reported frequency for any solid-state laser that operates without the assistance of an external magnetic field. Although it is difficult to speculate the lowest frequency that can be obtained from the specific one-well injector designs discussed in this chapter, it can be expected that operation at $\nu \sim 1$ THz might be possible with future design iterations, especially since the one-well injector QCLs operating at frequencies of 2.1 THz, 1.9 THz, 1.7 THz, and 1.6 THz, respectively, were all obtained with the first designs that were tried (as discussed in section 3.5.2).

The development of the metal-metal (MM) waveguides for terahertz QCLs have had an equally significant role to play as compared to the design of the active region, in obtaining the existing record temperature performances and operation at lower

frequencies. This is primarily due to their lower waveguide loss and stronger mode confinement at the terahertz frequencies as compared to the semi-insulating surface-plasmon (SISP) waveguides, which are other types of waveguides that are commonly used for terahertz QCLs. A comparison of the MM waveguides to the SISP waveguides was done in section 3.6.1. The MM waveguides will be discussed in a much greater detail in chapter 4.

It may be noted that despite their worse temperature performance due to a higher waveguide loss, the SISP waveguides can out-couple greater amount of optical power in comparison to the MM waveguides. Hence, for high-power operation, terahertz QCLs have so far been demonstrated with the SISP waveguides only, as described in section 3.6. Record high optical power output with a value of 248 mW in pulsed operation and 138 mW in cw operation (Fig. 3-50) has been demonstrated with the aforementioned FL designs when fabricated with the SISP waveguides. High optical power and good beam properties are critical for important applications such as terahertz imaging. Various techniques to improve the radiative properties of the MM waveguides will be described in chapter 4, with the objective of obtaining high-power operation with good radiation patterns while retaining their better temperature performance as compared to the SISP waveguides.

Chapter 4

Metal-metal waveguides for terahertz mode confinement

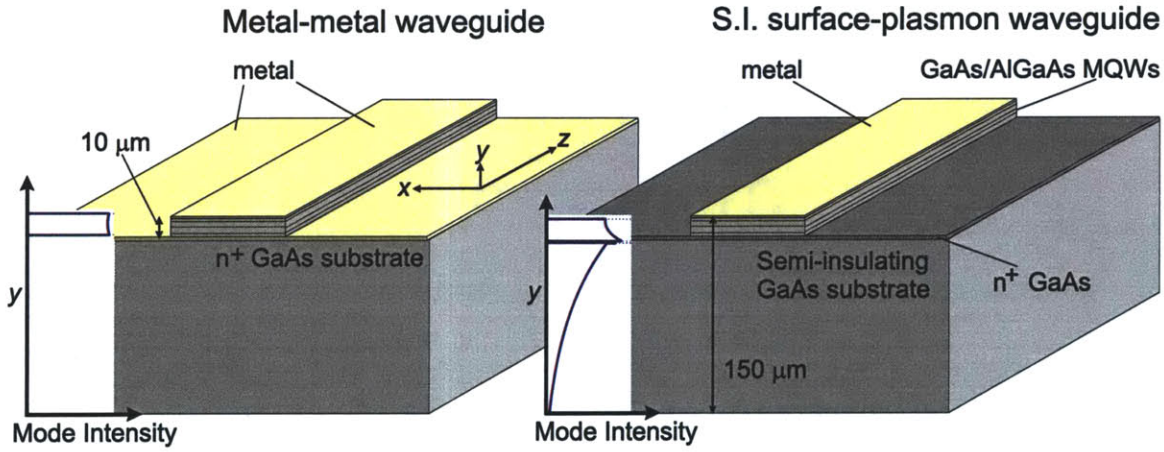


Figure 4-1: Schematic and mode-shape plots at $\nu \sim 3$ THz ($\lambda_{GaAs} \sim 28$ μm) for the two types of waveguides that are used for terahertz QCLs (reproduction of Fig. 2-3).

Figure 4-1 shows the schematics of the two types of waveguides that are used for terahertz QCLs. Some of the key distinctive electromagnetic properties of these waveguides were discussed in chapter 3 (section 3.6.1). Terahertz QCLs with the best temperature performance have been demonstrated in the metal-metal waveguides ($T_{max,pul} \sim 169$ K in Fig. 3-18a, $T_{max,cw} \sim 117$ K in Fig. 3-16b). A more conclusive evidence of this was shown by comparing the temperature performance of a $\nu \sim 4.2$ THz QCL gain-medium (wafer FL183R-2, section 3.6.2) that was fabricated with both

types of waveguides. Whereas a value of $T_{\text{max,pul}} \sim 113$ K was obtained for a semi-insulating surface-plasmon (SISP) waveguide laser (Fig. 3-51a), a much higher value of $T_{\text{max,pul}} \sim 165$ K was obtained for a metal-metal (MM) waveguide laser (Fig. 3-52).

The MM waveguides, which are similar to the microstrip transmission lines commonly used at the microwave frequencies, have efficient waveguiding properties at the terahertz frequencies primarily due to two reasons. First, they inherently have a lower waveguide loss α_{w} as compared to the SISP waveguides. Second, they provide near unity mode confinement even in cavities with length or width dimensions of the order of the wavelength of light inside the semiconductor material. This is due to the fact that the terahertz mode is confined in a sub-wavelength thickness inbetween the metal claddings, which causes a large mode mismatch between the waveguided mode and any free-space propagation modes, thereby leading to a negligible leakage of the waveguided mode into the free-space. However, this leads to poor out-coupling of light from the facets (and correspondingly high facet-reflectivities [88]), highly divergent radiation patterns from the edge emitting lasers, and easy excitation of the higher order lateral modes (i. e. along the width of the waveguide) that makes the lasers multi-moded even for very narrow waveguides. This, then necessitates the use of other techniques, such as implementation of distributed feedback, to modify the spectral and radiation properties of these waveguides for them to be useful for most of the practical applications.

This chapter discusses many different aspects of the MM waveguides including the pertinent fabrication methods. Following the discussion in section 3.6.1, the electromagnetic properties of the Fabry-Pérot ridge waveguides will be described in greater detail. Subsequently, various techniques for implementing both first-order and second-order distributed feedback (DFB) structures to improve the spectral and/or the radiative characteristics of these waveguides will be described. Many techniques of the conventional DFB theory, which have otherwise been successfully applied to the mid-infrared QCLs, are no longer applicable to the MM waveguides at the terahertz frequencies. Hence, the analysis and design of the DFB structures has to be significantly modified, as will become apparent in the following sections.

4.1 Surface-plasmon electromagnetic modes

Plasmons are quanta associated with longitudinal waves propagating in matter through collective motion of large numbers (i. e. a plasma) of electrons. Surface-plasmons (SP) are a subset of these charge oscillations, which are bound to regions at the interface between a dielectric and a conducting medium. In the low frequency limit, these surface charge oscillations, which carry an electric field of their own, can couple with the external electromagnetic fields to yield a polariton type excitation. These collective excitation modes are referred to as the surface-plasmon-polariton (SPP) modes [130, 19], by means of which the electromagnetic energy can propagate at the interface. Although not strictly correct, the SPP modes are often just called the surface-plasmon electromagnetic modes in literature, which is what they will be referred to in this chapter as well.

4.1.1 Single surface-plasmon electromagnetic mode propagation

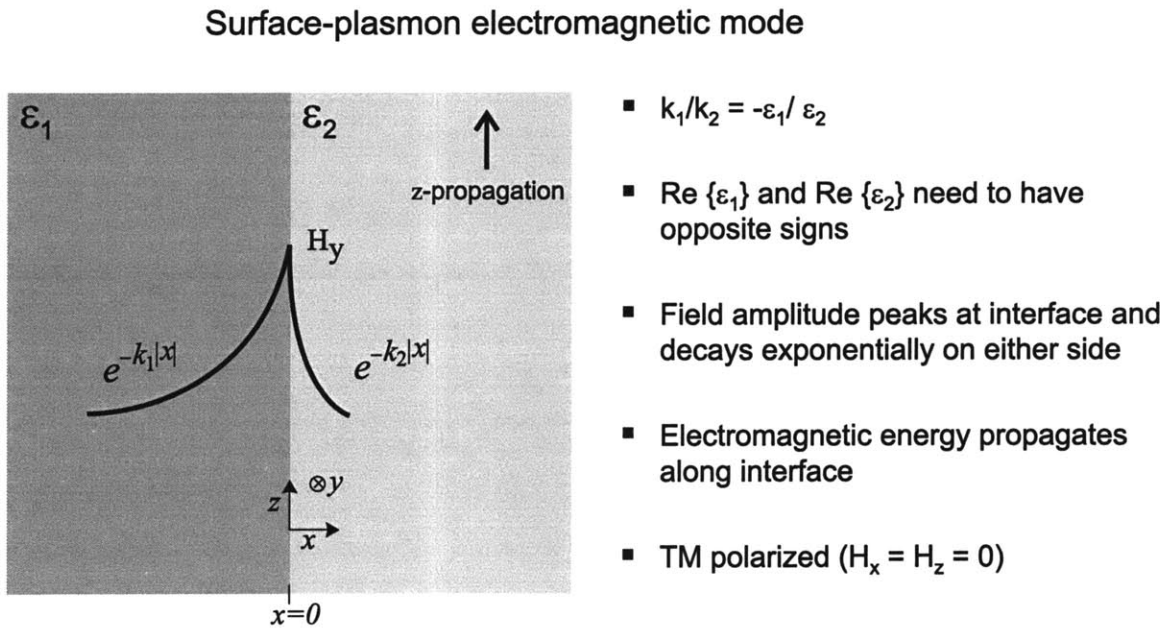


Figure 4-2: A schematic showing the concept of surface-plasmon electromagnetic wave propagation at the interface of a dielectric and a conducting medium.

The electromagnetic fields associated with a SP electromagnetic mode decay exponentially with distance from the interface into each of the bounding media. Such a mode is transverse-magnetic (TM) polarized and is shown schematically in Fig. 4-2. For the mode to exist, the real part of the dielectric constant on either side of the interface must have opposite signs. This condition can be satisfied at the interface of a highly conducting layer (such as a metal or a highly doped semiconductor) and a dielectric medium (such as a low-doped semiconductor, an insulator, or air). According to the Drude model [16], the contribution of the "free" electrons to the complex relative permittivity (in SI units) of a conducting medium can be included as

$$\begin{aligned} \epsilon_{\mathrm{r}} &= \epsilon_{\mathrm{core}} + i\frac{ne^2\tau}{\epsilon_0\omega m^*(1-i\omega\tau)} \\ &= \epsilon_{\mathrm{core}}\left(1 - \frac{\omega_{\mathrm{p}}^2\tau^2}{1+\omega^2\tau^2} + i\frac{\omega_{\mathrm{p}}^2\tau}{\omega(1+\omega^2\tau^2)}\right) \end{aligned} \tag{4.1}$$

where, ϵ_{core} is the relative permittivity of the material excluding the contribution due to the free electrons, m^* is the electron effective mass, ϵ_0 is the permittivity of vacuum, n is the electron concentration in the conducting medium, e is the electronic charge, τ is the effective (Drude) scattering time for the electrons, and ω is the frequency of the electromagnetic wave. ω_{p} is the plasma frequency in the material and is given by

$$\omega_{\mathrm{p}} = \sqrt{\frac{ne^2}{\epsilon_0\epsilon_{\mathrm{core}}m^*}} \tag{4.2}$$

For example, in Au $\omega_{\mathrm{p}}/2\pi \sim 2000$ THz ($\lambda_{\mathrm{p}} \sim 0.15$ μm) and for highly doped GaAs ($n \sim 5\times10^{18}$ cm^{-3}) $\omega_{\mathrm{p}}/2\pi \sim 20$ THz ($\lambda_{\mathrm{p}} \sim 15$ μm). At frequencies below approximately the plasma frequency, the real part of the dielectric constant becomes negative in the conducting medium, and therefore a SP electromagnetic mode can be sustained at the interface. For a metal with a complex relative permittivity ϵ_{m} and a dielectric with a complex relative permittivity ϵ_{d}, and for spatially varying magnetic field $H_y(x,y,z)$ of the form $e^{-k_{\mathrm{m}}|x|+i\beta_z z}$ in the metal and $e^{-k_{\mathrm{d}}|x|+i\beta_z z}$ in the dielectric,

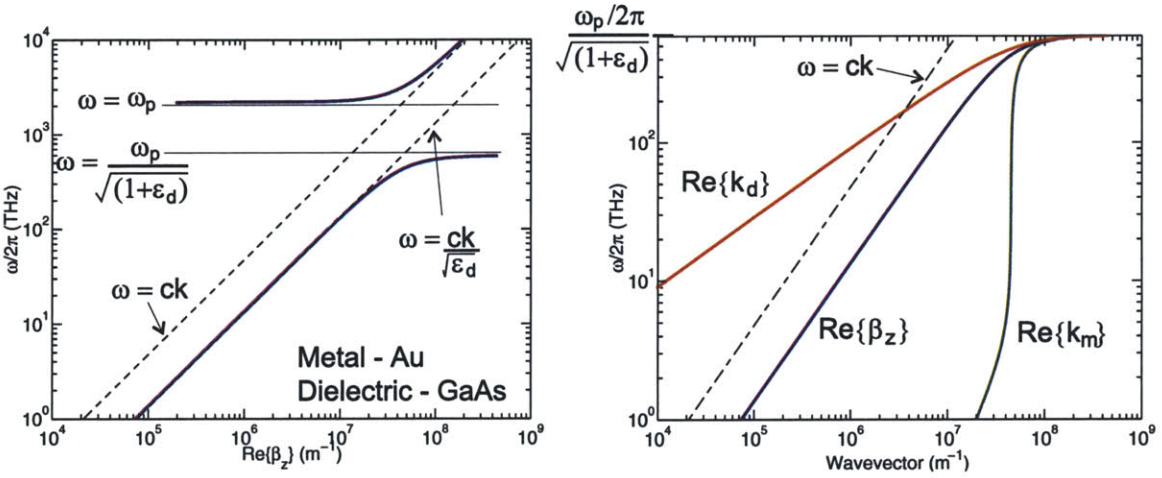


Figure 4-3: Dispersion relations for a surface-plasmon electromagnetic mode propagating at the interface of a metal (Au) and a dielectric (GaAs). The values of various parameters used in the calculation are mentioned in the caption of Fig. 4-4. Note that the wavevectors β_z, k_m, and k_d are complex quantities. While β_z is predominantly real at all frequencies, the quantities k_m and k_d have comparable real and imaginary components for $\omega\tau < 1$ (for $\omega\tau > 1$, the real components dominate).

respectively, the following relations are obtained by solving the Maxwell's equations

$$\frac{k_\mathrm{d}}{k_\mathrm{m}} = -\frac{\epsilon_\mathrm{d}}{\epsilon_\mathrm{m}} \tag{4.3}$$

$$\beta_z = \frac{\omega}{c}\sqrt{\frac{\epsilon_\mathrm{d}\epsilon_\mathrm{m}}{\epsilon_\mathrm{d}+\epsilon_\mathrm{m}}} \tag{4.4}$$

where c is the speed of light in vacuum, and

$$\begin{aligned} k_\mathrm{d}^2 &= \beta_z^2 - \frac{\omega^2}{c^2}\epsilon_\mathrm{d} \\ k_\mathrm{m}^2 &= \beta_z^2 - \frac{\omega^2}{c^2}\epsilon_\mathrm{m} \end{aligned} \tag{4.5}$$

Equation (4.4) leads to a polariton like dispersion and is plotted in Fig. 4-3. As can be noticed, the wavevector k_m in metal is much larger than that for the light-line ($k = \omega/c$), which means that the electromagnetic energy is confined within a dimension that is much smaller than the wavelength. This is the primary reason why surface-plasmon electromagnetic modes can propagate at subwavelength dimensions.

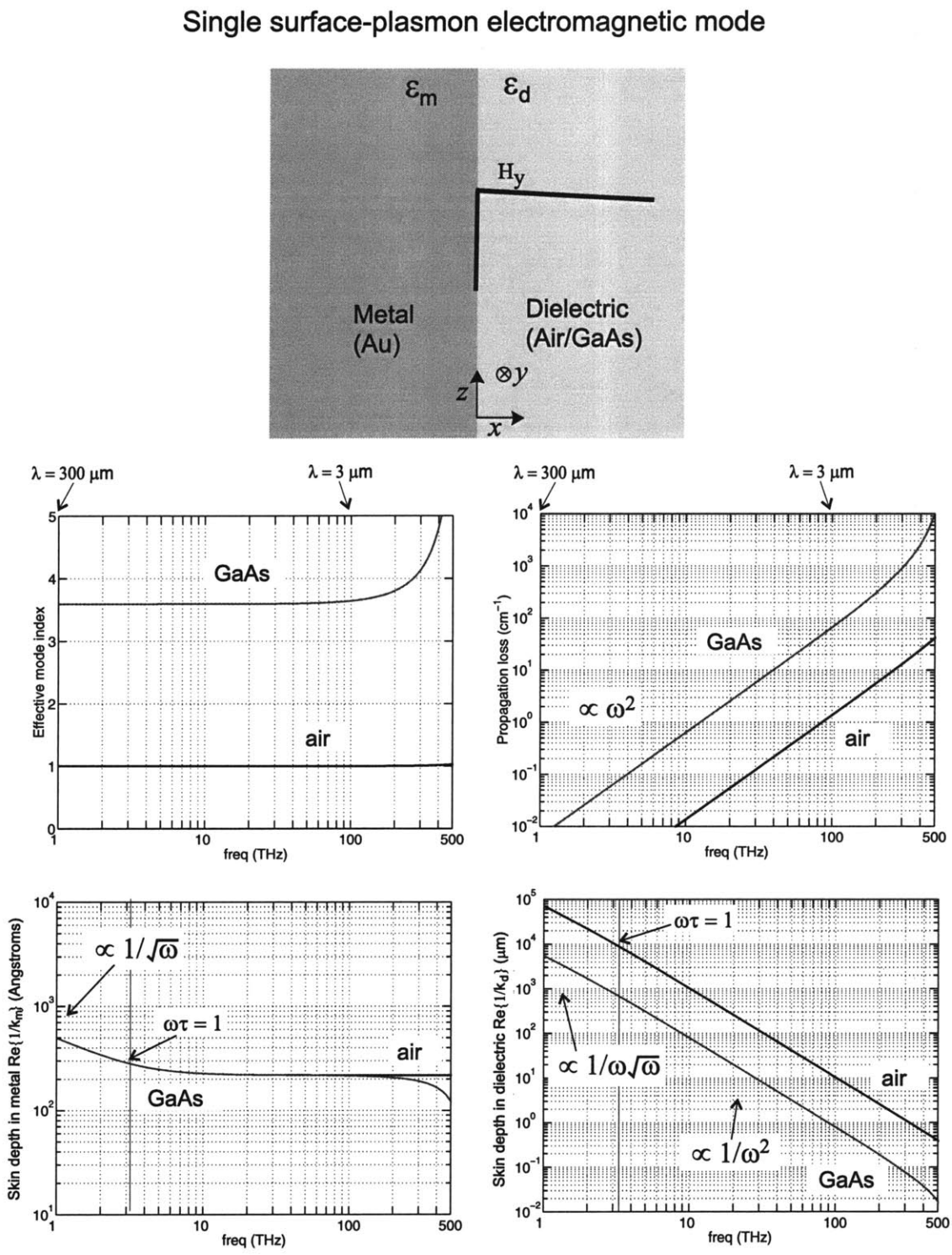


Figure 4-4: Schematic, mode propagation parameters, and field extinction lengths on either side of the interface for a single surface-plasmon electromagnetic mode. The values of $n = 5.9 \times 10^{22}$ cm^{-3}, $\tau = 50$ fs, $m^* = m_0$, $\epsilon_{\mathrm{m,core}} = 1$ are taken for Au [4, 166], and $\epsilon_{\mathrm{d}} = 12.9$, $m^* = 0.067\, m_0$ are taken for GaAs, respectively. It may be noted that these calculations are solely due to the free-carrier losses as calculated using a Drude model (equation 4.1), and contributions due to photon-phonon (polariton) coupling (i. e. the *Reststrahlen* band in the semiconductor), for example, are not included.

The other relevant propagation parameters for the single interface SP electromagnetic mode are plotted in Fig. 4-4. The mode propagation wavevector β_z has a small imaginary component that is caused by the free-carrier losses in the conducting medium (due to a finite Drude scattering time τ), which causes the mode intensity to decay exponentially as it propagates. The mode propagation loss is written as $\alpha = 2\mathcal{I}\mathrm{m}\{\beta_z\}$ and the effective mode index becomes $n_{\mathrm{eff}} = \mathcal{R}\mathrm{e}\{\beta_z\}\, c/\omega$. The field extinction lengths ($1/e$ decay lengths) are given by $1/\mathcal{R}\mathrm{e}\{k_{\mathrm{m}}\}$ and $1/\mathcal{R}\mathrm{e}\{k_{\mathrm{d}}\}$, for the metal and the dielectric, respectively, and are also plotted in Fig. 4-4. The extinction length in the metal is extremely small, and thus almost all of the modal volume exists in the dielectric medium. However, for a lossless dielectric, all the loss contribution is due to the small fraction of the mode that propagates in the metal.

Evaluating equation (4.4) for $\omega \ll \omega_{\mathrm{sp}}$, the following relations are obtained

$$\mathcal{R}\mathrm{e}\{\beta_z\} \propto \omega \tag{4.6}$$

$$\mathcal{I}\mathrm{m}\{\beta_z\} \propto \omega^2 \tag{4.7}$$

$$\mathcal{R}\mathrm{e}\{\beta_z\} \gg \mathcal{I}\mathrm{m}\{\beta_z\} \tag{4.8}$$

where ω_{sp} is the surface-plasmon resonance frequency given by

$$\omega_{\mathrm{sp}} = \frac{\omega_{\mathrm{p}}}{\sqrt{1+\epsilon_{\mathrm{d}}/\epsilon_{\mathrm{m,core}}}} \tag{4.9}$$

For the Au/GaAs interface $\omega_{\mathrm{sp}}/2\pi \sim 600$ THz. For $\omega \ll \omega_{\mathrm{sp}}$, the mode propagation loss $\alpha = 2\mathcal{I}\mathrm{m}\{\beta_z\} \propto \omega^2$ and the effective mode index $n_{\mathrm{eff}} = \mathcal{R}\mathrm{e}\{\beta_z\}\, c/\omega \approx \sqrt{\epsilon_{\mathrm{d}}}$; however, as $\omega \rightarrow \omega_{\mathrm{sp}}$, $\epsilon_{\mathrm{m}} \rightarrow (-\epsilon_{\mathrm{d}})$ and thus $(\beta_z,\, n_{\mathrm{eff}},\, \alpha) \rightarrow \infty$ (equation 4.4).

The variation of the skin-depth in dielectric $1/\mathcal{R}\mathrm{e}\{k_{\mathrm{d}}\}$ with ω can be obtained from equations (4.4) and (4.5). For $\omega \ll \omega_{\mathrm{sp}}$, the following expressions are derived

$$k_{\mathrm{d}}^2 \approx \begin{cases} \omega^4 \dfrac{\epsilon_{\mathrm{d}}^2}{c^2 \epsilon_{\mathrm{m,core}} \omega_{\mathrm{p}}^2} & \propto \omega^4 \quad \ldots \omega\tau \gg 1 \\ i\, 2\mathcal{R}\mathrm{e}\{\beta_z\}\mathcal{I}\mathrm{m}\{\beta_z\} & \propto \omega^3 \quad \ldots \omega\tau \ll 1 \end{cases} \tag{4.10}$$

which explains the frequency variation of $1/\mathcal{R}e\{k_d\}$ in Fig. 4-4. It may be noted that $\mathcal{I}m\{k_d^2\} = \mathcal{R}e\{k_d^2\}$ for $\omega\tau = 1$. The variation of the skin-depth in metal $1/\mathcal{R}e\{k_m\}$ with ω can be determined using equation (4.3) and the following expressions for the relative permittivity of metal

$$\epsilon_m \approx \begin{cases} -\epsilon_{m,core}\dfrac{\omega_p^2}{\omega^2} & \propto \dfrac{1}{\omega^2} \quad \ldots\omega\tau \gg 1 \\ i\,\epsilon_{m,core}\dfrac{\omega_p^2\tau^2}{\omega} & \propto \dfrac{1}{\omega} \quad \ldots\omega\tau \ll 1 \end{cases} \tag{4.11}$$

The frequency variations of the skin-depths $1/\mathcal{R}e\{k_d\}$ and $1/\mathcal{R}e\{k_m\}$ so determined (Fig. 4-4) explains the $\alpha \propto \omega^2$ variation since the fraction of the mode in the metal, which is the lossy medium, continually increases with frequency as ω^2 for $\omega \ll \omega_{sp}$.

It may be noted that the dominant component of the electric field in the metal is in the longitudinal (z) direction, since $|E_x| = |E_z\,\beta_z/k_m| \rightarrow 0$, and $E_y = 0$ due to the TM polarization. However, in the dielectric medium, the transverse component E_x is much larger than the longitudinal component E_z (since $|E_x| = |E_z\,\beta_z/k_d|$), which is compatible with the intersubband polarization selection rule (section 2.1.1, equation 2.25) for the QCLs.

4.1.2 Double surface-plasmon electromagnetic mode propagation

Although the single SP mode as shown in Fig. 4-4 promises very low propagation losses at the terahertz frequencies, there is no practical way of making a QCL with such a waveguide, especially since the extinction length in the semiconductor is very large (of the order of 1 mm). However, the semiconductor region can be sandwiched between metal layers on its either side to obtain the so-called metal-metal waveguides, which have a double surface-plasmon electromagnetic mode propagation as shown schematically in Fig. 4-5. Such a waveguide is particularly suited for QCLs since it is compatible with the requirement of the electromagnetic field to be TM polarized for intersubband transitions (section 2.1). By writing the double SP mode as a superposition of two single SP modes the following dispersion relation is obtained

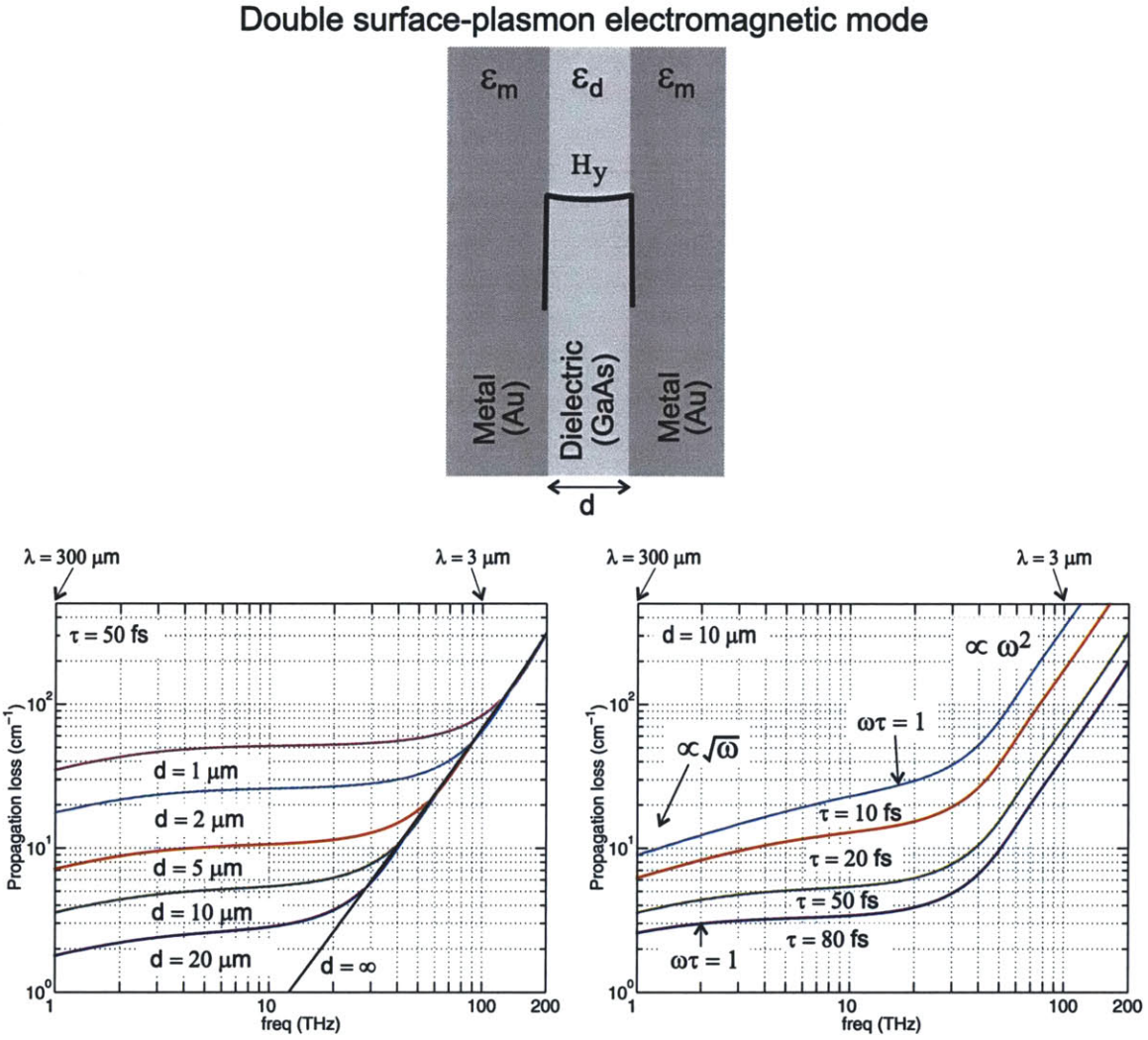


Figure 4-5: Schematic of a double surface-plasmon electromagnetic mode and the corresponding calculations of the mode propagation loss for different parameters. τ is the Drude scattering time in the metal. The values for the various other parameters remain the same as in Fig. 4-4. Similar to the calculations in Fig. 4-4, these calculations show the contribution due to the free-carrier losses only, neglecting other effects such as the photon-phonon (polariton) coupling (i. e. the *Reststrahlen* band in the semiconductor).

from a solution of the Maxwell's equations

$$\frac{\beta_z^2 - \frac{\omega^2}{c^2}\epsilon_\mathrm{d}}{\beta_z^2 - \frac{\omega^2}{c^2}\epsilon_\mathrm{m}} = \frac{\epsilon_\mathrm{d}^2}{\epsilon_\mathrm{m}^2}\left(\frac{1+e^{-k_\mathrm{d}d}}{1-e^{-k_\mathrm{d}d}}\right)^2 \tag{4.12}$$

where,

$$\begin{aligned} k_\mathrm{d}^2 &= \beta_z^2 - \frac{\omega^2}{c^2}\epsilon_\mathrm{d} \\ k_\mathrm{m}^2 &= \beta_z^2 - \frac{\omega^2}{c^2}\epsilon_\mathrm{m} \end{aligned} \tag{4.13}$$

The dispersion relation in equation (4.12) is a transcendental expression that can be solved analytically, the results of which are plotted for various parameters in Fig. 4-5. The skin depth in the semiconductor increases as $1/\omega^2$ or $1/\omega^{3/2}$ for lower frequencies (Fig. 4-4). For a typical value of $d \sim 10$ μm, the skin depth in the semiconductor becomes larger than d at the terahertz frequencies and the mode propagation loss α no longer varies as ω^2. At frequencies for which $\omega\tau \lesssim 1$, the conductivity of the metal ($\frac{ne^2\tau/m^*}{1-i\omega\tau}$) is a weak function of ω, the skin depth varies as $1/\sqrt{\omega}$ (Fig. 4-4), and thus the sheet resistance of the metal, and consequently α, vary as $\sqrt{\omega}$. However, for a typical value of $\tau \sim 50$ fs for the metal at the cryogenic temperatures, the $\alpha \propto \sqrt{\omega}$ variation happens for frequencies well below the terahertz frequencies, as can be seen from Fig. 4-5. Nevertheless, the losses decrease monotonously over the entire frequency range. Also note that α varies as $1/d$ at the terahertz frequencies since the mode almost uniformly fills the semiconductor region, whereby any reduction in the thickness of the semiconductor increases the fraction of the mode overlapping with the metal, and thus the propagation loss α, proportionately.

4.2 Metal-metal waveguide fabrication: In-Au and Cu-Cu bonding techniques

The MM waveguide fabrication process is relatively challenging since it involves substrate removal of the MBE grown wafer after a flip-chip wafer to wafer bonding step.

The substrate removal step typically involves a combination of mechanical lapping and wet-chemical etching during which the wafer to wafer bond should not degrade. Also, a good etch-stop layer is needed to stop the substrate removal immediately before the etchant reaches MBE grown layers. Once the entire fabrication is completed, the metal-metal bonding region must be robust not only mechanically (since wire-bonding on the fabricated devices puts significant strain on the bonding layers) but also a good conductor both thermally and electrically. Moreover, the metal layer(s) to be used in proximity of the active region should have a low loss at the terahertz frequencies, or in other words, should have a high electrical conductivity.

Figure 4-6 summarizes an In-Au reactive bonding based fabrication technique for the MM waveguides, which was used for the first demonstration of a terahertz QCL with MM waveguides [170]. The fabrication procedure is described in detail in chapter 5 of Williams [166] and only a brief description will be provided here. The original In-Au metallic wafer bonding was done by preparing the MBE-grown wafer with a Ti/Au layer sequence, and an n^+ GaAs receptor substrate wafer with a Pd/Ge/Pd/In/Au layer sequence (with the corresponding layer thicknesses as indicated in Fig. 4-6). The purpose of the Pd/Ge/Pd multilayer was to improve the electrical contact to the receptor substrate [162]. The topmost gold layer on the receptor substrate was included to minimize indium oxidation, which happens when the In surface is left open to the ambient atmosphere. Wafer pieces of about 1 cm^2 size were cleaved, aligned, and bonded on a hot plate at 250 °C for 10 min while pressure was applied to the stack manually. In this process, bonding takes place above the melting point of In (156.6 °C) as it wets the surface to fill in any crevices, and then diffuses into the gold layer to reactively form a variety of In-Au alloys [163, 102]. By careful choice of layer thicknesses, all the Indium is consumed, and the bonding layer remains robust up to the melting points of the In-Au alloys ($\sim$ 450 °C).

After the wafer to wafer bonding, the metal-metal waveguide fabrication proceeds with the following steps, which have remain standard even as the metal-metal bonding techniques have evolved during the course of this work. The GaAs substrate on the MBE-grown wafer is mechanically lapped down to a total thickness of $\sim$ 50 μm.

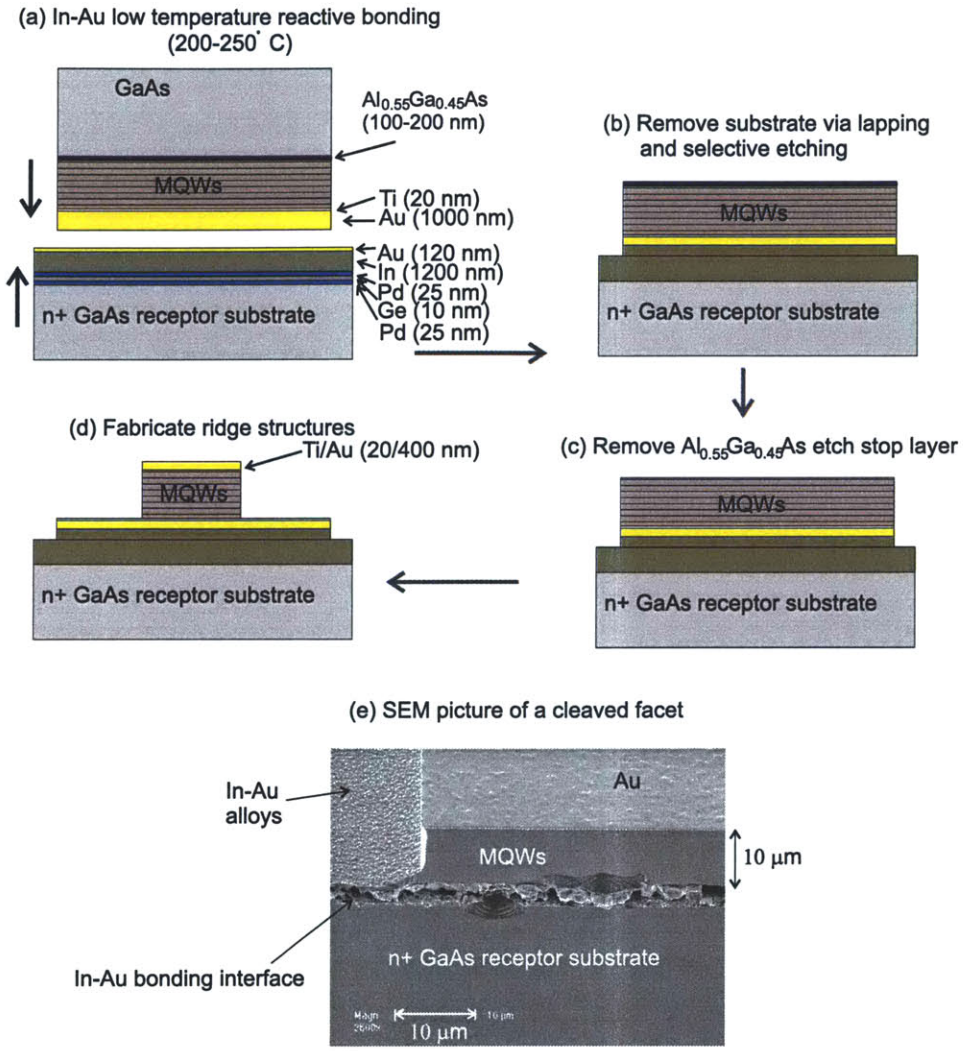


Figure 4-6: Schematic showing the fabrication process for terahertz QCLs with metal-metal waveguides based on a In-Au reactive bonding method. This was the process used for realizing the first terahertz QCL with metal-metal waveguides [170]. An SEM picture of a cleaved facet from a Fabry-Pérot ridge waveguide is also shown.

Afterwards, it is chemically etched in a $NH_4OH : H_2O_2$ 1:19 solution (etch rate of $1-3$ μm/min). This is a selective etchant that stops at the $Al_{0.55}Ga_{0.45}As$ layer, which is subsequently removed in HF acid. In the next step, the highly-doped top contact layer may or may not be completely removed (using a $NH_4OH : H_2O_2 : H_2O$ 10:6:480 etchant with an etch rate of $0.2 - 0.3$ μm/min), as was discussed in section 3.3.3. Ti/Au contacts are then evaporated using a lift-off process and then used as self-aligned etch masks to define ridge structures of various widths. Electron cyclotron resonance reactive-ion etching (RIE) in a $BCl_3 : N_2$ gaseous plasma is then used to etch down to the underlying metal. The processed wafer is then mechanically lapped again to thin the substrate of the receptor wafer, which improves heat-sinking during cw operation. After lapping the wafer down to a final thickness of $\sim 150 - 250$ μm, Ti/Au is deposited on its backside and the devices are ready to be cleaved into desired lengths to obtain Fabry-Pérot cavities.

Terahertz QCLs fabricated with MM waveguides using the layer sequence outlined in Fig. 4-6(a) achieved a value of $T_{\text{max,pul}} \sim 137$ K with the FL178C-M1 design [171] (Fig. 3-12b). This was a maximum temperature record for terahertz QCLs in pulsed operation that stood for over a year and a half from the time of its demonstration. However, lasing in those devices would cease after the first $10-30$ μs of pulse duration, and thus operation above a duty-cycle of few percent was never obtained. This was mainly attributed to poor thermal conduction at the In-Au bonding layer interface. The electrical properties of the bonding layer seemed to be much less affected since the lasers with the MM waveguides led to a significantly better temperature performance as compared to those with the SISP waveguides ($T_{\text{max,pul}} \sim 92$ K).

The thermal properties of the In-Au bonding layer improved dramatically with a minor modification to the layer sequence, as shown in Fig. 4-7(a). In this case, the topmost Au layer on the receptor wafer stack (Fig. 4-6a) was omitted. Although this layer was intended to minimize In oxidation, it was found to reactively consume much of the In prior to bonding, whereby, the resulting In-Au alloys had high melting points ($\gtrsim 450\,^{\circ}$C) and prevented the remaining In from wetting the interface uniformly during the wafer bonding. This led to the incorporation of voids at the bonding interface,

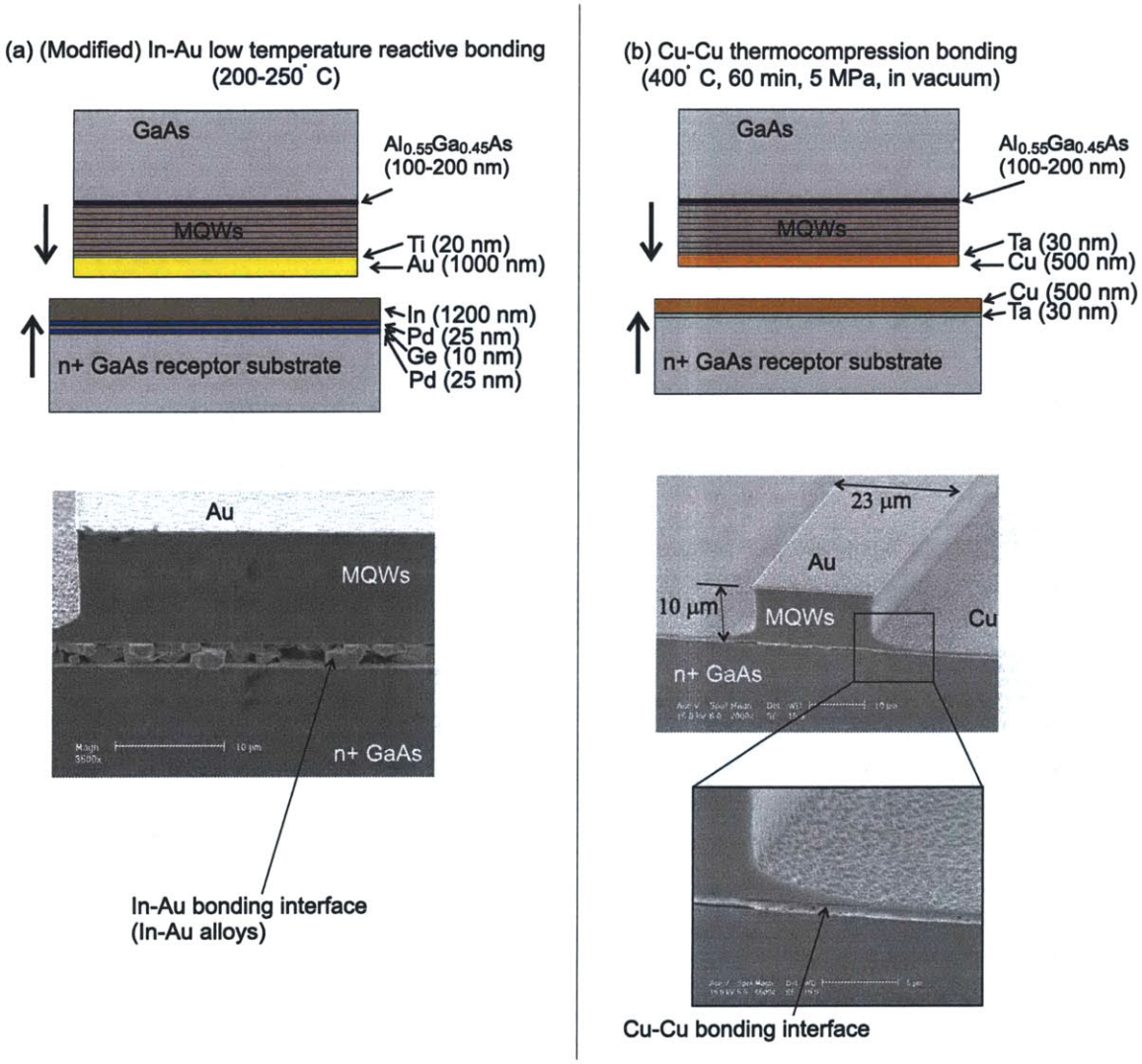


Figure 4-7: Layer sequences for (a) a modified In-Au reactive wafer bonding process [94], and (b) a Cu-Cu thermocompression wafer bonding process [174]. SEM pictures of cleaved facets from Fabry-Pérot ridge waveguides fabricated with each process are also shown. The height of the waveguides remains the same as in Fig. 4-6(e).

which possibly caused the large thermal resistance to prevent cw operation. Based on the modified layer sequence of Fig. 4-7(a), cw performance above the technologically important liquid nitrogen temperature of 77 K was achieved for the first time for terahertz QCLs [94] with the FL177C-M5 design (Fig. 3-9, Fig. 3-13a). The FL177C-M5 QCLs obtained a value of $T_{\mathrm{max,cw}} \sim 93$ K even though their $T_{\mathrm{max,pul}} \sim 134$ K was similar to the value of $T_{\mathrm{max,pul}}$ for the FL178C-M1 lasers, which had failed to lase in cw mode originally[1].

The thermal quality of the In-Au bonding layer depends largely on the metallurgy, and careful control of layer thicknesses as well as their composition is needed to prevent the formation of voids in the bonding layer. The In-Au bonding results were generally found to be less repeatable from one processing run to another. Moreover, void formation was present even in the devices fabricated with the modified technique, as can be seen from a SEM picture in Fig. 4-7(a). Therefore, an alternative wafer-bonding technique based on Cu-Cu thermocompression bonding was investigated [38], which has the potential for allowing improved cw operation due to the high thermal conductivity of copper. In this method, both the device and the receptor wafers are prepared by evaporating Ta/Cu layers as shown in Fig. 4-7(b). The Ta layer serves as an adhesion layer and prevents Cu diffusion into the GaAs/AlGaAs epitaxial layers [37]. Compared to the In-Au reactive bonding method, which can be performed by hand on a hot plate, the Cu-Cu method is considerably more demanding in terms of the high pressures and temperatures required and is more sensitive to particulate surface contamination. For bonding, an EV Group 501 wafer bonder machine has been used under vacuum at 400 °C for 60 min at a pressure of approximately 5 MPa. Following cooling, the devices are typically annealed for 30 min at 400 °C in an N_2 atmosphere. Devices that undergo the post-bond anneal display noticeable fewer stress cracks and defects in the epitaxial layer, which is consistent with previously observed strain relaxation in the copper layer [38].

[1]The FL178C-M1 devices were re-fabricated with the modified In-Au bonding technique and cw performance was eventually obtained with a value of $T_{\mathrm{max,cw}} \sim 67$ K. This value is lower than that for the FL177C-M5 lasers due to larger current densities in FL178C-M1 devices that cause extra electrical heating during cw operation.

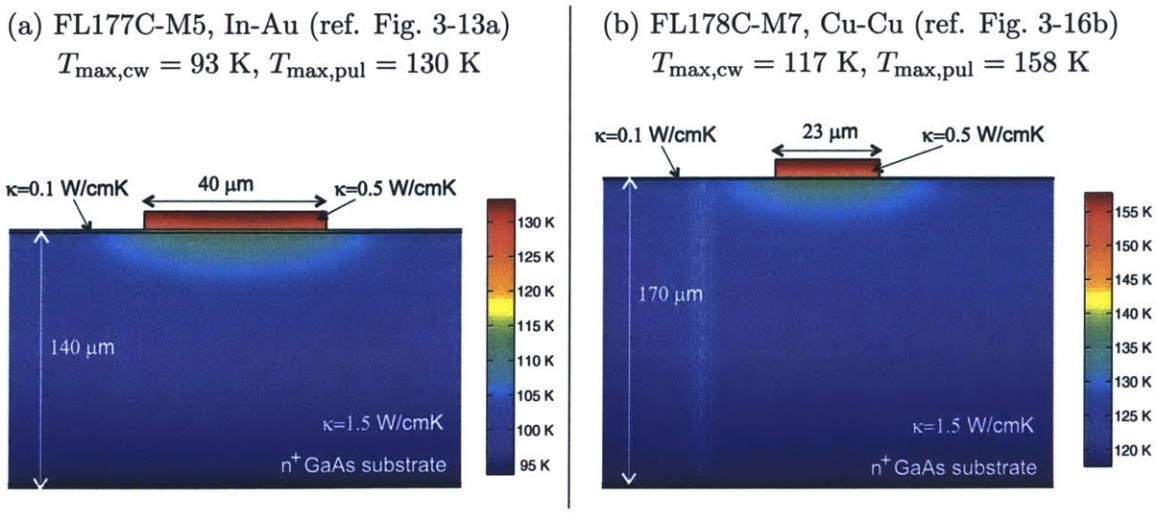


Figure 4-8: Two-dimensional heat flow calculation using a finite-element solver for the best performing In-Au and Cu-Cu lasers in cw operation. The metal-metal bonding layer is ~ 2 μm thick for FL177C-M5 and ~ 1 μm thick for FL178C-M7. The thermal conductivity of the MQW active region is modeled as $\kappa_{active} \sim 0.5$ W/cm·K to correspond with measurements of κ_{active} on a similar device [35], and the n^+ GaAs substrate is modeled with an approximate value of $\kappa \sim 1.5$ W/cm·K ($\kappa_{GaAs} \sim 2 - 1$ W/cm·K for $100 - 150$ K [23]). The κ for the bonding layer is arbitrarily chosen to match the active region temperature to $T_{max,pul}$ for a heat-sink temperature of $T_{max,cw}$.

In order to evaluate the thermal properties of the bonding interface, heat-flow out of the laser ridge was modeled by solving the heat-equation in two dimensions using a finite-element solver [2]. A calculation for the best performing In-Au and Cu-Cu lasers is shown in Fig. 4-8. The thermal conductivity of the bonding layer is selected to match the results of the calculation with the observed experimental results. A value of $\kappa \sim 0.1$ W/cm·K is obtained for both types of interfaces, which is considerably smaller than that for the respective metals ($\kappa_{In} \sim 1$ W/cm·K, $\kappa_{Cu} \sim 4$ W/cm·K). In spite of there being an uncertainty in the value κ_{active}, the calculations roughly suggest that the the bonding interface is not yet optimum for efficient heat-removal. By increasing the thermal conductivity of the bonding interface by a factor of ~ 10 the estimated value of $T_{max,cw}$ goes up by ~ 10 K for both devices, and by making the bonding interface perfectly conducting $T_{max,cw}$ goes up by additional ~ 5 K. The calculations in Fig. 4-8 are done for the best devices for each type of bonding; in general the experimental results from the QCLs fabricated with the Cu-Cu method have consistently shown improved, or similar results as compared to the best devices

fabricated by the In-Au method. This is mainly attributed to the Cu-Cu method being more consistent and robust as compared to the In-Au one.

Finally, it may be noted that during the initial development phase of the Cu-Cu thermocompression technique, a similar Au-Au bonding method was also tried. However, very thick ($\sim$ 1.1 μm) Au layers were deposited during that processing run, which caused cleaving problems after the devices had been processed. This is shown in the SEM pictures of two different cleaved facets from those devices (wafer FL177C-M5) in Fig. 4-9. Only a few of those devices were experimentally tested. Although the devices achieved lasing their transport characteristics were not optimum, which makes it is difficult to draw any conclusions from their temperature performance. Use of thinner Au layers would possibly solve the cleaving problem, however, due to the development of the relatively robust Cu-Cu bonding technique the Au-Au method has not yet been tried again.

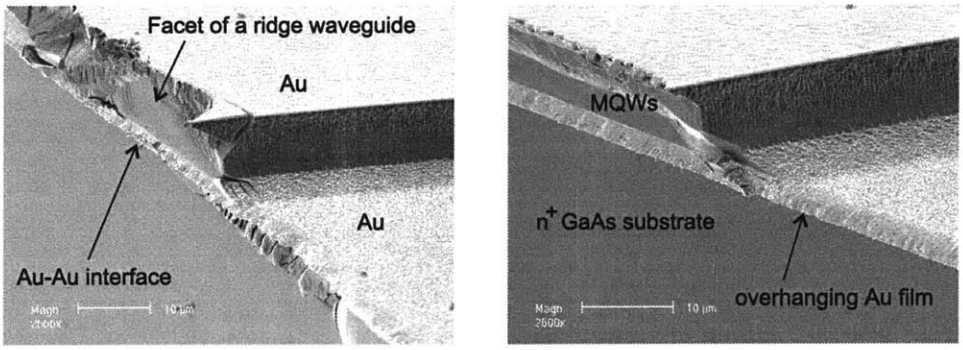


Figure 4-9: SEM pictures of badly cleaved devices fabricated using an Au-Au thermocompression bonding method, with a process similar to that of the Cu-Cu bonding method. The Ta/Cu (300/5000 Å) layer sequence in Fig. 4-7(b) was replaced by a Ti/Au (200/11300 Å) layer sequence. The bad cleaves are believed to be due to the use of very thick layers of Au, which did not break along the cleave but tore off uncontrollably at a random location due to Au being very malleable.

4.3 Electromagnetic characteristics of metal-metal waveguides due to subwavelength waveguiding

The MM waveguides for terahertz QCLs confine the terahertz fields in a subwavelength vertical dimension ($\lambda_{GaAs} = \lambda/n_{GaAs} \sim 17-83$ μm for $\nu \sim 5-1$ THz, whereas the thickness of the waveguides is typically $\sim$ 10 μm). This is made possible by some unique properties of the surface-plasmon electromagnetic modes as discussed in section 4.1.1. The subwavelength mode confinement causes the electromagnetic properties of these waveguides to be distinctively different than those of the conventional dielectric waveguides that are commonly used at shorter wavelengths. This section discusses some of these pertinent characteristics with the aid of finite-element calculations that are performed using a commercial software [2].

4.3.1 Strong mode confinement: lasing in microcavities

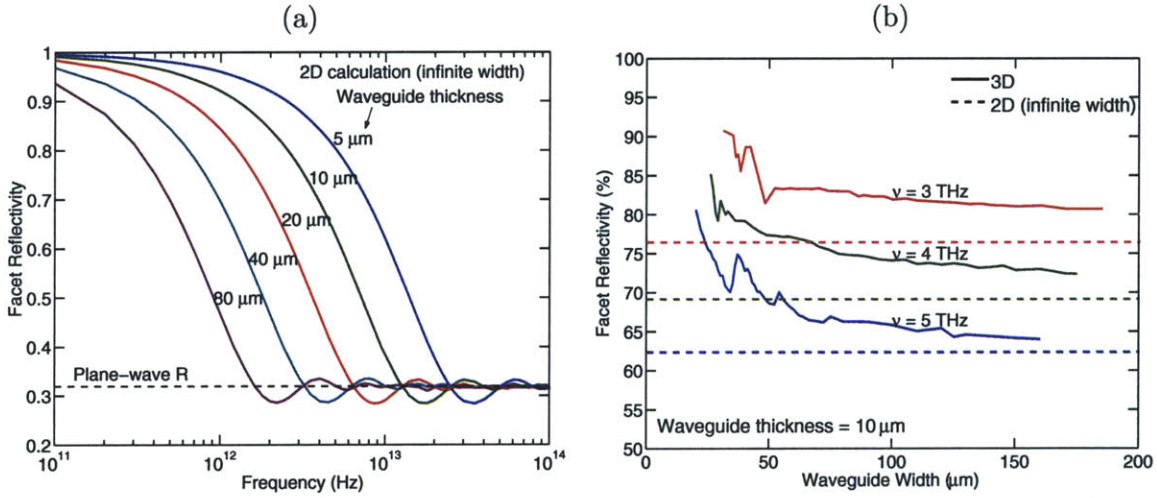


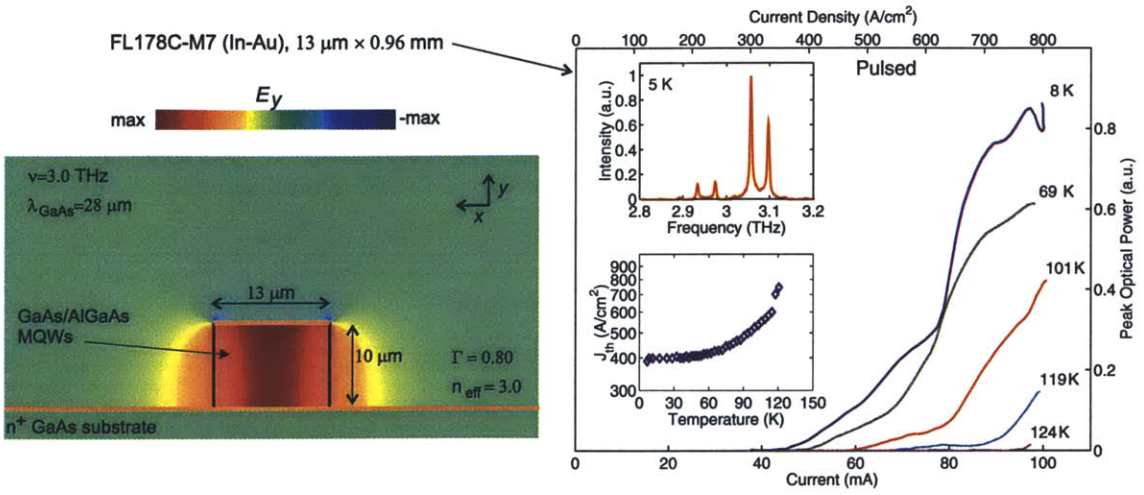
Figure 4-10: Finite-element facet-reflectivity calculations for Fabry-Pérot MM waveguides. Plot (a) is for waveguides of infinite width that have different thicknesses. The value of the facet-reflectivity corresponding to the normal incidence of a plane-wave at the air/GaAs boundary ($R = \left(\frac{n_{GaAs}-1}{n_{GaAs}+1}\right)^2 \sim 0.32$) is indicated by a dashed line. Plot (b) is for a 10 μm thick waveguide for different waveguide widths. The dashed lines are the corresponding values for a waveguide of infinite width. Both of these plots are reproduced from Ref. [88].

The computation of facet-reflectivities for the waveguided modes is a complex problem to which analytical closed form solutions are often not available. This problem has been treated with various different methods for the semiconductor dielectric waveguides at the optical frequencies, and for the metallic parallel-plate waveguides and microstrip transmission lines at the microwave frequencies. A selected few of the representative techniques could be found in Refs. [75, 69, 77, 79]. Due to a large mode mismatch between the waveguided mode and any free-space propagation modes, the electromagnetic mode in terahertz QCLs has negligible leakage from the facets of a Fabry-Pérot MM waveguide. Kohen *et al.* [88] calculated the facet-reflectivities of such cavities using a finite-element method. The results of their calculation deviate significantly from the value predicted by assuming plane-wave propagation as shown in Fig. 4-10[1]. Due to the high reflectivity values, the facets of the MM ridge waveguides need not be high-reflectivity coated, and lasing can be obtained in cavities with lengths down to the the order of the wavelength inside the semiconductor. Additionally, for Fabry-Pérot waveguides, the cavities can be made much narrower than the wavelength while still maintaining a near unity value of the mode confinement factor Γ. These characteristics allow lower lasing thresholds in terahertz QCLs fabricated with MM waveguides. Also a smaller volume of the active material lowers the energy requirements for electrical biasing, allows high temperature cw operation without the need of complicated techniques such as electroplating for heat removal, and allows multiple lasers to be fabricated in a relatively small area.

Experimental results from a FL178C-M7 (In-Au) laser that had a width of ~ 13 μm are shown in Fig. 4-11. A relatively large value of $\Gamma \sim 0.80$ is calculated for this laser even though the waveguide width is smaller than half of the wavelength inside the material. The temperature performance of this laser ($T_{\mathrm{max,pul}} \sim 124$ K) is not as good as that from the corresponding wider devices ($T_{\mathrm{max,pul}} \sim 148$ K, Fig. 3-20a2). This can be partially attributed to a large surface to volume ratio in this laser. The

[1]It may be noted that the facet-reflectivities for MM waveguides depend (weakly) on the outside environment of the laser, which determine the mode-matching conditions for the waveguided mode to the outer propagating modes. Hence, these calculations are strictly valid only for the geometries similar to the one described in Ref. [88].

(a) Subwavelength width Fabry-Pérot cavity laser operating at $\nu \sim 3.0$ THz



(b) Cylindrical microdisk cavity laser operating at $\nu \sim 2.0$ THz

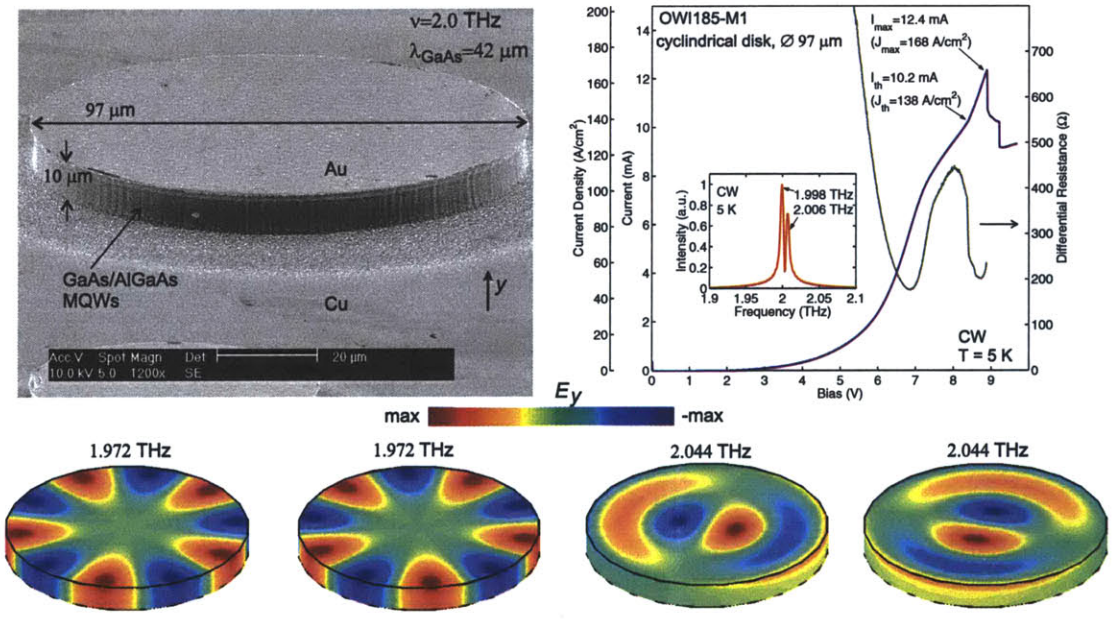


Figure 4-11: Experimental results and finite-element electromagnetic mode calculations for two different types of metal-metal waveguide terahertz QCLs with small cavity dimensions. The mode calculation in (a) is for two dimensions (i. e. infinite length waveguide), whereas the calculations in (b) are for three dimensions. The device in (a) is from the same processed die as the one in Fig. 3-20(a2), and the laser in (b) is from the same processed die as the one in Fig. 3-44(b). Hence, the value of $J_{\mathrm{max,5K}} - J_{\mathrm{th,5K}}$ for the "microcavity" lasers in this figure can be compared to the corresponding values for the larger devices referred above. *L-I*s for the laser in (b) were not measured hence no information is available about its temperature performance.

dry-etching causes surface roughness of the order of $\sim$ 10 nm in these waveguides, which, along with the possible surface damage induced by the physical nature of the etching could reduce gain in the quantum-well regions close to the sidewalls. Moreover, the value of $J_{\max}$ in the narrower ridges is typically lowered due the occurrence of an early NDR, possibly due to a non-uniform current density distribution along the waveguide height since the sidewalls of the cavities are not perfectly vertical.

Figure 4-11(b) shows the experimental results and eigenmode calculations for a cylindrical cavity ("microdisk") laser from the OWI185-M1 wafer (section 3.5.2). These circularly symmetric cavities can lase in many different kinds of modes including the so called "whispering-gallery modes" that propagate around the edges of the cavity [60]. The laser in Fig. 4-11(b) supports approximately $\text{Area}/\lambda^2_{\text{GaAs}} \approx 4$ modes close to its operating frequency of $\sim$ 2 THz. For this waveguide configuration, two doubly-degenerate pairs of modes, as shown in Fig. 4-11(b), exist in the frequency range of $1.9 - 2.1$ THz (which is the typical spectral range for OWI185-M1 lasers). It is likely that one of the pair of modes is represented by the two closely spaced modes in the lasing spectrum of this device. The value of $J_{\text{th,5K}} \sim 138$ A/cm^2 for this device is similar to that obtained from a larger device (Fig. 3-44b), which is due to the fact that the value of $J_{\text{th,5K}}$ is determined by the low-bias parasitic current channels in this design, as was discussed in section 3.5.2. The smaller value of $J_{\text{max,5K}}$ for the device in Fig. 4-11(b) as compared to the one in Fig. 3-44(b) is attributed to its larger surface to volume ratio, which can reduce the lasing range in current, as was mentioned in the preceding paragraph.

It is noteworthy that attempts to find the shortest/narrowest Fabry-Pérot MM waveguides to suppress lasing at T= 5 K from a lasing QCL active region did not succeed. This fact conforms to the strong mode confinement provided by the MM waveguides at terahertz frequencies, due to which the out-coupling losses in a cavity cannot be made large simply by reducing its dimensions. The fundamental mode for the MM waveguides with open ended sidewalls is TEM like, and thus, there is no cut-off width for such a mode to not exist at a given frequency. Waveguides narrower than 13 μm could not be tested since any attempts to bond wires on such ridges

destroyed them mechanically due to the relatively large size of the bonding wires (diameter $\gtrsim 25$ μm). To test shorter cavities, a FL177C-M5 laser with dimensions of 100 μm $\times$ 100 μm was tested at the liquid nitrogen temperature in cw mode, at which point it was still lasing. The best of the FL177C-M5 lasers tested (with a dimension of 40 μm $\times$ 1.35 mm) had a value of $T_{\text{max,cw}} \sim 93$ K (Fig. 3-13a), which suggests that the laser with a 100 μm long cavity did not suffer much of a performance degradation.

4.3.2 Effective mode index and frequency spacing of the Fabry-Pérot cavity modes

Corresponding to a given lateral mode of an infinite length waveguide, the resonance frequencies ν_{m} and the longitudinal mode spacing $\Delta\nu$ in a finite length Fabry-Pérot cavity are given by

$$\nu_m = \frac{mc}{n_{\text{eff}}\, 2L_{\text{cav}}} \tag{4.14}$$

$$\begin{aligned} \Delta\nu &= \nu_{m+1} - \nu_m \\ &= \frac{c}{n_{\text{eff}}\, 2L_{\text{cav}}} \end{aligned} \tag{4.15}$$

where, $n_{\text{eff}} = \mathcal{R}\text{e}\{\beta_z\}\, c/\omega$ is the effective mode index for an electromagnetic mode propagating in the z direction with the z dependence as $e^{i\beta_z z}$, L_{cav} is the length of the cavity, c is the speed of light in vacuum, and m is an integer that specifies the number of half-wavelengths along the length of the cavity for the mode at frequency ν_m (m may not be strictly an integer depending on the phase of the fields at the end-facets, however, this fact is irrelevant to the discussions that follow). Thus, the effective mode index can be estimated experimentally using the expression

$$n_{\text{eff}} = \frac{c}{\Delta\nu\, 2L_{\text{cav}}} \tag{4.16}$$

In Fabry-Pérot cavity terahertz QCLs with widths $\lesssim \lambda_{\text{GaAs}}$, a significant fraction of the electromagnetic mode may propagate outside the active region, thereby making the effective mode index n_{eff} much smaller than n_{GaAs}. This can be seen in Fig. 4-11(a)

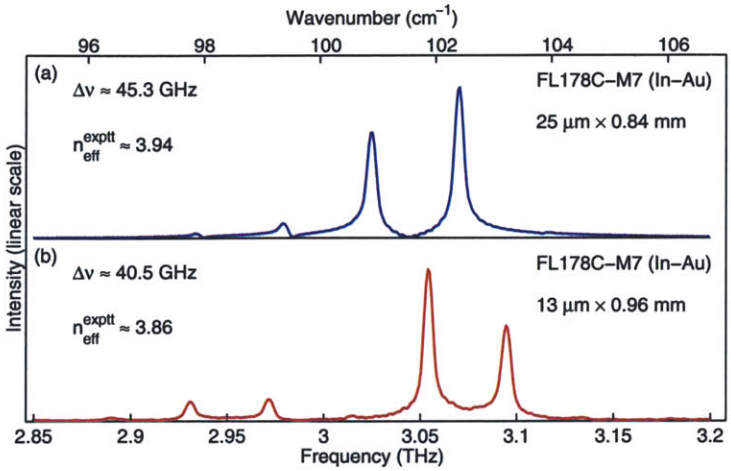


Figure 4-12: Estimation of the effective mode index for two different terahertz QCLs from the frequency spacing $\Delta\nu$ of the lasing modes according to the equation (4.16). From the uniformly spaced frequency spectrum and due to the narrow width of the waveguides, the modes can be attributed to the Fabry-Pérot cavity modes corresponding to the fundamental lateral mode (mode shape for the energy density is shown in Fig. 3-48 and for the vertical component of the electric-field is shown in Fig. 4-11a).

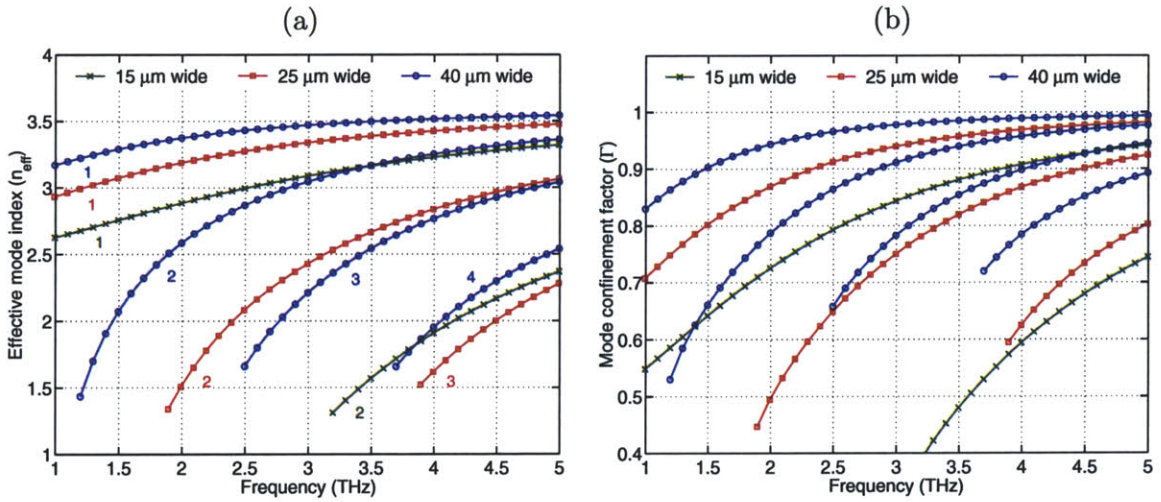


Figure 4-13: Effective mode index (n_{eff}) and mode confinement factor (Γ) for Fabry-Pérot MM waveguides of different widths, calculated for $\nu = 1 - 5$ THz using a finite-element mode solver. The ridge waveguides are taken to be 10 μm thick and with open ended sidewalls. For a given waveguide width, different curves are for different lateral modes where the lower valued curves (for both n_{eff} and Γ) are for progressively higher order lateral modes. Mode shapes for the different lateral modes in a MM ridge waveguide, albeit a one that has metal coated sidewalls, are shown in Fig. 4-30. For a n-th order lateral mode, the number n is indicated for each curve in plot (a), where $n = 1$ implies the fundamental mode. $n_{\text{GaAs}} = 3.6$ is assumed for these calculations.

where $n_{\text{eff}} \sim 3.0$ for a ridge width of 13 μm, and in Fig. 3-48 where $n_{\text{eff}} \sim 3.35$ for a ridge width of 25 μm, when calculated at a frequency of $\nu = 3.0$ THz ($\lambda_{\text{GaAs}} \sim 28$ μm) and for $n_{\text{GaAs}} = 3.6$. However, the values measured experimentally by using the expression in equation (4.16) are much higher, as is shown for two different MM waveguide lasers in Fig. 4-12[1]. To understand this behavior better, finite-element calculations for n_{eff} and Γ versus ν are shown in Fig. 4-13. It is noticeable that there is a significant frequency dispersion in the value of n_{eff}, especially for the narrower waveguides, which is due to the subwavelength width dimension in those waveguides. Hence, the equations (4.14) and (4.15) need to be modified as

$$\nu_m = \frac{mc}{n_{\text{eff}}(\nu_m)\, 2L_{\text{cav}}} \tag{4.17}$$

$$\begin{aligned} \Delta\nu &= \nu_{m+1} - \nu_m \\ &= \frac{c\left[(m+1)\frac{n_{\text{eff}}(\nu_m)}{n_{\text{eff}}(\nu_{m+1})} - m\right]}{n_{\text{eff}}(\nu_m)\, 2L_{\text{cav}}} \\ &\approx \frac{c\left[1 - \frac{m\Delta\nu}{n_{\text{eff}}(\nu_m)}\frac{dn_{\text{eff}}}{d\nu}\right]}{n_{\text{eff}}(\nu_m)\, 2L_{\text{cav}}} \end{aligned} \tag{4.18}$$

Consequently, the experimentally calculated value of the effective mode index appears "enhanced" according to the expression

$$\begin{aligned} n^*_{\text{eff}} &= \frac{c}{\Delta\nu\, 2L_{\text{cav}}} \\ &\approx \frac{n_{\text{eff}}}{\left[1 - \frac{m\Delta\nu}{n_{\text{eff}}}\frac{dn_{\text{eff}}}{d\nu}\right]} \end{aligned} \tag{4.19}$$

Due to a relatively large frequency dispersion in n_{eff}, the factor $\frac{m\Delta\nu}{n_{\text{eff}}}\frac{dn_{\text{eff}}}{d\nu}$ can be a few percent, which can modify the experimentally observed effective mode index $n^{\text{exptt}}_{\text{eff}}$ appreciably. Moreover, the enhancement factor depends on the length of the cavity (through m and $\Delta\nu$), and also on the width of the waveguide as well as the order of the lateral mode (through $\frac{dn_{\text{eff}}}{d\nu}$). Using equation (4.19), the values of $n^*_{\text{eff}} = 3.70$ and

[1]It is typical to observe multi-mode behavior in MM waveguides as shown in Fig. 4-12. However, MM waveguides with widths $\gtrsim 25$ μm typically excite higher order lateral modes also, the most common evidence of which is a non-uniformly spaced frequency mode spectrum for the laser. The higher order lateral modes are discussed further in section 4.5.3.

$n^*_{\text{eff}} = 3.58$ are calculated for the lasers in Fig. 4-12(a) and Fig. 4-12(b), respectively, which are closer to those experimentally observed ($n^{\text{exptt}}_{\text{eff}} = 3.94$ and $n^{\text{exptt}}_{\text{eff}} = 3.86$). The cause for the remaining discrepancy is not well understood; it is likely that the refractive index of the active region, which is assumed to be the same as the static value for GaAs ($n_{\text{GaAs}} = \sqrt{\epsilon_{\text{GaAs,DC}}} = 3.6$) in these calculations, has to be modified due to the vicinity of the operating frequencies to the *Reststrahlen* band in GaAs ($\hbar\omega \sim 33 - 36$ meV, $\nu \sim 8 - 9$ THz) [117, 126, 166].

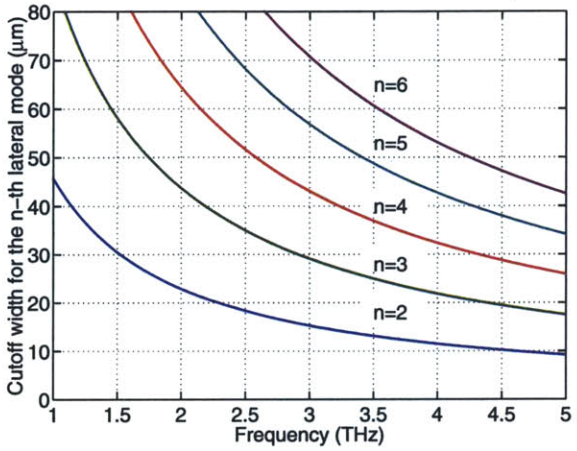


Figure 4-14: Plot of the maximum width of a rectangular MM waveguide that does not support the n-th lateral mode (calculated using equation 4.20 for $n_{\text{GaAs}} = 3.6$).

Finally, it may be mentioned that an *empirical* relation has been derived through finite-element simulations for the maximum width of a rectangular MM waveguide with open ended sidewalls that does not support the n-th order lateral mode (excluding the fundamental mode $n = 1$, which is always supported). At this width, the effective mode index for the n-th lateral mode approaches a value of 1 (the waveguide is assumed to be surrounded by vacuum/air). The empirical expression, as given below, is valid irrespective of the thickness of the waveguide (assuming only the fundamental transverse mode is supported in the vertical direction).

$$\text{Cutoff width for the } n\text{−th lateral mode} \approx \frac{\lambda_{\text{GaAs}}}{2}(n - 0.9) \tag{4.20}$$

where, $\lambda_{\text{GaAs}} \equiv \lambda/n_{\text{GaAs}}$ is the wavelength inside the semiconductor. This expression is plotted in Fig. 4-14 for the terahertz frequencies. For a rectangular waveguide with

metal on the sidewalls, where the n-th lateral mode is commonly referred to as the TE_{n0} mode, the expression in equation (4.20) becomes $\frac{\lambda_{GaAs}}{2}n$.

4.3.3 Radiation patterns from Fabry-Pérot cavity lasers

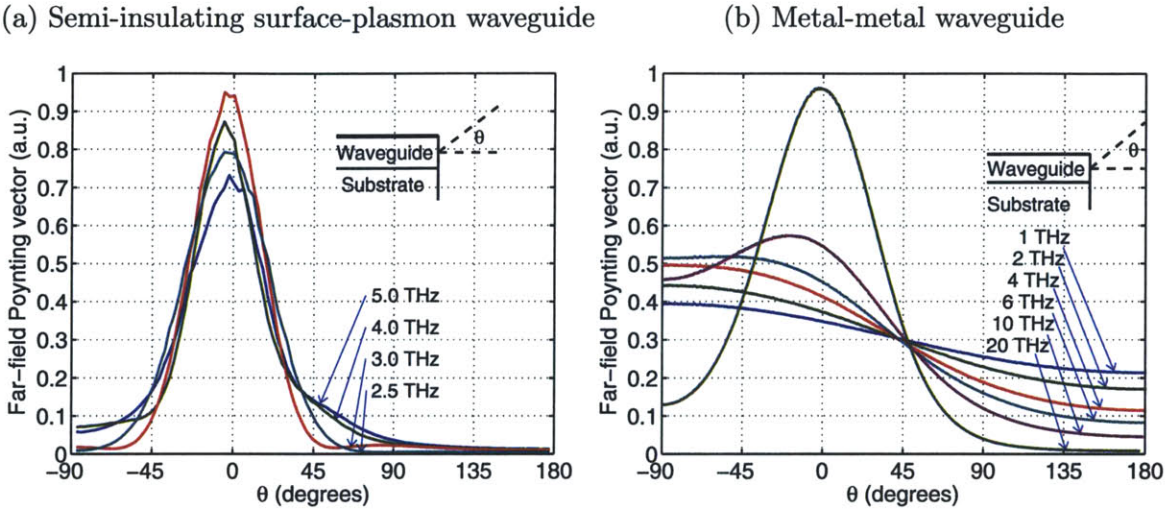
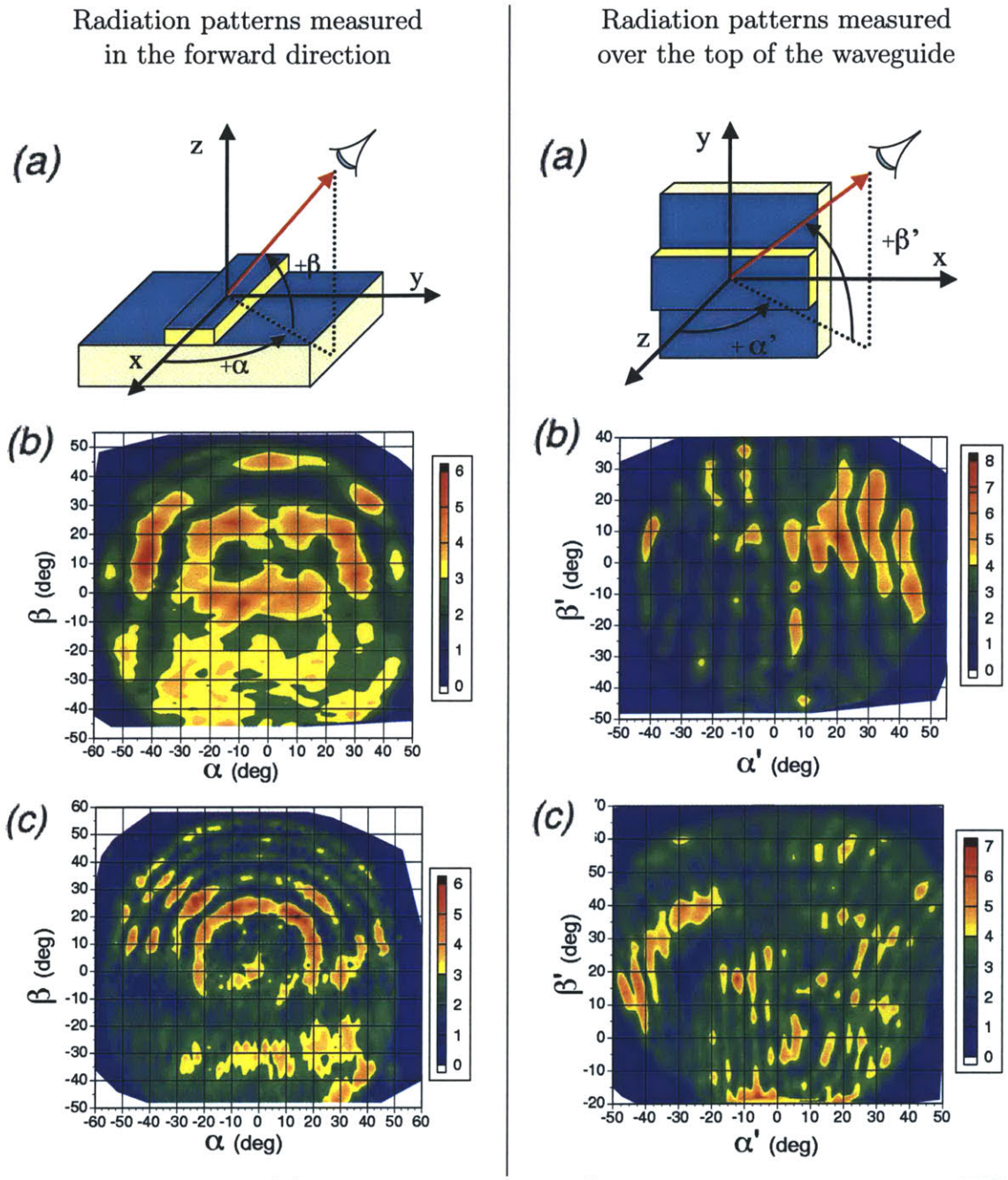


Figure 4-15: Two-dimensional (infinite waveguide width) finite-element calculations for the far-field radiation patterns from edge-emitting Fabry-Pérot cavity lasers with (a) the SISP, and (b) the MM waveguides. The thickness of the waveguides is 10 μm. Interference effects due to emission from the rear-facet are not included. These plots are reproduced from Ref. [87].

The aperture of an edge-emitting Fabry-Pérot MM waveguide has a subwavelength dimension at the terahertz frequencies. This makes the aperture highly reflective as was discussed in section 4.3.1. The small amount of radiation that does get out-coupled through the aperture is highly divergent. The radiation pattern of the emitted light is not simply a case of Fraunhofer diffraction, even though the problem might appear to be similar to that of a plane-wave uniformly illuminating an aperture (since the mode in a MM waveguide is TEM-like). The difference arises because the surface currents on the metal claddings and also excitation of the higher order lateral modes in the vicinity of the emitting aperture introduce additional boundary conditions. The problem at hand is somewhat similar to that of the metallic parallel-plate waveguides or the microstrip transmission lines (for example, see Ref. [181, 160]); however, none of those results are directly applicable to the frequency range and geometries of interest.



QCL device in plots (b): FL176C-M4-2 (Cu-Cu), 25 μm $\times$ 0.67 mm, $\nu \sim 2.92$ THz
QCL device in plots (c): FL176C-M4-2 (Cu-Cu), 25 μm $\times$ 1.50 mm, $\nu \sim 2.75$ THz

Figure 4-16: Experimentally measured far-field radiation patterns from Fabry-Pérot MM waveguide terahertz QCLs of two different cavity lengths [9]. The details of the experimental setup are described in the cited reference. The waveguides have open ended (cleaved) facets at both the ends. The thickness of the waveguides is 10 μm and that of the n^+ GaAs substrate is $\sim$ 210 μm. The device corresponding to plots (b) is the same as the one for Fig. 3-13(b), although the frequency of its lasing mode is different in that figure due to a different operating temperature.

Figure 4-15 shows finite-element simulation results obtained by Kohen [87] for the far-field radiation patterns of the SISP and the MM waveguides at some specific terahertz frequencies. Whereas the radiation pattern for the SISP waveguides are similar to that expected by Fourier optics (i. e. a spatial Fourier transform of the mode shape inside the waveguide), they are highly divergent for the MM waveguides, wherein significant amount of optical power is directed both towards the substrate ($\theta \rightarrow -90°$) and back above the top metal contact ($\theta > 90°$). This is believed to be due to a strong diffraction from the emitting aperture, and also due to strong coupling to the surface-plasmon electromagnetic modes of the type discussed in section 4.1.1, which can travel in vacuum at the interface of the top metal and the bottom doped substrate. Consequently, the already low optical power from the MM waveguides is reduced further since only a small fraction of it can be collected. This is why much higher optical power is measured from terahertz QCLs with SISP waveguides despite their smaller lasing range in current (section 3.6)[1].

Far-field radiation patterns from terahertz QCLs with MM waveguides have been experimentally measured. The results from the first of such measurements [9] for two nominally similar QCLs, but with different cavity lengths, are reproduced in Fig. 4-16. As is noticeable, significant amount of radiation is measured over the top of the waveguide, which agrees qualitatively with the calculation in Fig. 4-15(b). However, the measured radiation patterns show ring-like intensity maxima, which was not predicted by the calculation in Fig. 4-15(b). In the forward direction measurements, the observed rings are centered on the longitudinal axis of the QCL (i. e. $\alpha = \beta = 0°$). The rings are incomplete, appearing only in the upper part of the pattern ($\beta > 0°$). The angular distance between the first two maximum intensity rings for the shorter cavity laser ($\sim 20°$) is larger than that for the smaller cavity one ($\sim 9°$). For the radiation measured from the side of the waveguide, a set of approximately parallel lines of maximum intensity is observed. Just as for the other orientation, the angular distance between the lines is larger for the shorter cavity laser ($\sim 9°$) as compared to

[1]Nevertheless, the significantly better temperature performance of the MM waveguide terahertz QCLs makes them preferable for many applications of interest as will be discussed in section 4.4.

that for the longer cavity one ($\sim 5°$). Note that lasing in a higher order lateral mode can be ruled out in these lasers due to their narrow waveguide widths and a relatively low operating frequency, which makes the mode confinement factor small for such a mode (Fig. 4-13b). Moreover, a higher order lateral mode will not cause a ring-like radiation pattern by symmetry arguments.

The experimentally measured radiation patterns for these MM waveguide QCLs show strong intensity modulations where the directions of the extremes form cones around the longitudinal axis of the waveguide. In the simplest sense, these interference patterns could be explained as a result of two coherent point source situated at the two laser end-facets, radiating into the free space at a wavelength λ [9]. In that case, conc like intensity maxima will occur for directions such that

$$\cos(\theta_m) = m\frac{\lambda}{L} \tag{4.21}$$

where m is an integer, θ_m is the angle of elevation with respect to the x axis, and L is the length of the waveguide. Using the data on L and λ in Fig. 4-16, the angular distance between the first two maximum intensity rings for the short cavity and the long cavity laser are calculated to be 17° and 10°, respectively, in the forward direction ($\theta_m \sim 180°$), and 10° and 4°, respectively, when measured directly over the top of the waveguides ($\theta_m \sim 90°$). These simplistic estimates fit remarkably well to the angular distribution in the observed intensity pattern.

To investigate this further, the radiation pattern calculation of Fig. 4-15 was redone for some specific cases by using a different technique, which allowed calculation of the interference effects due to emission from the rear facet. Figure 4-17 shows the results of the simulation for Fabry-Pérot modes occurring approximately at $\nu \sim 4.35$ THz in a 1 mm long MM waveguide (plots a and b), and a 0.2 mm long SISP waveguide (plot c)[1]. These calculations were done with slightly improved boundary conditions, in that, whereas the absorbing far-field boundaries used for the

[1]The simulation was performed for only a small length of the SISP waveguide due to some convergence problems in the simulations for waveguides of practical lengths, mostly due to the fact that a large mesh is required to model the highly doped "plasmon" layer at the bottom of the active region. The metal is modeled as a perfect electrical conductor for both types of waveguides.

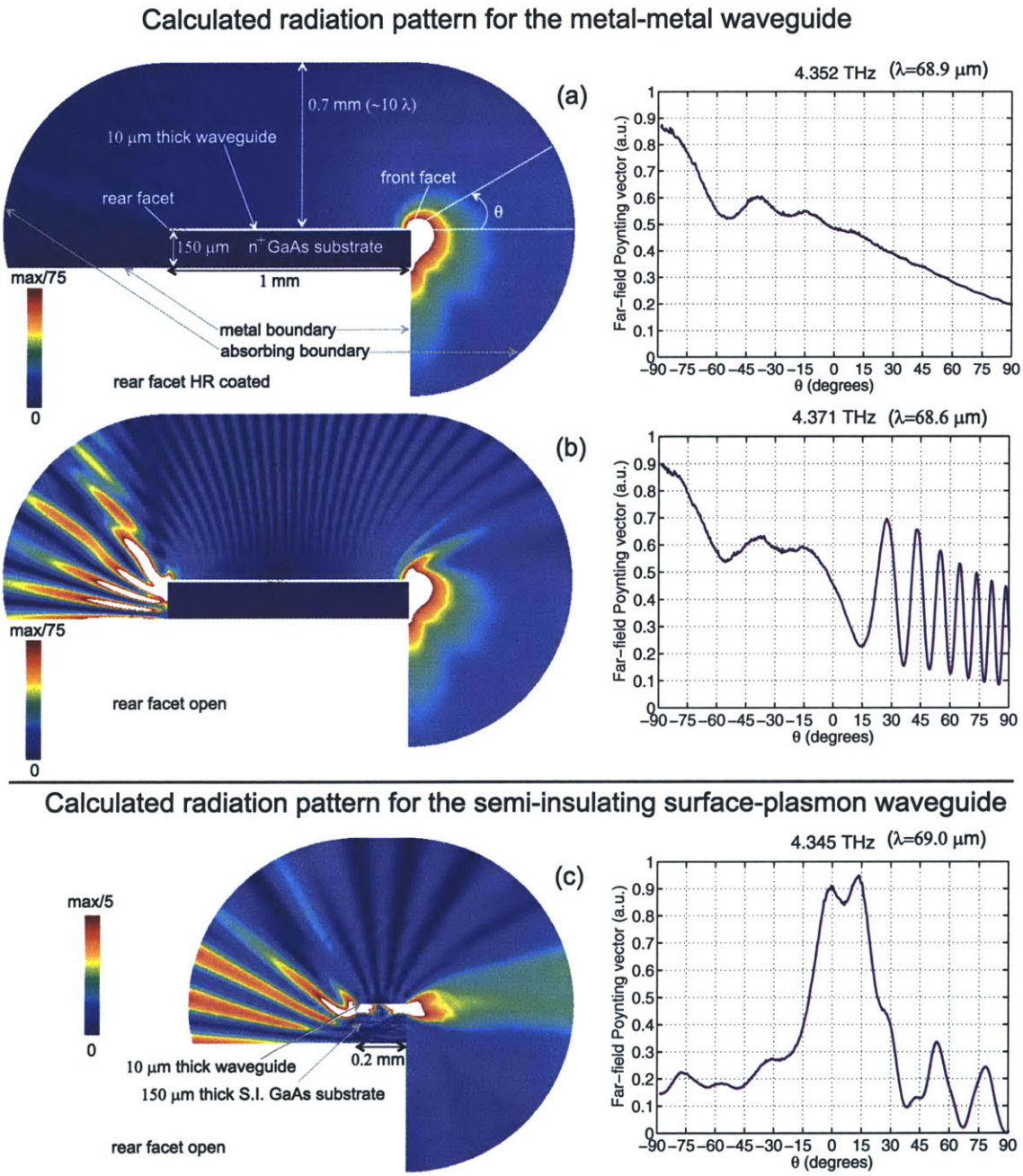


Figure 4-17: Two-dimensional (infinite waveguide width) finite-element calculations for the power flow around the waveguide, and the far-field radiation patterns in the frontal half-plane for the MM and the SISP waveguides, respectively. Unlike the calculations in Fig. 4-15, interference effects due to emission from the rear-facet are included. The plots on the left are for the magnitude of the Poynting vector. Although its direction is not indicated, it points approximately radially outwards from the facets on either side of the waveguide. Only a partial range of the magnitudes is plotted to better highlight the power flow; the regions with values over the maximum plotted range appear as white in color. The calculation of plot (c) is done for a SISP waveguide with a thin highly doped GaAs layer at the bottom of the active region with the layer parameters same as those for the FL183R-2 design (Fig. 3-48).

Far-field radiation patterns from adjacent Fabry-Pérot modes in a MM waveguide (as obtained from finite-element calculations similar to that in Fig. 4-17b)

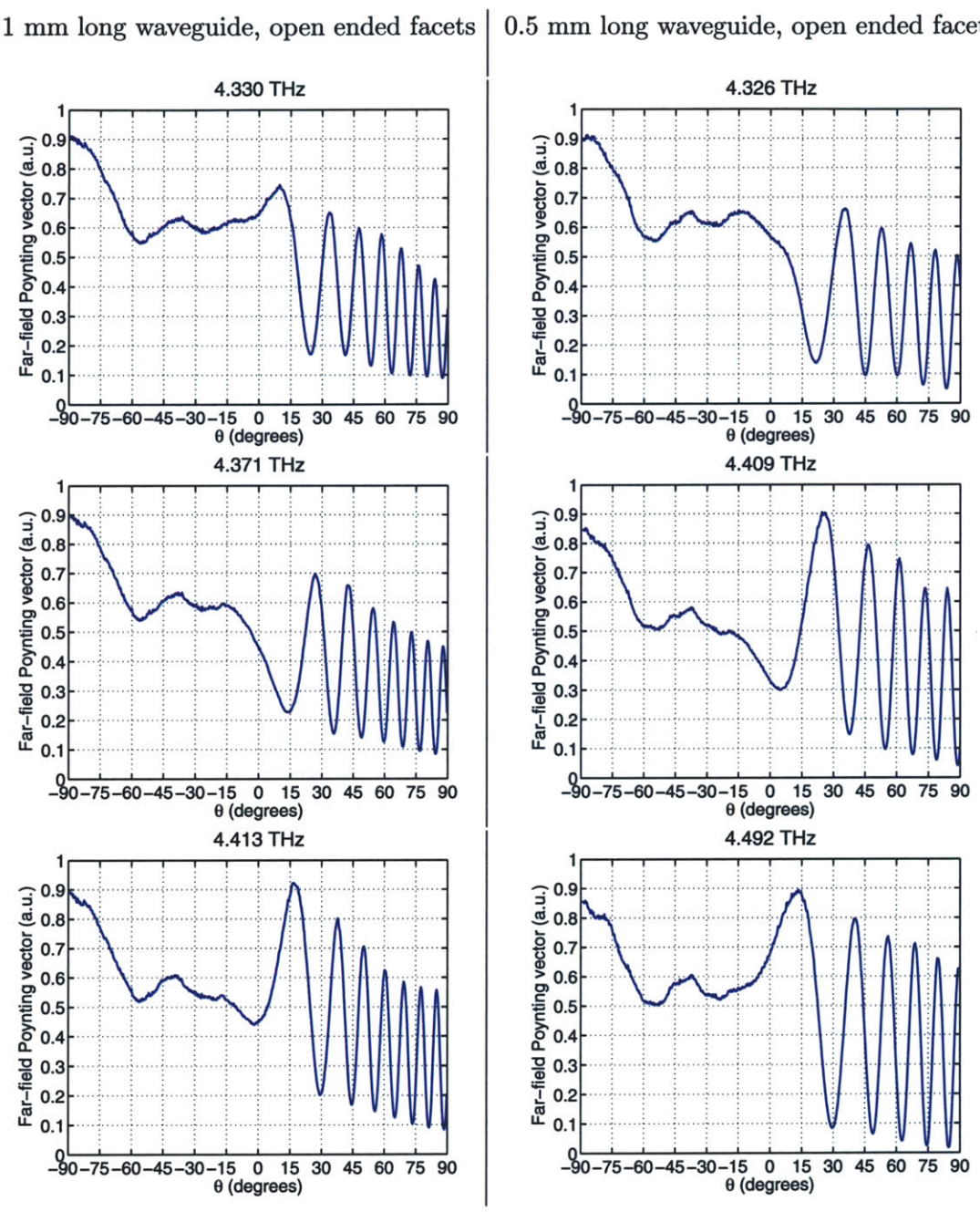


Figure 4-18: Far-field radiation patterns for three adjacent longitudinal modes in Fabry-Pérot MM waveguides of lengths 1 mm and 0.5 mm, respectively. The results are obtained from finite-element calculations similar to that in Fig. 4-17(b).

calculations in Fig. 4-15 could absorb only plane-waves, the absorbing boundaries used for calculations in Fig. 4-17 could attenuate arbitrary waveforms more effectively, thereby minimizing spurious reflections at those boundaries (more is explained about this in a following paragraph). The simulation in Fig. 4-17(a) is done with a HR coated rear-facet and no interference fringes occur in the radiation pattern. However, for the geometry in Fig. 4-17(b) where the rear facet is also emitting, a radiation pattern that is qualitatively similar to the experimentally one was obtained. In comparison, for the SISP waveguide, the relative strength of the interference fringes is small and most of the power is concentrated in the forward direction in spite of an open rear-facet[1].

Figure 4-18 shows far-field radiation patterns calculated for three adjacent Fabry-Pérot cavity modes from MM waveguides of length of 1 mm and 0.5 mm, respectively. The simulations predict sparsely spaced fringes for the shorter cavity laser, as is indeed observed in the experimental results of Fig. 4-16. Furthermore, the angular distribution of the fringes changes with the frequency of the cavity mode, which is because the ratio λ/L changes (as applied to the simple expression in equation 4.21, and the corresponding arguments therewith).

Note that the relative amplitude of the interference fringes in Fig. 4-18 is not any smaller for the longer waveguide as compared to the shorter waveguide. The origin of these fringes can be explained by the strong diffraction of light emitted from the end apertures of the waveguide, which then couples into the single surface-plasmon propagating modes at the metal-air interface on the top-metal. These modes have a negligible propagation loss as can be deduced from Fig. 4-4. Thus, for waveguides that are few milli-meters long, the relative intensity of the interference fringes is likely to depend very weakly on the length of the waveguides.

The calculations presented above are for infinite width waveguides and include effects due to emission from the end-facets only. However, for terahertz QCLs with MM waveguides the ridge widths are often of the order of the wavelength, in which

[1]It may be noted that weak interference fringes have indeed been observed experimentally in the radiation patterns for terahertz QCLs with SISP waveguides [73].

case significant evanescent mode exists outside the waveguide. Hence, for accurate calculation of the beam-pattern from such a waveguide the effect of the longitudinal field distribution inside the waveguide must be included, especially for waveguides of subwavelength widths. Orlova *et al.* [124] accounted for these effects by expressing the far-field radiation field in terms of equivalent current sources inside the volume of the cavity and obtained a close match to the experimental results in Fig. 4-16. An interesting phenomenon is predicted for a case when the effective mode index inside the the waveguide becomes close to unity, whereby emission in a very narrow beam is possible even though the emitting aperture is of subwavelength dimensions. However, this condition is difficult to meet in terahertz QCLs even with very narrow MM waveguides. For example, a 10 μm thick and 5 μm wide MM waveguide has a value of $n_{\text{eff}} \sim 2.25$ at $\nu \sim 3$ THz, which is still much greater than unity.

A discussion about the methods of the finite-element simulations that are used for these calculations is in order. The calculations in Fig. 4-15 were done using a "Harmonic-Propagation" solver [87], wherein a wave is launched using a fixed source boundary condition at one end of the waveguide towards the emitting facet. Any power that is reflected back towards the source is completely absorbed by the source boundary. This is relatively simple to do mathematically for a guided mode with a known propagation index [2]. The power radiated outside the waveguide is terminated into a boundary placed several wavelengths away from the facet. For the calculations in Fig. 4-15, that boundary was defined such that it would absorb any plane-waves impinging on it completely, whereas boundary reflections would arise for any other types of waves. The simulations in Fig. 4-17 and Fig. 4-18 were instead done using an "eigenfrequency" solver, which calculates the eigenmodes for the complete specified geometry. Out of the several modes that are obtained as a solution, a few of them correspond to the desired waveguided modes (i. e. the Fabry-Pérot resonances in the cavity). However, for this simulation to work, the boundaries have to be close to perfectly absorbing for any radiation, otherwise accurate results are not obtained. For the simulations corresponding to Fig. 4-17 and Fig. 4-18, an absorbing region of a thickness of $\sim 5\ \lambda$ was placed outside the boundaries that are shown in the plots

of Fig. 4-17. The dielectric constant in that region was linearly graded with distance by adding a complex component proportional to the radial distance in the absorbing boundary. This gradually attenuates any power that enters the absorbing region and minimizes any spurious reflections. Since this relatively simple model for the absorber region gave good results (which could be confirmed by the absence of and interference ripples in the electromagnetic fields close to the absorbing boundaries), more sophisticated "Perfectly Matched Layer" (PML) [81] implementations were not used for these two-dimensional simulations[1].

4.4 First-order edge-emitting distributed feedback (DFB) lasers

One of the more immediate application for terahertz QCLs is as local oscillators in heterodyne receiver systems for operation above 2 THz [59, 73], which are of great interest in astronomy and space science. In such systems, high optical power is not necessarily a requirement since only a few microwatts of optical power is needed to pump superconducting mixers. However, stable, narrow-linewidth, single-mode continuous-wave operation is required from the laser. For mid-infrared QCLs single mode operation is typically accomplished using a distributed feedback (DFB) grating, where index or gain/loss modulation is introduced via etching into the upper waveguide cladding or gain region, sometimes in combination with a metalized grating [62]. Recently, a corrugated lateral grating fabricated using high aspect-ratio dry-etching was demonstrated for a QCL operating at 11 μm [63]. The use of lateral distributed feedback allowed the grating to be defined lithographically and decoupled from the design of the vertical waveguiding structure and epitaxial growth.

This section describes the operation of first-order DFB terahertz QCLs fabricated using lateral corrugation in a metal-metal (MM) ridge waveguide [173]. The opera-

[1]Note that the the radiation pattern in Fig. 4-17(a) for a waveguide with HR coated rear-facet should have been identical to the corresponding result in Fig. 4-15 because of the similarities in the geometries used; however the slightly different results are due to the different boundary conditions used in these two calculations.

tion and design of such a DFB structure in MM waveguides is significantly different from the loss-coupled DFB QCLs that have been demonstrated in the semi-insulating surface-plasmon (SISP) waveguides [111, 11].

4.4.1 Grating design and experimental results

The FL176C-M4-2 design (section 3.3.1) was used as the active region for the DFB lasers presented in this section. DFB fabrication proceeded as follows. The MM waveguides were fabricated using the Cu-Cu thermocompression wafer bonding technique as described in section 4.2. Following wafer bonding and substrate removal, corrugated Ti/Au metal contacts (200/5000 Å) were defined on the 10-μm-thick epitaxial active region using contact lithography. The corrugated ridge waveguides were then defined by dry etching by using the top metallic contact as a self-aligned etch mask (section 4.2) and the receptor substrate was lapped to 180 μm to improve heat sinking.

Distributed feedback lasers are often designed with anti-reflection (AR) coatings, in order to prevent reflections from the cleaved facets from competing with the grating feedback, which can alter the relative lasing thresholds of various modes and cause lasing of unwanted modes. The precise effect is related to the phase $\sim 2\pi\ell/\Lambda$ of the effective complex facet-reflectivity, where ℓ is the facet length with respect to the grating period Λ (see Fig. 4-19b). The problem is potentially more severe in the terahertz, since AR coatings are more difficult to fabricate. Furthermore, due to the the tightly confined subwavelength mode, metal-metal waveguides have much larger cleaved facet-reflectivities ($R \sim 0.7 - 0.9$ for 10-μm-thick active region) as was described in section 4.3.1 (Fig. 4-10). However, because of the longer wavelength of terahertz radiation it is feasible to define the waveguide facet (and its reflectivity phase) in a controllable way using dry etching (see Fig. 4-19a).

For testing, the receptor substrate was cleaved approximately 100 μm away from the facet, and the device was indium soldered on a copper chip carrier in a vacuum cryostat, and wire bonds were made directly on top of the corrugated ridges. The ridges were 49-μm wide at the widest sections, with 10-μm lateral corrugations in

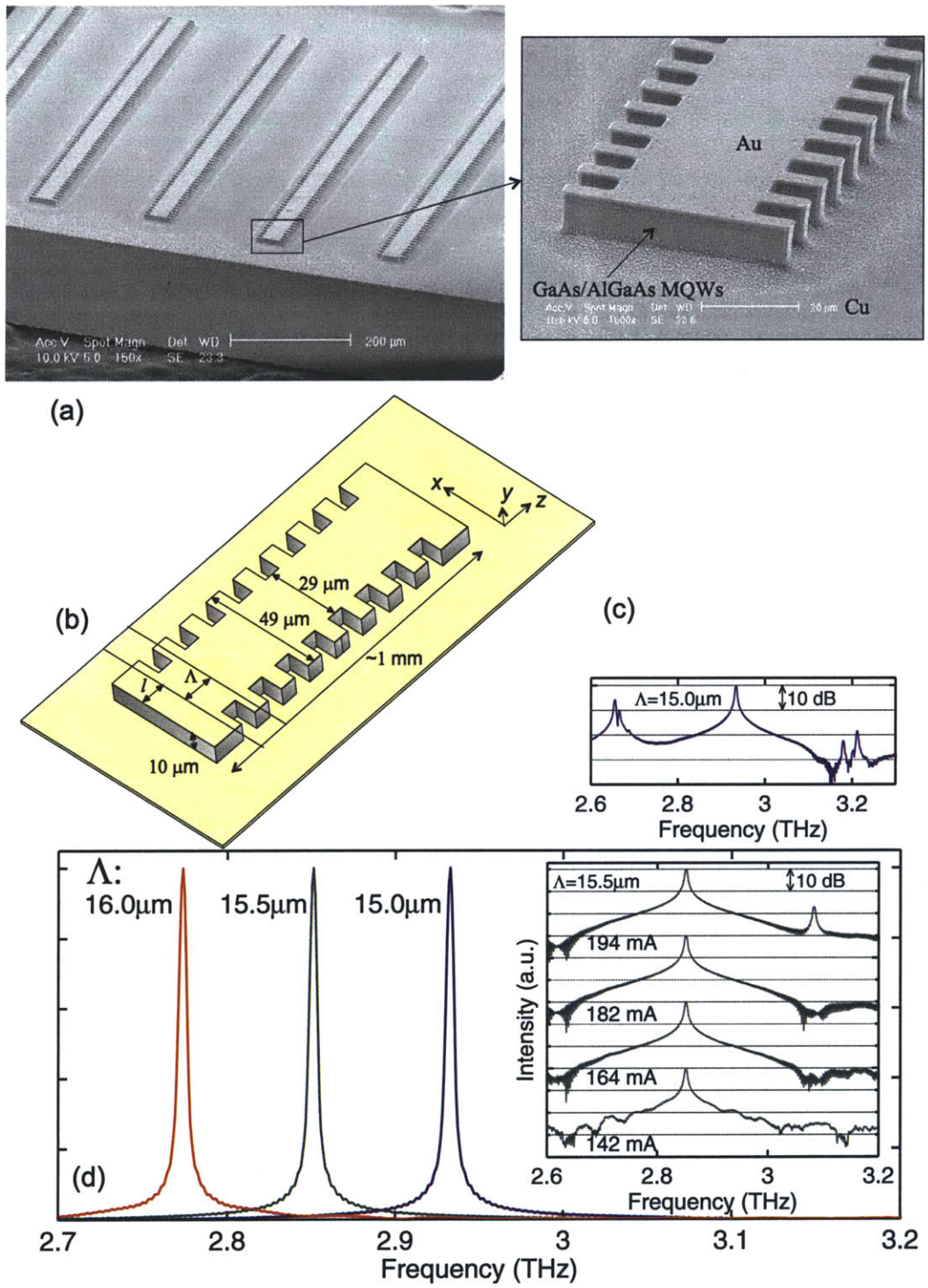


Figure 4-19: (a) Scanning electron micrograph of metal-metal corrugated ridge DFB with dry-etched facets of extension length ℓ. (b) Schematic showing part of the DFB structure near its end-facets. (c) Spectra from $\Lambda = 15.0$ μm, $\ell = -\Lambda/4$ device at high bias. (d) Spectra from DFB devices with several grating periods Λ, with $\ell = \Lambda/4$ taken at 8 K at moderate bias. The inset contains spectra taken at multiple biases from the $\Lambda = 15.5$ μm device.

either side. The corrugation duty cycle was nominally 50%, but due to limitations of the lithography, a duty cycle of approximately 44% was obtained. All ridges were roughly 1.0 mm long. Spectra were taken under a nitrogen gas purge with a Nicolet Fourier transform spectrometer (resolution = 3.75 GHz).

Devices were tested for several different facet lengths and several different grating periods. At low bias, single mode behavior was observed for most devices, but some of the best results were obtained from devices with a facet length of $\ell = \Lambda/4$. Continuous-wave spectra taken at moderate injection current at a heat-sink temperature of 8 K from three devices with Λ = 15.0, 15.5, and 16.0 μm are shown in Fig. 4-19(d). These devices lased at 2.935 THz, 2.853 THz, and 2.776 THz respectively, which scales well with the grating period. For comparison, Fabry-Pérot devices fabricated from the same wafer tended to lase between 2.7 and 2.95 THz (Table 3.1). The peak power emitted from the DFB devices was relatively low (100–300 μW per facet) in comparison to similar Fabry-Pérot devices, which typically emitted $\sim$1 mW.

As the bias was increased far beyond threshold, most devices lased in additional modes due to spatial hole burning. This can be seen in the spectra shown in the inset of Fig. 4-19(d) taken from the device with Λ = 15.5 μm, with a facet extension of ℓ = 2.9 μm. As is discussed below, the principal mode at 2.853 THz and the higher frequency mode at 3.085 THz can be associated with the edges of the grating stopband, which gives a bandgap of approximately Δf = 0.23 THz. Stopbands of similar magnitude (Δf = 0.21 THz for Λ = 15.0 μm, Δf = 0.25 THz for Λ = 16.0 μm) were observed for devices with different grating periods. The size of the stopband can be used to estimate the coupling coefficient κ used in the coupled-wave model [86] according to $|\kappa| = \pi n_{\text{eff}} \Delta f / c$, where the effective refractive index is $n_{\text{eff}} = \lambda_B / 2\Lambda \approx 3.26$ and λ_B is the Bragg wavelength taken at the center of the observed stopband. This expression gives a value of $|\kappa| \approx 79$ cm^{-1}, which indicated a much "stronger" grating than expected from the usual estimate of $\kappa \approx 2\Delta n_{\text{eff}} / \lambda_0 = 23$ cm^{-1}, where $\Delta n_{\text{eff}} = 0.12$ is the difference in effective mode indices in the wide and narrow regions of the waveguide (as determined using finite-element calculations). However, this perturbative calculation fails because the surface-plasmon modes are strongly coupled to the

metallic corrugation, and thus there is a large mode mismatch at each grating step, which causes a much stronger grating feedback. An additional high bias spectrum from a $\Lambda = 15.0$ μm, $\ell = -\Lambda/4$ device is plotted in Fig. 4-19(c), which supports modes separated by 0.55 THz. This large gain bandwidth suggests the potential for large tunability if the gain medium is placed in a tunable cavity.

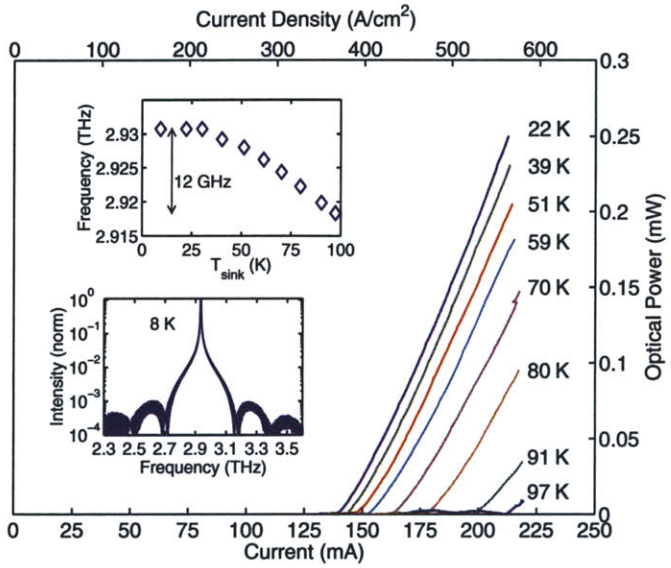


Figure 4-20: CW optical power versus current for device with $\Lambda = 15.0$ μm and $\ell = 6.75$ μm, which lased in a single mode over entire bias range (inset). The upper inset displays the frequency tuning achieved by changing the heat-sink temperature.

Single-mode behavior over the entire bias range was observed in a different device with $\Lambda = 15$ μm and $\ell = 6.75$ μm. The collected single-facet optical power is plotted versus current for several heat sink temperatures in Fig. 4-20. In cw operation, the threshold current density for this device was 370 A/cm^2 at 8 K, and it lased up to a heat sink temperature of 97 K, over which 12 GHz of temperature tuning was obtained. This performance, which was typical for all the grating devices tested, was as good as or better than comparable Fabry-Pérot devices (for example, see Fig. 3-13(b) and Fig. 3-15 for experimental results on the best Fabry-Pérot cavity lasers that were tested from this wafer). This is unlike the case for loss-coupled gratings where the waveguide losses for the DFB lasers are relatively higher, and thus the temperature performance is worse as compared to that of the Fabry-Pérot cavity lasers.

4.4.2 Finite-element calculations

A more complete picture of the DFB operation was obtained by performing full-wave 3-D finite-element (FEM) [2] simulations for the case of a $\Lambda = 15.5$ μm device with $\ell = 2.9$ μm, which corresponds to the device tested in Fig. 4-19(d). Because of memory limitations, only a device 465 μm in length could be simulated. Furthermore, the longitudinal and lateral symmetry of the problem was used to reduce the scale of the problem, so that in fact only one quarter of the structure was truly simulated. The simulation set-up was as follows. A "Harmonic-Propagation" solver was used, in which a 2-D mode could be excited at a given surface boundary of the 3-D geometry of the problem, and a self-consistent solution is calculated for the 3-D geometry to satisfy the excitation boundary condition. For the DFB geometry, an arbitrary (constant) electric-field was excited half-way along the length of the waveguide at the cross-section obtained by imagining a virtual cleave at that location. Afterwards, the frequency of excitation was swept and the output power at the end-facet of the waveguide was recorded at each frequency step. Note that the type and location of the parameter that is recorded is not critical. At some frequency close to a resonance frequency of a grating mode, the mode amplitude inside the waveguide increases with respect to the excitation amplitude, and correspondingly the value of recorded parameter also becomes large. This is because the excitation point is a node (zero) of the standing wave inside the waveguide at the resonance frequency. To find the gain threshold, a similar sweep was done by varying the gain inside the waveguide for all the resonance frequencies determined by the aforementioned method. The gain was added as an imaginary component to the dielectric constant of the waveguide. For waveguide gain close to the value of the threshold gain, the mode amplitude in the waveguide increases sharply, and hence the threshold gain value can be determined for that particular resonance mode[1]. All of the aforementioned steps were repeated with a magnetic field excitation instead of an electric-field one to calculate resonance

[1]Note that the mode amplitude reduces again for a gain higher than the threshold gain, since the excitation point is a node only for the resonance condition, which happens when the gain in the waveguide equals the threshold gain of the grating mode exactly.

modes corresponding to the second longitudinal symmetry condition[2].

The resonance frequencies and threshold material gains $g_{\rm th}$ for the various DFB modes thus calculated are plotted in Fig. 4-21. The metal was treated as a perfect conductor and the semiconductor was undoped, so $g_{\rm th}$ only reflects radiative losses, and not free-carrier waveguide losses (expected to be of the order of 10 $\rm cm^{-1}$). Part (a) of the figure displays those modes which are laterally symmetric, and part (b) displays the laterally antisymmetric modes. Coupled-mode analysis suggests that the results for a longer device (L = 1 mm as tested) should be qualitatively similar to the simulated results, and indeed there is good agreement between the observed lasing modes (2.85 THz and 3.09 THz), and the lower and upper band-edge symmetric modes at 2.87 THz ($g_{\rm th}$ = 0.1 $\rm cm^{-1}$), and 3.11 THz ($g_{\rm th}$ = 1.9 $\rm cm^{-1}$) respectively. The two modes in the center of the stopband have large $g_{\rm th}$ values because they are localized near the facets. The low-threshold modes with odd lateral symmetry (Fig. 4-21b) do not lase for devices with $\Lambda = 15.5$ μm because they are too far away from the gain peak; however they were observed in spectra from some $\Lambda = 15.0$ μm devices (see Fig. 4-19c). Although the $\ell = \Lambda/4$ facet location is nominally in phase with the grating, the effective facet-reflectivity is somewhat out of phase with that of the grating, due to the fact that the facet was fabricated shorter (2.9 μm) than its designed length, and also because of a residual phase shift at the facet of the metal-metal waveguide. These effects break the symmetry across the bandgap, whereby the upper modes couple more strongly to the facets and thus have higher $g_{\rm th}$, while the modes on the lower bandgap edge maintain lower thresholds.

The vertical component of the electric-field (which is the predominant electric-field component in the waveguide) for the modes with odd lateral symmetry can be written as

$$E_y(x,y,z) \;=\; c_1(z)\,E_{y,1}(x,y) + c_3(z)\,E_{y,3}(x,y) \tag{4.22}$$

where, $E_{y,1}(x,y)$ and $E_{y,3}(x,y)$ are the mode-shapes for the first (fundamental) and

[2]For the resonant-modes of longitudinally symmetric waveguides, the amplitude of either the electric-field, or the magnetic-field is zero for the standing wave established inside the waveguide.

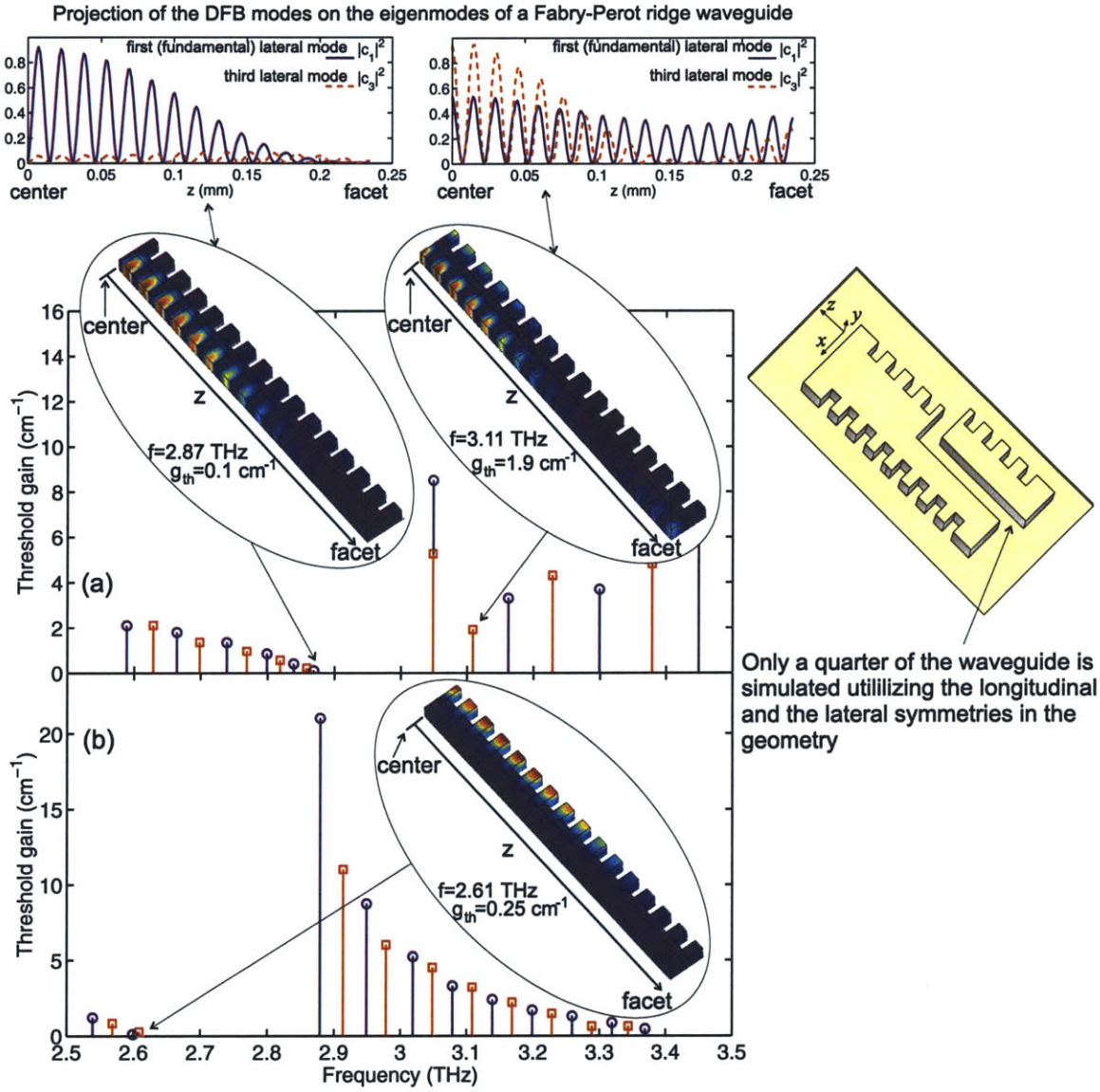


Figure 4-21: Modal threshold gains calculated using a 3-D FEM solver for a $\Lambda = 15.5$ μm, 465-μm-long DFB structure with $\ell = 2.9$ μm. Modes with even and odd lateral symmetry are shown in (a) and (b) respectively, and modes with even and odd longitudinal symmetry are marked with squares and circles respectively. Also shown are the plots for the magnitude squared vertical component of the electric field ($|E_y|^2$) for the band edge modes (note that only one quarter of the waveguide is plotted). The two plots on the top corresponding to the band-edge modes of plot (a) indicate the contribution of the different order lateral modes to the standing wave established in the DFB waveguide (equation 4.22). Note, however, that due to the odd lateral symmetry of the mode, there is no coupling to the second lateral lateral mode since it is even. The mode content of the band-edge mode corresponding to plot (b) is primarily due to the second-lateral mode, since the fourth-lateral mode is not supported by the waveguide.

the third order lateral modes, respectively, in a Fabry-Pérot ridge waveguide of the same width as the wider part of the corrugated DFB waveguide. Examination of the modal field patterns in Fig. 4-21(a) reveals that a considerable higher order mode content can exist in the DFB modes, which is due to the reflections at each grating step that causes coupling between the different lateral modes. The modal field patterns also reveal that the band edge modes are highly localized in the center of the waveguide, which accounts for their low measured power levels, and is consistent with the understanding that non-uniform longitudinal field profiles occur for κL products larger than unity (L is the cavity length) [86]. For this simulated case $\kappa L \sim 4$; for the experimental case $L = 1$ mm and $\kappa L \sim 8$, which suggests that the longitudinal mode profiles are even more localized than the simulated case, and that these devices might be vulnerable to spatial hole burning. However, given the similarity of the longitudinal mode envelopes of the upper and lower stopband modes, the appearance of the upper mode is more likely due to spatial hole burning on the scale of the wavelength, since $\Lambda/2$ is more than an order of magnitude larger than the diffusion length associated with the intersubband population inversion ($\sim$200 nm). Finally, even though the experimental results are in good agreement with the simulations, the possibility of localized modes that can form around unintentional grating disorder, such as the wire bond attachment points, cannot be completely excluded.

The ability to easily customize the grating corrugation and facet phase conditions via lithography allows the engineering of the relative gain thresholds and longitudinal mode profiles. For example, longitudinal field uniformity and output power levels could be improved by reducing the corrugation depth (or device length) to reduce the κL product. Although not explored here, power could be coupled preferentially out of a single facet by introducing controlled phase shifts into the length of the DFB or independently adjusting the facet phases.

4.5 Second-order surface-emitting distributed feedback lasers

4.5.1 Surface-emitting DFB advantage for terahertz QCLs with metal-metal waveguides

Some of the most important applications for terahertz QCLs are expected to be in heterodyne receivers as local oscillators, or as sources for imaging and spectroscopy [59, 73, 101]. Whereas stable continuous-wave (cw) single-mode emission is required for local oscillators, several milliwatts of optical power in a narrow beam is desired to obtain large dynamic range and high-spatial resolution in an imaging system. Moreover, stringent frequency requirements may exist for both spectroscopy and imaging, such as to target a particular spectroscopic feature, or to avoid certain atmospheric water absorption spectral bands for imaging over a distance of few tens of meters [100].

The best terahertz QCLs in terms of high temperature operation have been demonstrated in the metal-metal (MM) waveguides as opposed to the semi-insulating surface-plasmon (SISP) waveguides, which was affirmed through experimental results in section 3.6. Besides providing a lower waveguide loss α_w (section 3.6.1), the MM waveguides provide strong mode confinement for the terahertz fields even in cavities with subwavelength dimensions (section 4.3). However, their efficient waveguiding is accompanied by highly divergent radiation patterns, low out-coupling of the optical power, and multi-mode behavior due to easy excitation of the higher order lateral modes even for very narrow (< 50 μm) Fabry-Pérot cavities.

Whereas the first-order edge-emitting DFB lasers, which were described in section 4.4, can achieve single-mode emission without any degradation in the temperature performance, the problem of divergent radiation patterns remains and the out-coupling of the optical power is further reduced in comparison to the already low values of the Fabry-Pérot cavity lasers.

The objective of developing second-order surface-emitting DFB lasers is to preserve the low-loss advantage of the metal-metal waveguide structures, but increase the

output power levels and improve the beam patterns along with single-mode operation. For this purpose, it is instructive to analyze the effect of two waveguide parameters, the waveguide loss α_w and the facet (mirror) loss α_m, on the lasing threshold (and therefore the maximum operating temperature) and the output power levels. To a leading degree, the lasing threshold is proportional to $(\alpha_w + \alpha_m)/\Gamma$, where Γ is the mode confinement factor and ≈ 1 for the MM waveguides. On the other hand, the output power is proportional to the slope efficiency, which is proportional to the factor $\alpha_m/(\alpha_w + \alpha_m)$. Because of the high facet-reflectivities (section 4.3.1), the mirror loss α_m in the MM waveguides is small. Approximately, α_m is of the order of 1 cm^{-1} in comparison to the waveguide loss α_w that is of the order of 10 cm^{-1} (for a more detailed analysis, see section 3.6.1). This suggests that by increasing the facet loss, the output power levels can be increased proportionally, whereby increasing the lasing threshold only marginally as long as $\alpha_m < 5$ cm^{-1}. In microwave engineering, this increase of the facet loss (or the out-coupling efficiency) is usually done by using a tapered mode matching structure to adiabatically transform the low impedance of the microstrip transmission line to the free-space impedance. At the output of the mode matching structure, the cross section is much greater than the wavelength, resulting in a much tighter far-field beam pattern. In principle, such a technique could also be applied to the MM waveguide terahertz QCLs. However, because of the shorter wavelengths at terahertz frequencies, the fabrication and assembly of such mode-matching structures will be quite demanding.

An alternative method to increase the mirror loss is by using a second-order distributed feedback (DFB) grating to couple the laser beam out from the surface. The out-coupling loss α_m can be increased in a controllable way by choosing the grating duty cycle, phase of reflection at facets, and a central phase-shift in the grating. A large surface area for emission provides the possibility of an improved beam pattern in addition to higher optical power levels. Furthermore, the DFB nature of this structure allows robust single-mode operation and frequency selectivity, which are required for all sensing applications involving narrow target spectral features. High-power surface-emitting quantum-cascade lasers using second-order DFB gratings have

been well developed at mid-infrared frequencies [70, 128]. However, implementation of such gratings in MM waveguides at terahertz frequencies turned out to be much more challenging because of mode competition from strongly confined higher order lateral modes and high reflection from end-facets of the cavity (reflectivity $\sim$ 0.7–0.9, Fig. 4-10). At the mid-infrared frequencies, silicon-nitride that is used to cover the side walls for electrical isolation, strongly absorbs radiation [65]. This, in addition to the fact that higher order lateral modes have a greater overlap with the lossy doped cladding layers, effectively eliminated all such modes either by accident or by design. At terahertz frequencies, the absorption by the insulator (silicon-nitride or silicon-dioxide) is much weaker, and higher lateral modes are easily excited due to strong mode confinement in MM waveguides. As discussed in the following sections, new methods are needed to develop surface-emitting terahertz quantum-cascade lasers with the desired grating modes in such waveguides.

The following sections discuss design and fabrication techniques to eliminate the higher order lateral modes and to control the phase of reflection at the facets so that it does not interfere with DFB operation [95]. Control of the facet-phase also allows engineering of the mode-shape for the excited mode to minimize possibility of longitudinal spatial-hole burning that can cause multiple-mode behavior at high-bias conditions. In these devices, robust single-mode operation is demonstrated under all bias-conditions at frequencies over a range of $\sim$ 0.35 THz centered around 2.9 THz for a set of devices with various grating periods. A single-lobed far-field radiation pattern is achieved using a π phase-shift in center of the grating [105, 93]. In pulsed operation, the grating device lasing near the peak frequency of the gain spectrum lased up to 149 K, while the grating device operating away from the gain peak still lased up to 141 K. In comparison, multi-mode Fabry-Perot ridge structures on the same die lased up to 153 K, thus demonstrating a relatively small degradation in temperature performance from the implementation of surface-gratings. The measured cw power level, $\sim$ 6 mW, is $\sim$ 2 times higher than that measured from a single facet of edge-emitting lasers with a comparable area.

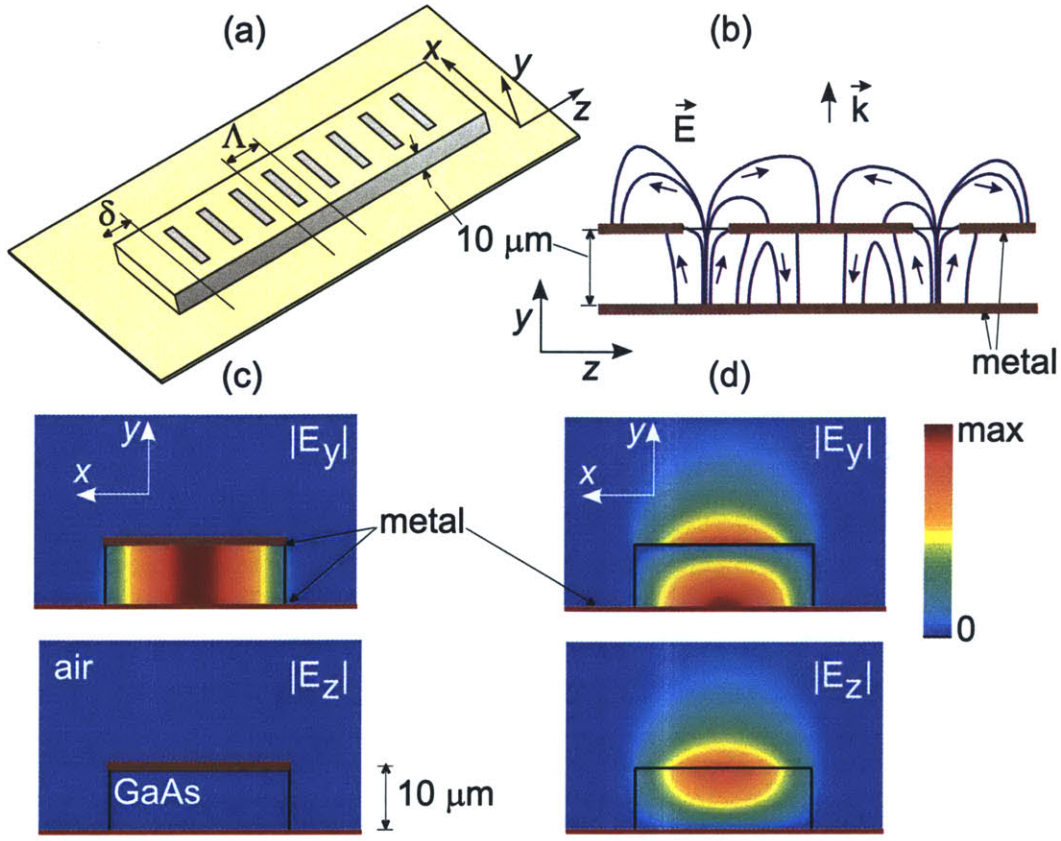


Figure 4-22: (a) Three dimensional schematic of the grating structure. (b) Electric-field lines for a grating mode showing the grating induced change in field polarization to achieve surface emission. (c),(d) Electric-field profiles for the fundamental propagating mode in an infinitely long, 100 μm wide waveguide with and without top-metal, respectively, at $\nu = 3$ THz. The corresponding propagation mode indices n_{eff} are 3.58 and 2.66 where n_{GaAs} is taken to be 3.6. The aspect ratio for the plotted geometry in the x and y directions is not to scale.

4.5.2 Grating design

While metallic gratings for surface emission have been demonstrated in SISP waveguides [41], the design, analysis, and operation of such gratings is considerably different for MM waveguides. Figure 4-22(a) shows the chosen scheme for implementing second order DFB in MM waveguides [93]. The 10 μm thick GaAs/$Al_{0.15}Ga_{0.85}As$ multiple-quantum well (MQW) gain region is sandwiched between metal on top and bottom. Due to intersubband polarization selection rule, in order to couple with the material gain, the electric field must be polarized normal to the MQW layers. Hence, for unpatterned Fabry-Perot ridge lasers electric-field in the mode is vertically polarized and only edge emission is possible. However, as shown in Fig. 4-22(b), field polarization could be "bent" by having apertures in the top-metal. The thin highly doped GaAs contact layer, which is often used at the top of the MQW stack to facilitate electrical contact, must be removed from the aperture regions so that it will not cause a significant cavity loss due to the presence of E_z field in apertures. For a second-order Bragg grating with grating-period Λ the same as the wavelength of propagating mode inside the semiconductor ($\approx \lambda_{GaAs} \sim 28$ μm at 3 THz), the grating acts like a phased-array antenna and the E_z fields in all the apertures are in phase to produce vertical emission. Additionally, such a grating provides a strong feedback into the cavity for DFB action.

The standard coupled-mode theory is no longer applicable for such a grating due to sub-wavelength mode confinement in the vertical direction. This can be seen from propagating mode solutions for structures with and without top-metal as shown in Fig. 4-22(c), and Fig. 4-22(d). There is a large difference in the mode-shapes for the two cases, which causes a strong mode-mismatch at each grating step. Thus numerical analysis is necessary to understand DFB operation. As will be explained below, grating operation depends critically on the end-length δ which sets the phase of reflection at the facets. The facets are high-reflectivity (HR) coated to prevent edge-emission, which is highly divergent and can couple back into the cavity externally to interfere with grating modes [88].

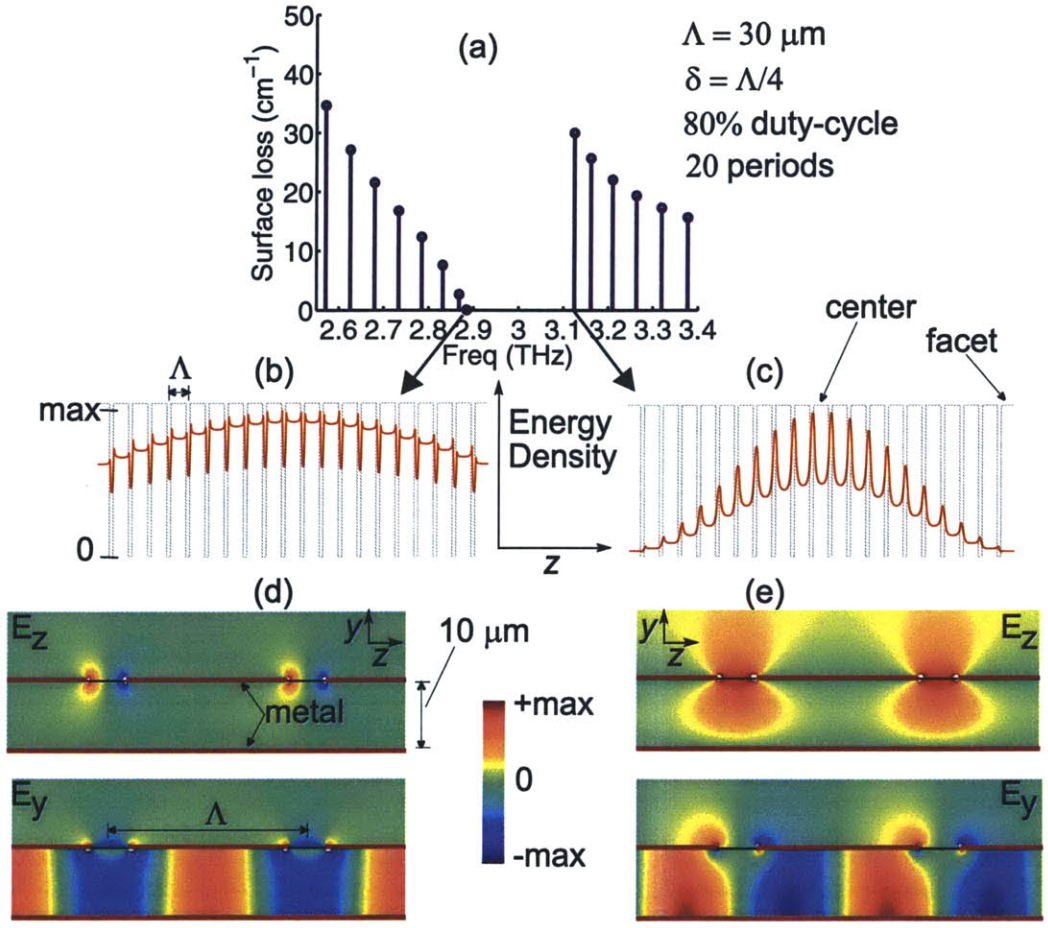


Figure 4-23: (a) Mode-spectrum for a finite length (infinite-width) grating structure of the type shown in Fig. 4-22. Plotted is propagation loss inside the waveguide due to surface out-coupling only. (b),(c) Energy-density averaged along waveguide height for lower and upper band-edge modes plotted along the length. The gray rectangular lines are shown as guides to locate grating apertures. (d),(e) Electric-field profiles near center of grating for lower and upper band-edge modes respectively.

Propagating mode solutions for an infinite-length periodic grating structure can be found numerically as the Floquet-Bloch eigen-modes [122, 55, 140]. When the propagation wavevector $2\pi n_{\text{eff}}/\lambda_0$ equals the Bragg wavevector $2\pi/\Lambda$, the mode is in resonance with the grating and two standing wave solutions exist, each at two slightly different frequencies, thus forming a bandgap at resonance. This is analogous to the energy splitting at Brillouin-zone boundary for electronic dispersion in a periodic lattice. For propagation vectors close to $2\pi/\Lambda$, the modes of a finite-length DFB structure could be written as a linear superposition of the two resonant band-edge modes of the infinite-length grating, each multiplied by slowly-varying envelopes along the length of the grating [140]. This is a good approximation so long as the envelopes do not change rapidly over a distance of Λ. This approximation will, however, break down for small cavities or for highly localized modes that can exist for some specific boundary conditions.

Finite-element (FEM) simulations were performed to analyze the grating operation [2]. FEM simulations were especially useful to understand grating behavior in finite-length structures for different boundary conditions (with and without HR-coated facets) and with additional aperiodicity added in the center of the gratings. A complex eigenfrequency solver was used to solve the electromagnetic eigenvalue problem formulated by surrounding the grating structure with lossy absorbing regions. To compute surface-loss only, both metal and active region were taken to be lossless. For an eigenmode solution with a complex eigenfrequency $\omega_{\text{complex}} = \omega_{\text{real}} + \text{i}\omega_{\text{imag}}$, ω_{real} corresponds to resonant frequency of the mode and $2\omega_{\text{imag}} n_{\text{GaAs}}/c$ corresponds to surface-loss per unit length for mode-intensity inside the waveguide. The challenge lies in the design of boundaries that can absorb radiation isotropically with minimum reflection. For two-dimensional simulations, an absorber was implemented by adding a linearly graded imaginary component to the dielectric constant of vacuum as a function of radial distance from the waveguide, as was described in section 4.3.3. For three-dimensional simulations, a standard perfectly-matched layer (PML) with anisotropic permittivity and permeability was implemented and its parameters coarsely adjusted to achieve minimum reflection at terahertz frequencies [81].

The grating mode-spectrum, and mode characteristics for the lower and upper band-edge modes for a finite-length grating structure with end-lengths $\delta = \Lambda/4$ are shown in Fig. 4-23. For the lower band-edge mode the energy density is localized more in regions with top-metal. Consequently, the mode character is more like Fig. 4-22(c), and E_z for this mode is small resulting in negligible surface-loss. Conversely, the upper band-edge mode is localized more under the apertures, and like in Fig. 4-22(d), can support large E_z in the apertures resulting in greater surface-loss. This also means that the propagation mode index n_{eff} for the upper-band edge mode is smaller since a larger fraction of the mode propagates in air, which is why this mode occurs at a higher frequency. A bandgap of 0.24 THz, which is $\sim 8\%$ of the center frequency, is observed for such a grating indicating very strong DFB action.

For an infinite-length grating the lower band-edge mode is antisymmetric in E_z, which interferes destructively and results in zero radiation loss. However, the upper band-edge mode is symmetric in E_z, which results in constructive interference leading to strong vertical radiation. In a finite-length structure, facets break periodicity of the grating and hence, grating modes acquire a mixed character and a non-uniform envelope shape. For the lower band-edge mode, $|E_y|$ has maxima at centers of the apertures and at middle points between two adjacent apertures, as shown in Fig. 4-23(d). Hence, an end-length of $\delta = \Lambda/4$ causes the location of the HR-coated facet to coincide with a null of E_y, thus automatically satisfying the facet boundary condition. This causes least perturbation to the infinite-length lower band-edge mode character and the envelope shape in Fig. 4-23(b) is relatively uniform along the length. Such an end-length is desired for robust single-mode operation as it minimizes longitudinal spatial-hole burning. However, the highly antisymmetric mode character yields very small surface-loss and low output power.

Figure 4-24, along with Fig. 4-23(a) and Fig. 4-23(b), capture the importance of facet boundary condition for this type of grating. Mode behavior for a grating will remain mostly unchanged if δ is replaced by $\delta \pm n\Lambda/2$. As the end-length δ is increased from $\Lambda/4$ to $\Lambda/2$, the lower-band edge mode deviates further from that of the infinite-length solution and it acquires a more non-uniform mode shape and

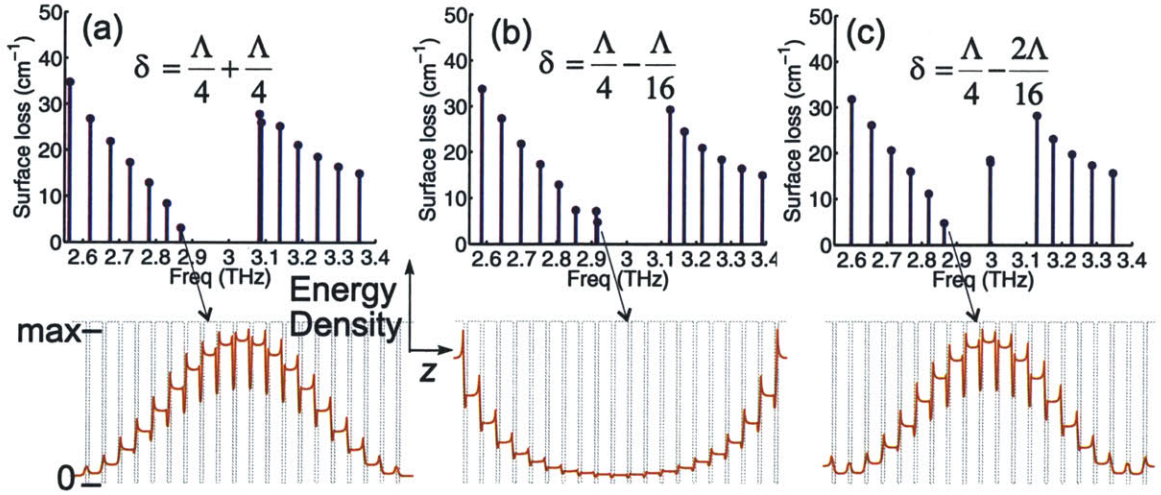


Figure 4-24: (a),(b),(c) Grating mode-spectrums and average energy-density plots along waveguide length for the lowest-loss modes corresponding to three different end-lengths. Other grating parameters are the same as in Fig. 4-23.

a greater surface-loss. For $\delta = \Lambda/2$, the envelope amplitude is zero at the facets (Fig. 4-24a). Figure 4-24(b) shows the mode-behavior when δ is made less than $\Lambda/4$. Only by slightly reducing δ, two modes that are mostly localized near the facets and close to each other in frequency become the lowest loss modes (mode-shape of only one of the two adjacent modes is plotted). Such facet localized modes are not desirable since they will lead to a divergent beam-pattern along the longitudinal (z) direction. For δ further away from $\Lambda/4$ in the negative direction, the facet-localized modes move through the band-gap (as in Fig. 4-24c) and eventually for $\delta = 0$, the mode spectrum is similar to that for $\delta = \Lambda/2$.

E_z for the lower band-edge mode is antisymmetric in z within the grating apertures (Fig. 4-23d). For apertures located symmetrically about the center of a finite length ridge, apertures in one half of the ridge length emit in opposite phase to those in the other half. This causes the envelope for E_z to be antisymmetric along the grating length, as shown in Fig. 4-25(a). Consequently, the far-field radiation pattern for the lower band-edge mode is predominantly two-lobed with a null in the center. Polarity of the E_z envelopes could be flipped from antisymmetric to symmetric and vice-versa for all modes by introducing a π phase shift in the grating center by means of an additional $\pm(2n+1)\Lambda/2$ waveguide section [105]. This is essentially equivalent to

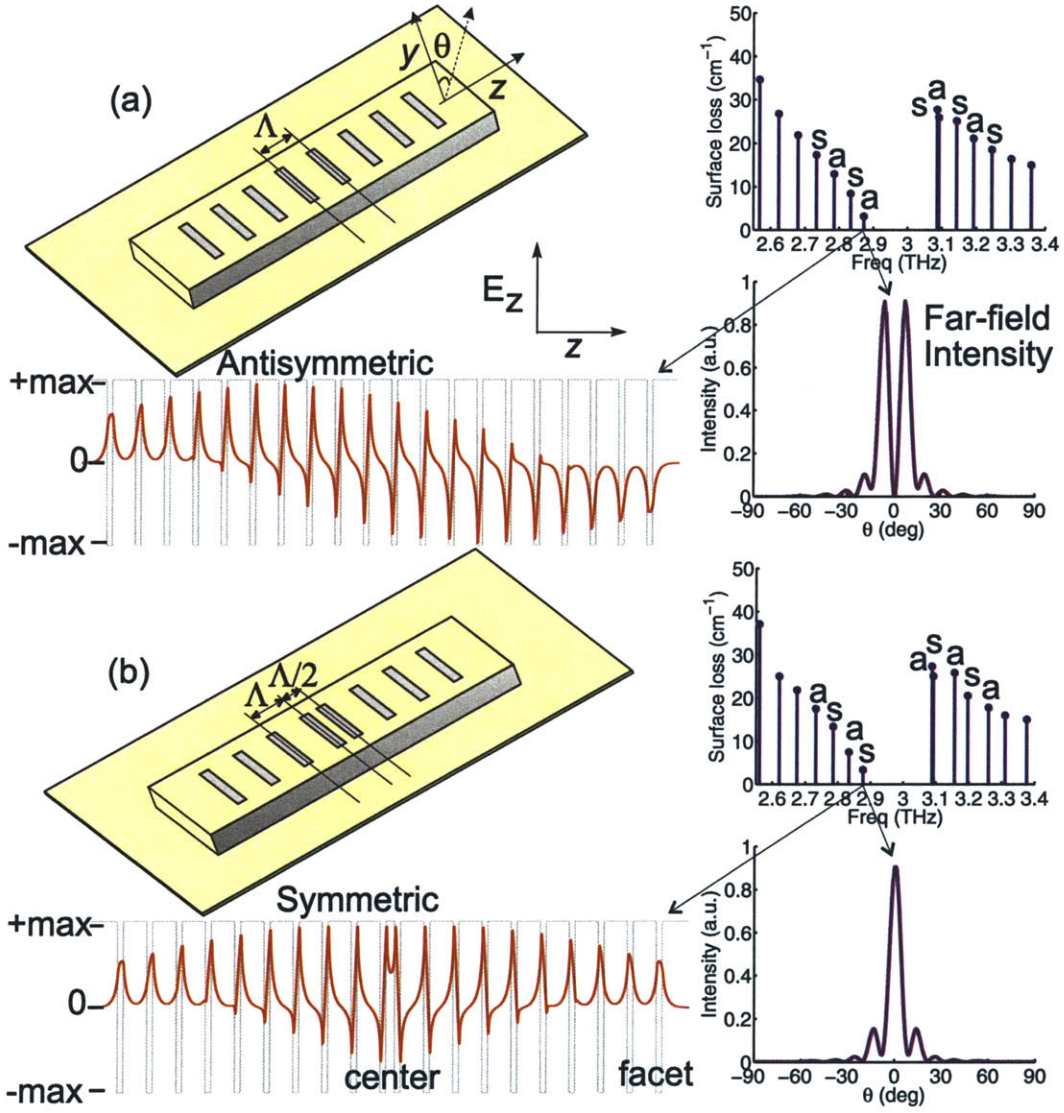


Figure 4-25: (a) Grating mode spectrum, and vertically averaged E_z plotted along the length with the corresponding far-field radiation pattern for the lower band-edge mode. These are calculated for end-length $\delta = \Lambda/2$ to show the mode behavior clearly since this δ produces a relatively bigger surface loss. Other grating parameters are the same as in Fig. 4-23. E_z envelopes are alternately antisymmetric ('a') and symmetric ('s') along the length for adjacent modes. (b) Same quantities plotted for a similar structure that has a $\Lambda/2$ section of waveguide removed from the center to create a $-\Lambda/2$ phase-shift in the grating. Consequently, polarities for E_z envelopes are switched from 'a' to 's' and vice-versa.

flipping the phase of emission from all apertures in one half of the grating length by π with respect to the other half. The mode-spectrum, energy-density mode-shapes, and grating behavior for different end-lengths essentially remain the same as that of a grating without the phase-shift. Figure 4-25(b) shows one such case, when due to a $-\Lambda/2$ central phase-shift, entire grating emits in phase and a single lobed far-field pattern is obtained with a maxima in the center. Such a radiation pattern is highly desirable for practical applications.

4.5.3 Early attempts: problems with the excitation of higher order lateral modes

In the early attempts, grating structures with varying cavity dimensions, grating periods, end-lengths, and grating duty-cycles were fabricated by dry-etching [93]. The FL179C-M9-2 MBE wafer was used as the gain medium (section 3.3.1, Table 3.1). Wafer bonding for MM waveguide fabrication was done using a Cu-Cu technique as detailed in section 4.2. After the grating-metal lithography, ridges were defined by reactive-ion etching in a BCl_3:N_2 plasma with PECVD silicon-nitride used as an etch-mask. While the facets were HR-coated with silicon-nitride/Ti/Au, the sidewalls were left open.

Figure 4-26(a) shows the scanning-electron micrograph (SEM) of a typical fabricated device. Out of the many devices tested, few lased in single-mode. A spectrum of one such device, that emitted 1.5 mW of cw power at 5 K, is shown in Fig. 4-26(b). However, the radiation pattern of this device measured using a microbolometer camera showed multiple lobes with a null in the center along the lateral (x) direction, as shown in Fig. 4-26(c). This beam pattern indicates that the lasing mode is a higher order lateral mode. The emission frequency for single-mode devices usually did not scale according to the grating-period Λ. These observations suggested that the desired grating mode (with the fundamental lateral mode) was not excited.

Due to the surface-plasmon nature of the modes that are attached to the top and bottom metal plates, MM waveguides provide strong confinement for multiple

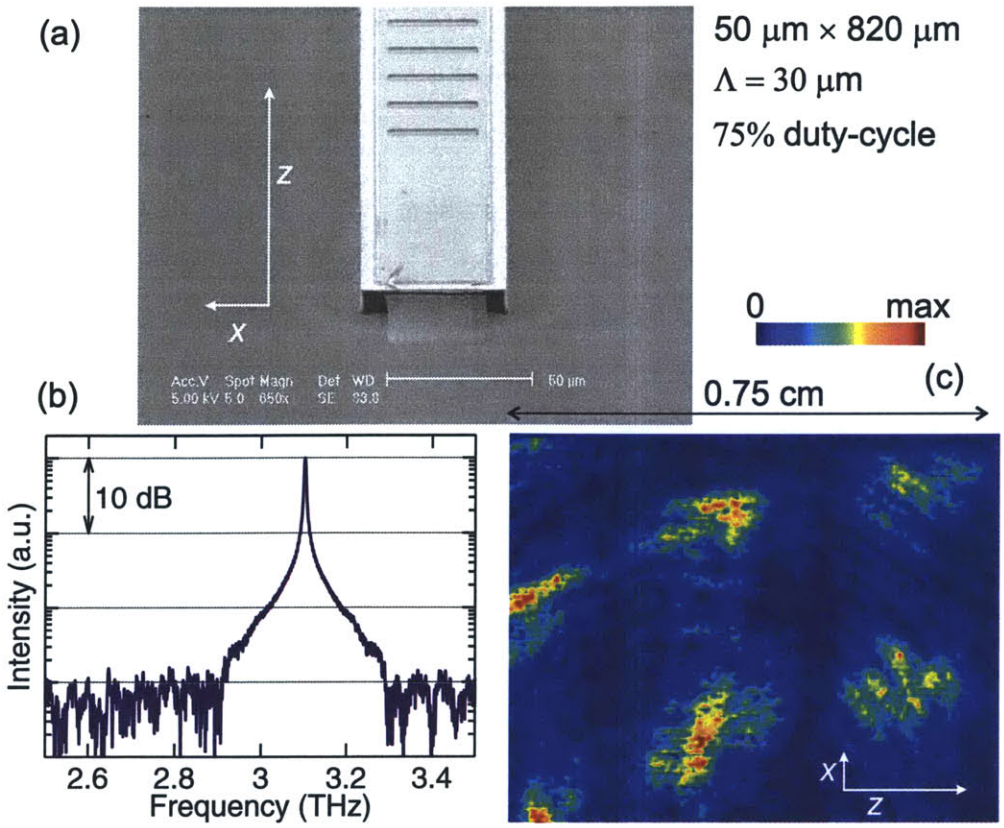


Figure 4-26: (a) SEM image of a grating device fabricated by dry-etching during the early attempts of grating design. (b) Single-mode cw spectrum from one such grating device measured at 5 K (which is not believed to the desired grating mode). (c) Far-field intensity pattern for the same device measured at a distance of ≈ 2.5 cm with a microbolometer camera [101].

higher lateral modes. For example, in a 40 μm wide ridge waveguide at 3 THz, modal-confinement for the first two lateral modes is greater than 90 %. For facet-reflectivities that are close to unity, it is quite common experimentally to observe higher order lateral modes in Fabry-Perot cavities. To investigate the behavior of the higher order lateral modes in the grating structures, FEM simulations were done in three-dimensions. Figure 4-27 shows computed grating mode-spectrum for fundmental and second lateral modes for one such grating. As observed, surface-loss for the second lateral modes in a 40 μm wide waveguide is lower than that of the fundamental modes. Consequently, these modes will reach the lasing threshold first and then gain clamping prevents the device from operating at the desired fundamental modes. For a general grating structure, the lowest-loss modes (due to surface coupling) are the

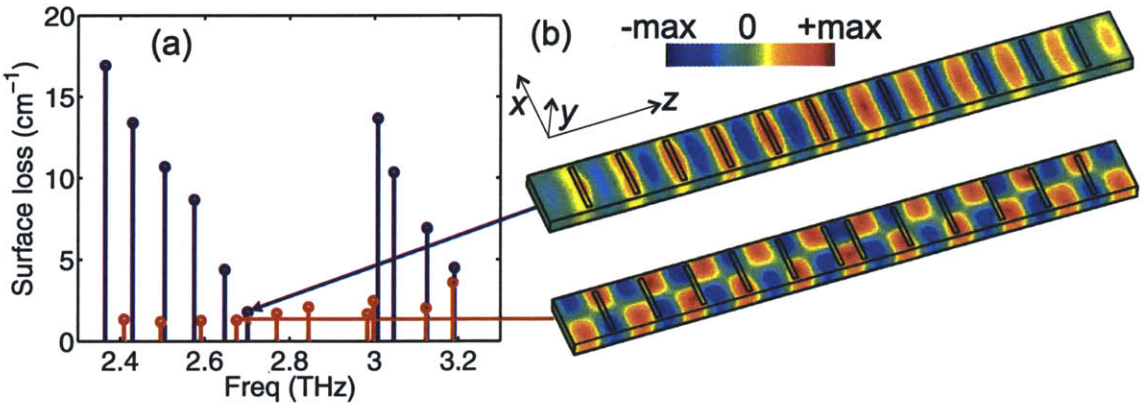


Figure 4-27: (a) Grating mode spectrum obtained by three-dimensional simulation of a 12-period grating with $\Lambda = 32$ μm, $\delta = \Lambda/4 + \Lambda/8$, 90 % duty-cycle, and $-\Lambda/2$ phase-shift in center. The cavity facets are covered with metal while the sidewalls are left open. Ridge dimensions are 40 μm $\times$ 392 μm $\times$ 10 μm. Modes in blue are fundamental lateral modes and the ones in red are second lateral modes. (b) E_y plots in the waveguide for the lower band-edge fundamental lateral mode, and for a second lateral mode close to it in frequency.

highest-order lateral modes that are supported by the waveguides. For those modes, their longitudinal wavevector k_z becomes much smaller than the total wavevector k_{GaAs} ($= 2\pi n_{\mathrm{GaAs}}/\lambda_0$) in the semiconductor due to a significant component k_x along the lateral dimension. For such modes, the surface out-coupling from the grating is much weaker, since E_z is proportional to k_z ($E_z \propto \partial H_x/\partial y \propto k_z \partial E_y/\partial y$) for a given mode intensity inside the waveguide. In general, mode competition from higher lateral modes makes operation of this type of gratings with high-reflectivity at the facets less predictable.

The dry-etching technique used for fabricating these structures is also not without problems. Firstly, nitride coverage on edges of the grating apertures might not be isotropic making the active region in apertures prone to etch-damage. Secondly, the bottom metal surface reached at the completion of dry-etch is rough, making it hard for any bonding pads to stick on it. This puts an additional requirement on the grating to have long end-lengths (as shown in Fig. 4-26a) so that electrical contact could be made by wire bonding on top of the mesa. This is not desirable since wire-bonds could pose a strong perturbation to the grating structure due to a significant fraction of the mode that exists near the end facets. The following section describes a modified

fabrication process with wet-etching that helps counter the aforementioned problems.

4.5.4 Selective excitation of the fundamental lateral mode using metal covered sidewalls

The higher lateral modes can selectively be made more lossy by covering waveguide sidewalls with insulator/metal and extend these layers to bonding pads, as will be described in this section. To do this viably, outward-sloped sidewalls are desirable, which is possible to obtain by wet-etching. Furthermore, wet-etching results in a smooth bottom surface that is better suited for bonding pad fabrication so that wire bonds could be made away from the mesas. For GaAs, the wet-etch profile is usually different in different lattice directions for most reaction-rate limited etchants. Figure 4-28(a) and Fig. 4-28(b) show etch profiles in the two different directions obtained typically with H_2SO_4:H_2O_2:H_2O or H_3PO_4:H_2O_2:H_2O etchants for etch masks aligned parallel to the cleave ($\langle 110 \rangle$) directions. For a cavity that needs to be coated with metal on all the sides, the inward slope in Fig. 4-28(b) will not be suitable for metal patterning by liftoff lithography.

This problem was avoided by aligning the etch mask at 45° angle to the cleave directions (i.e. along $\langle 100 \rangle$) on the (100) GaAs substrate [178]. Figure 4-28(c) shows the etch profile obtained with a square shaped etch-mask. Outward-sloped sidewalls are obtained on all four sides. Once the bottom-metal is reached, the mesa could be over-etched to obtain a more vertical side-profile since the bottom of the mesa etches faster compared to the top due to a galvanic effect in the presence of the bottom-metal. For an etch mask with 45° angle alignment, several H_2O_2 based etchants were tried with H_3PO_4, H_2SO_4, HCl, and NH_4OH in different concentrations. The HCl based etchants caused severe undercut at the bottom of the mesas, while the NH_4OH based etchants damaged the sidewalls making them jagged at several locations, as is shown in Fig. 4-29. The etch profiles for H_2SO_4 and H_3PO_4 based etchants were very similar, although etching was two to three times faster with H_2SO_4. An etch solution of H_2SO_4:H_2O_2:H_2O 1:8:80 was finally selected for its relatively fast etch-rate

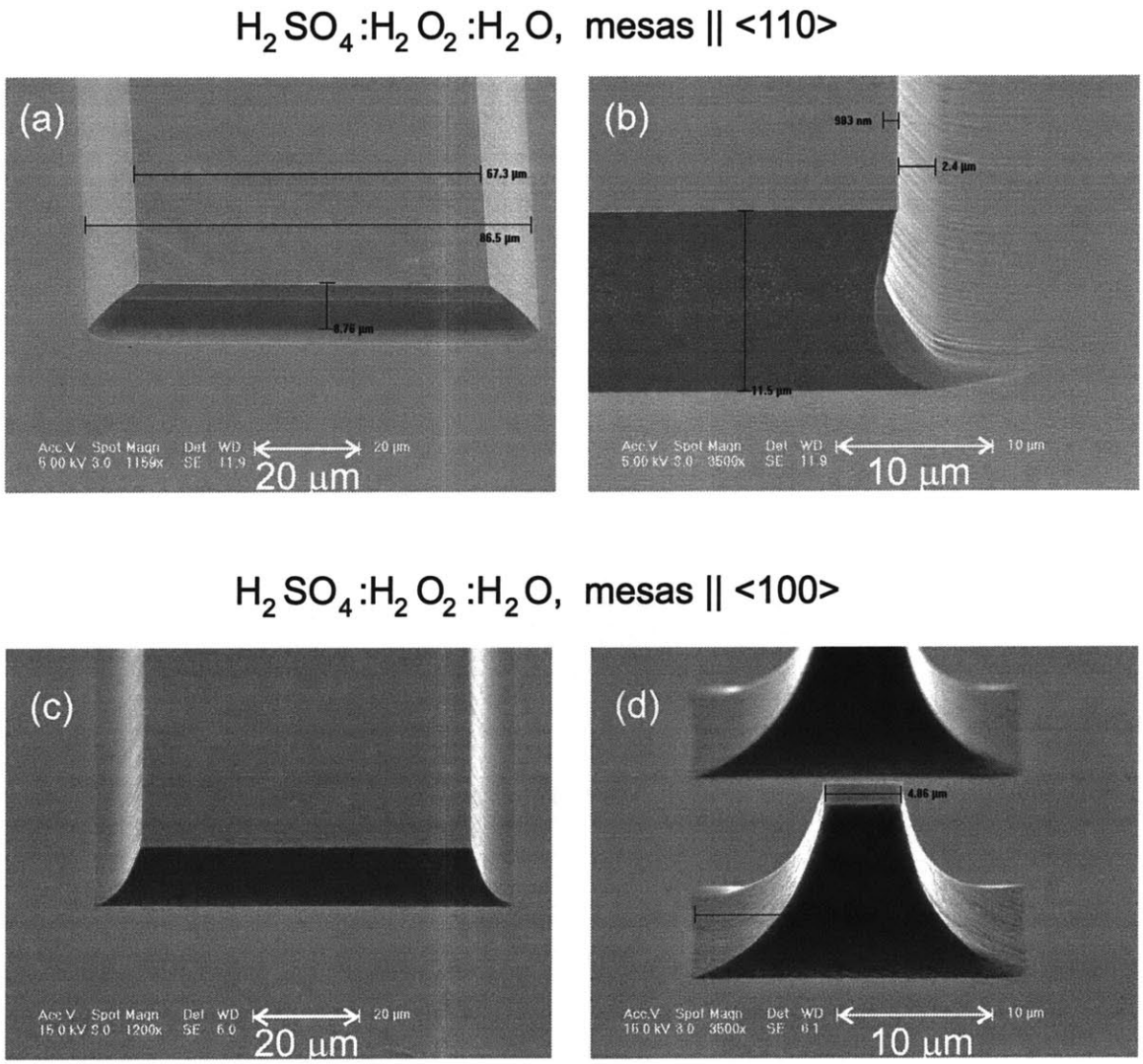


Figure 4-28: (a),(b) SEMs of a wet-etched GaAs/$Al_{0.15}Ga_{0.85}As$ MQW region grown on a (100) GaAs substrate taken along longitudinal and lateral directions, respectively, of a rectangular mesa. Etch-mask for the mesa was aligned parallel to the $\langle 110 \rangle$ directions. The bottom surface is GaAs. (c),(d) SEMs of two different mesas when etch-mask was aligned parallel to the $\langle 100 \rangle$ directions. The etchant used was H_2SO_4:H_2O_2:H_2O 1:8:80 and etch-depth was $\approx$ 10 μm. The bottom surface is metal in this case; similar profiles are obtained with a GaAs substrate.

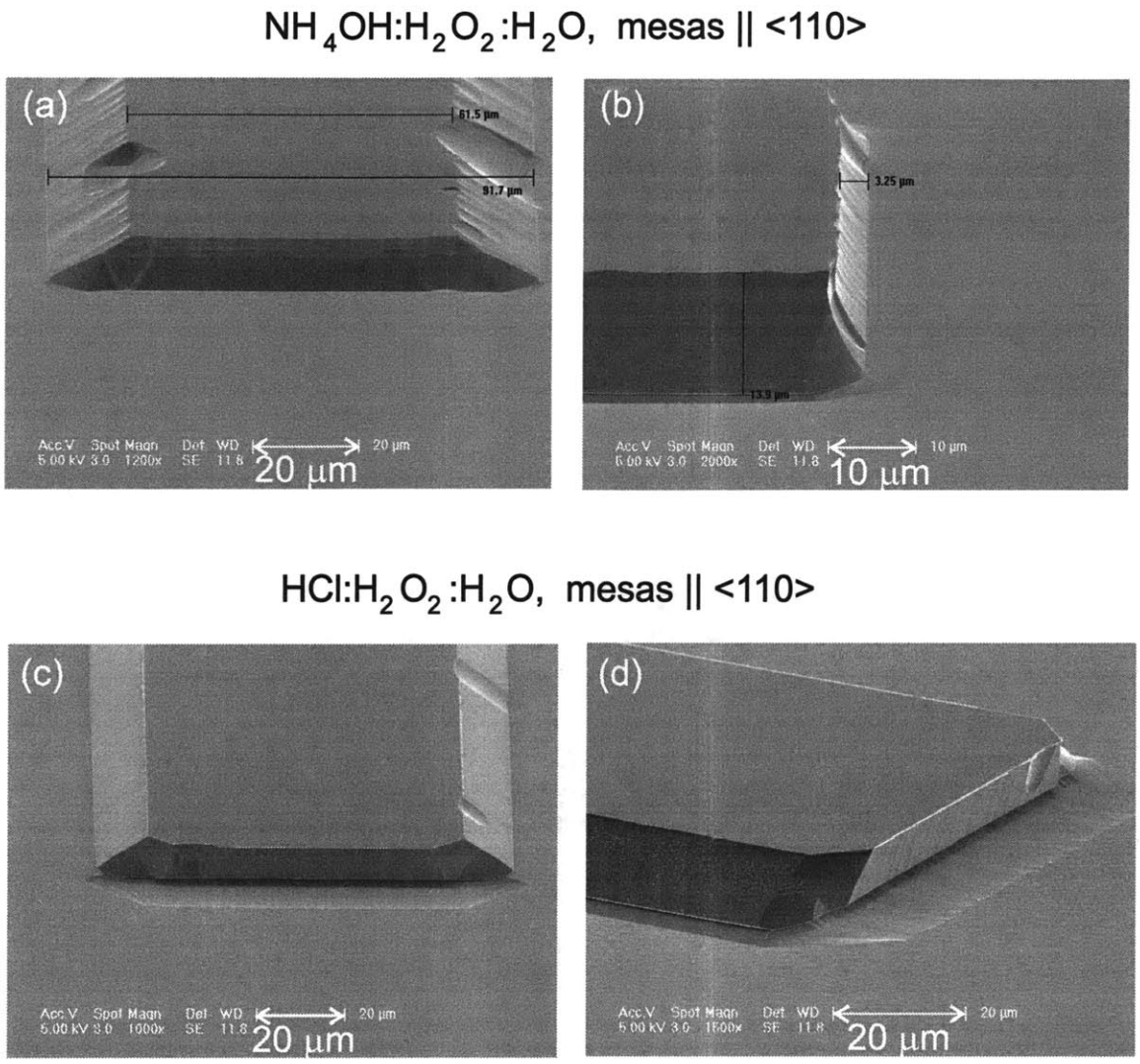


Figure 4-29: (a),(b) SEMs of a wet-etched GaAs/$Al_{0.15}Ga_{0.85}As$ MQW region grown on a (100) GaAs substrate taken along longitudinal and lateral directions, respectively, of a rectangular mesa. Etch-mask for the mesa was aligned parallel to the $\langle 110 \rangle$ directions. A $NH_4OH:H_2O_2:H_2O$ 11:6:490 solution was used for this etch. Outward sloped sidewalls were obtained when the mesas were made parallel to the $\langle 100 \rangle$ directions (similar to those in Fig. 4-28) but the jaggedness on the sidewalls remained similar to as shown. (c),(d) Same as in (a) and (b) except that a $HCl:H_2O_2:H_2O$ 1:1:9 solution was used for this etch. The undercut at the bottom of the mesas remained similar to as shown when the mesas were made parallel to the $\langle 100 \rangle$ directions. The bottom surface is GaAs in all these images.

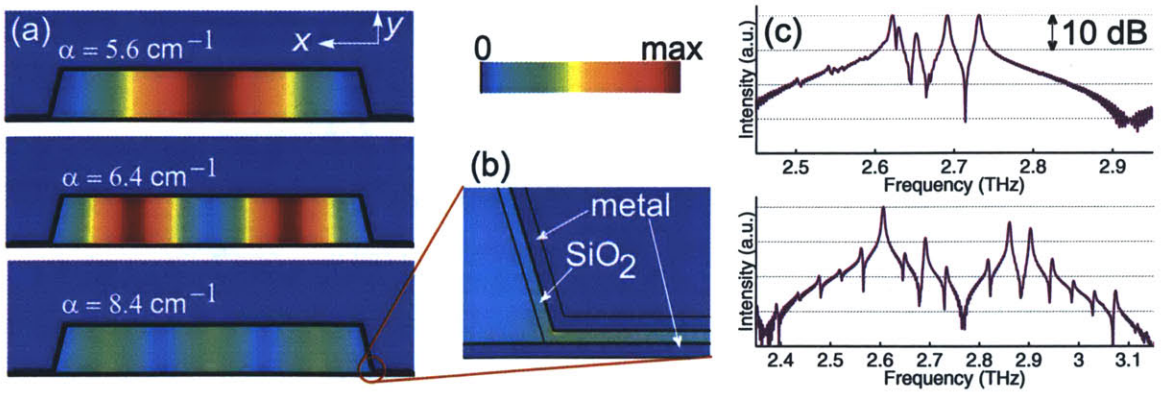


Figure 4-30: (a) Energy-density profiles for different lateral modes at 2.9 THz in an 80 μm wide metal-metal waveguide covered with SiO_2/Ti/Au on the sidewalls. Calculated propagation loss α due to metal only is indicated. The aspect ratio of the plotted geometry is not to scale. (b) Close-up for the third-lateral mode near the bottom edge of the ridge. (c) Measured cw spectra on log scale at 5 K from two Fabry-Perot ridge lasers with cleaved facets (size 80 μm $\times$ 0.92 mm) located close to each other on the same die. The two devices are nominally identical except the one without metal on the sidewalls yielded higher order lateral modes (top, where the mode spacing is nonuniform); while the one with metal on the sidewalls yielded only fundamental lateral mode (bottom).

$(0.5 - 0.7\ \mu\text{m/min})$ and smooth sidewalls.

Sidewall coverage with insulator/metal can selectively make higher lateral modes more lossy for MM waveguides at terahertz frequencies. The waveguide width is comparable to wavelength, and hence the higher lateral modes can couple to propagating modes in the bonding pad region, which is quite lossy since it is essentially a parallel plate waveguide with small plate separation. Figure 4-30(a) shows propagating mode solutions for an 80 μm wide ridge waveguide covered with 300/30/350 nm of SiO_2/Ti/Au on the sidewalls. The active region is taken to be loss less and the metal is approximately modeled with the following Drude parameters: $n_{\text{Ti}} = 4.6 \times 10^{21}\ \text{cm}^{-3}$, $n_{\text{Au}} = 5.9 \times 10^{22}\ \text{cm}^{-3}$, and τ for both metals is taken as 50 fs, a value greater by a factor of ~ 2 compared to the published room temperature values [123] (which is justified by the fact that typically for these metals, the published low-frequency mobilities at ~ 100 K are greater by a factor of $2-5$ compared to their room temperature values). The extra loss for higher lateral modes is due to bonding-pad regions that could develop high fields at sharp corners, rather than insulator/metal on the

sidewalls where field is almost zero. Figure 4-30(c) shows experimental confirmation of this behavior. An edge-emitting Fabry-Perot laser with insulator/metal on its sidewalls exhibited a spectrum with equal frequency spacing between its modes. From this it can be inferred that the device lased in Fabry-Perot modes corresponding to the fundamental lateral mode. However, a device without metal on its sidewalls had non-uniformly spaced frequencies in its spectrum, which suggests that more than one type of lateral mode was lasing.

Experimental results

The techniques outlined in the previous section were implemented and the modified grating devices have performed as expected. Figure 4-31 shows SEMs of grating devices fabricated by wet-etching with cavity dimensions of 65 μm$\times$ $\sim$ 0.9 mm $\times$ 10 μm. 30-period gratings with Λ in the range of 28–32 μm were implemented with $\delta \approx \Lambda/4 + 2$ μm, 80 % duty-cycle, and a $-\Lambda/2 - 0.5$ μm phase-shift in the center. The phase-shift was chosen slightly smaller than $-\Lambda/2$ since it reduces the relative amplitude of side-lobes in the far-field pattern of Fig. 4-25 by apodizing the mode shape inside the waveguide. Gratings without a phase-shift were not implemented because of lack of available space on mask. The choice of the end-length δ, as noted previously, predominantly determines the mode-shape for the lasing mode inside the waveguide (Fig. 4-23(b) and Fig. 4-24). While $\delta = \Lambda/4$ gives the most uniform mode-shape, the corresponding surface out-coupling is negligible and the mode spectrum depends very sensitively to δ. By increasing δ, surface out-coupling could be increased at the cost of a more non-uniform mode inside the waveguide, which is more prone to longitudinal spatial-hole burning. To achieve the primary purpose of robust single-mode operation, δ in this design was only slightly detuned from $\delta = \Lambda/4$. The choice of a large grating duty-cycle was based more on electrical transport requirements rather than grating operation. A smaller duty-cyle provides greater surface out-coupling; however, due to absence of the highly doped GaAs contact layer in the apertures, current injection becomes more non-uniform. In that situation, different modules of the QCL may be biased differently and push the device into a negative-differential

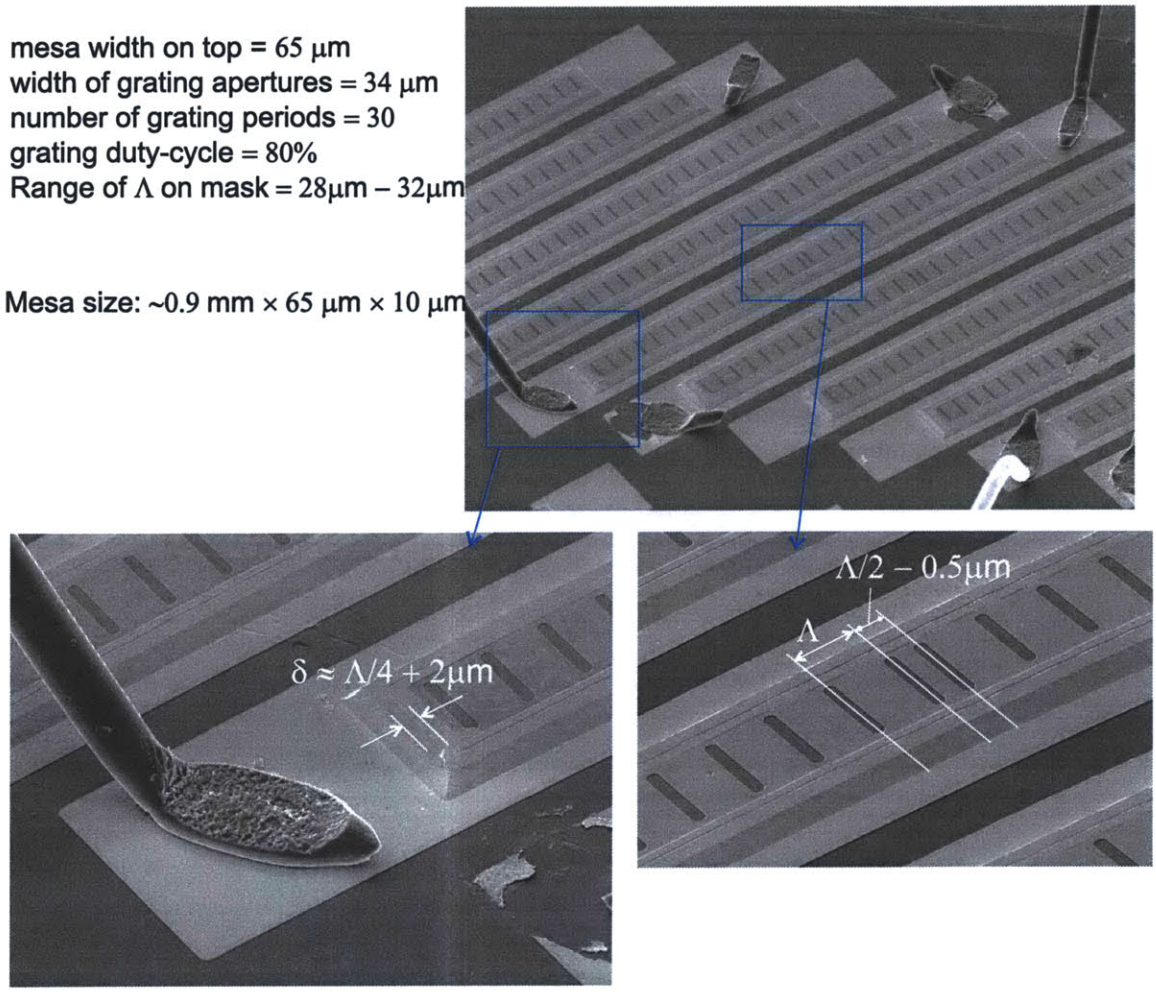


Figure 4-31: SEMs of metal-metal grating devices fabricated with mesas along ⟨100⟩ directions. The sidewalls and facets are covered with 300/30/350 nm of SiO_2/Ti/Au. Al wire bonds are made on bonding pads away from the mesas to electrically bias the lasers. SiO_2 isolates the bonding pads from the bottom Ta/Cu.

resistance region at an early bias, thereby limiting its lasing dynamic range.

A 2.9 THz QCL active region grown on a (100) GaAs substrate (design FL179C-M9, growth EA1189, Table 3.1)was used for fabrication of these grating devices. The fabrication sequence was as follows. The Cu-Cu thermocompression bonding technique as detailed in section 4.2 was used as a first step towards the MM waveguide fabrication. Following back-substrate removal by lapping and selective etching, the 100-nm thick highly doped top contact GaAs layer, was completely removed by wet-etching. Ti/Au (20/350 nm) metal-gratings were then defined on the 10 μm thick epitaxial active region by contact-lithography, with mask edges aligned along $\langle 100 \rangle$ directions. A second lithography step to cover the grating-metal was performed for the photoresist to be used as a mask for wet-etching of mesas. Wet-etching followed in a H_2SO_4:H_2O_2:H_2O 1:8:80 solution for $\approx$ 25 minutes. This included a few minute over-etch after all bottom metal was exposed to reduce the slope on the sidewalls and allow for a more uniform current density through the height of the mesa. 300 nm of PECVD silicon-dioxide was then blanket-deposited everywhere at a temperature of 300°C. The next lithography step was to open windows on top of mesas to remove the silicon-oxide by wet-etching in buffered-HF. The final lithography was for Ti/Au (30/350 nm) deposition on sidewalls and facets, as well as to form bonding pads adjacent to bottom of the mesas. The outward sloped sidewalls and facets in this process allow this deposition to be done in a single lithography step. Before backside-metalization of the wafer, the substrate was thinned to 160 μm to improve heat-sinking. A die consisting several tens of such grating devices along with some Fabry-Perot ridge waveguides was then In-soldered on a Cu chip carrier and mounted for testing on a cold-plate in a vacuum cryostat. The emitted laser light was collected normal from the surface without any optical components inside the cryostat.

Figure 4-32(a) shows cw spectra from different surface-emitting grating devices within the range of grating-periods that were implemented on the fabrication-mask ($\Lambda = 28-32$ μm). The spectra were measured at 5 K using a Nicolet 850 spectrometer (purged with N_2 gas) with a room-temperature deuterated triglycine sulfate (DTGS) pyroelectric detector at a resolution of 0.125 cm^{-1} (= 3.75 GHz). Most devices tested,

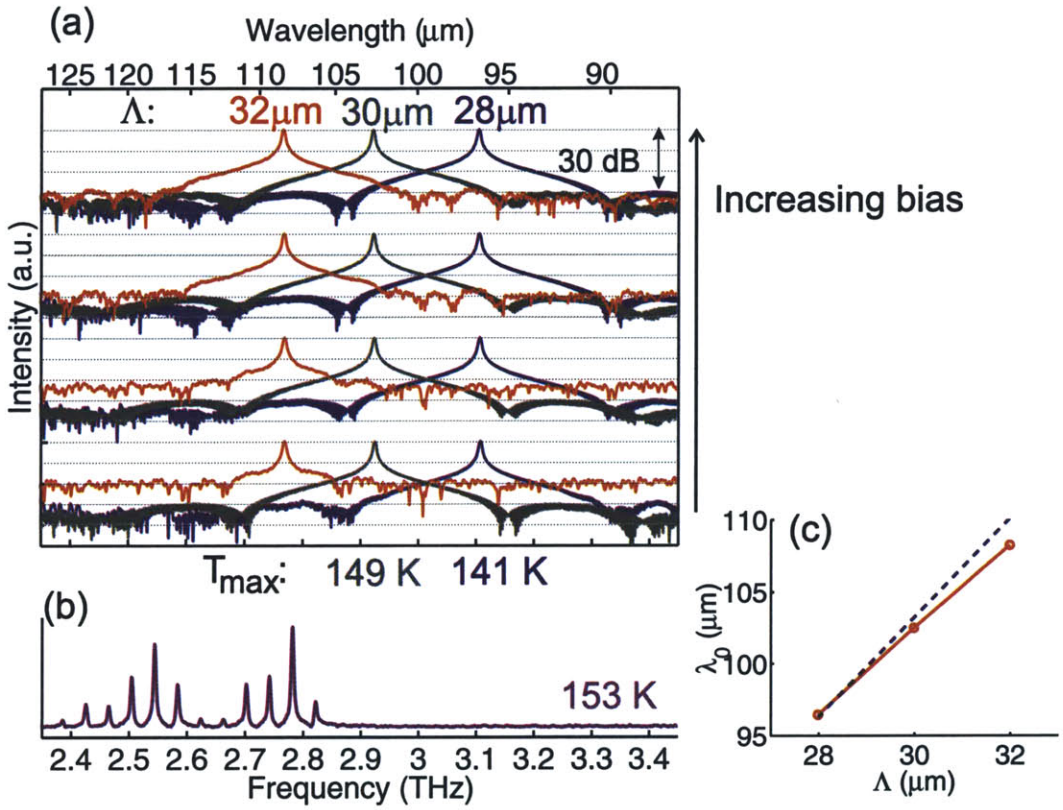


Figure 4-32: (a) 5-K cw spectra for three different grating devices (each color corresponding to a device with different Λ), plotted on log scale. Spectra for different bias are plotted starting from near-threshold bias at bottom to peak-bias at top. Signal recorded from the $\Lambda = 32$ μm device, which lased at 2.77 THz, was very weak due to it being on a strong atmospheric water-absorption line. Maximum temperature of pulsed operation are also indicated. (b) CW spectrum from a Fabry-Perot ridge laser that was located adjacent to grating devices on the same die, plotted on linear scale. (c) λ_0 versus Λ variation (in solid red). A line going through origin and corresponding to $n_{\text{eff}} \approx \lambda_0/\Lambda = 3.44$ is also plotted (in dashed blue).

like the ones reported in Fig. 4-32(a), lased in single-mode for all bias-conditions, down to the instrument limited noise floor of 30 dB below maximum. There were, however, a few random devices that had additional modes, which are attributed to lithographic fluctuations. The lasing wavelength λ_0 scales linearly with grating period Λ as shown by the solid red line in Fig. 4-32(c). The line, however, does not cross origin since the mode index n_{eff} for cavities with metal-covered sidewalls changes with wavelength when the wavelength is comparable to the width. The dashed line in blue shows a line passing through origin corresponding to a n_{eff} value of 3.44 for the $\Lambda = 28$ μm device. For a device at a longer wavelength, n_{eff} ($\propto k_z$) will drop as the lateral component k_x increases, which explains the deviation of $\Lambda > 28$ μm grating devices from the dashed blue line. For the range of grating periods available on the mask, an operation range of $\sim$ 0.35 THz is demonstrated for the single-mode grating devices with this gain medium. The multi-mode spectrum for a Fabry-Perot edge-emitting laser from the same die shown in Fig. 4-32(b), however, indicates the possibility of accessing an even wider range.

Radiation patterns for grating devices were measured at 5 K with a Ge:Ga photodetector in pulsed mode, and with a microbolometer camera in quasi-CW mode (25 % duty-cycle). Measurement results from the $\Lambda = 30$ μm device are shown in Fig. 4-33. The other two grating devices also had similar radiation patterns. This measurement was done with only a $\sim$ 300 μm thick polypropylene dewar window in between the laser and the detector. A single-lobed pattern in the longitudinal (z) direction is observed as expected. The beam is relatively divergent in the lateral (x) direction due to a waveguide width that is of the order of the wavelength. In the far-field image of Fig. 4-33(a), a single-lobed pattern is observed in the lateral direction as well; however, the camera image in Fig. 4-33(b), which is taken from a much closer distance (due to a small camera size), shows two additional lobes which are likely due to evanescent field.

Figure 4-34 shows measured optical power versus current (L-I) characteristics for the $\Lambda = 30$ μm grating device in both pulsed and cw mode of operation. Light was collected without use of any cone optics. In pulsed operation, the device lased

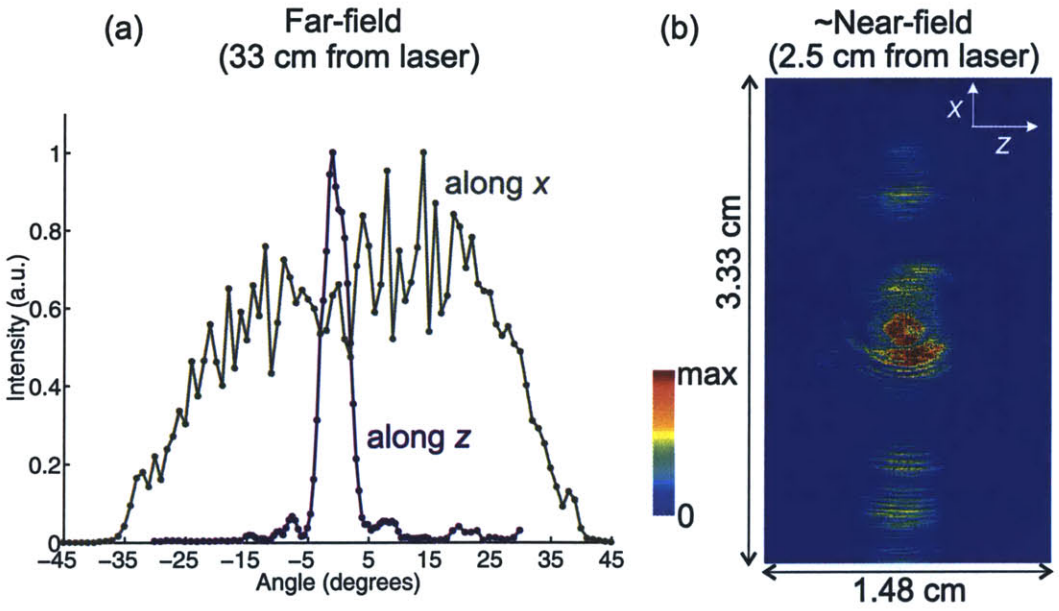


Figure 4-33: (a) Far-field radiation pattern measured from the $\Lambda = 30$ μm grating device with a He-cooled Ge:Ga photo-detector at an angular resolution of $< 1°$, and a distance of 33 cm from the laser. The full-width at half-maximum in the longitudinal (z) direction is $\sim 5°$. (b) Real-time snapshot of the radiation pattern at a distance of ~ 2.5 cm from the laser taken with a room-temperature 320×240 element (1.48×1.11 cm^2) microbolometer camera [101]. Three images taken at different x locations are stacked vertically to show the full image.

up to a maximum heat-sink temperature ($T_{\max}$) of 149 K and the threshold current density (J_{th}) at 5 K was 460 A/cm^2. A temperature tuning of 19.7 GHz was observed for the single-mode spectrum from 5 K to 147 K. In cw mode, the corresponding value for $T_{\max}$ was 78 K and J_{th} at 5 K was 560 A/cm^2. 6.1 mW of cw power was detected with a thermopile detector (Scientech, model AC2500) placed adjacent to the cryostat window. CW power per facet for the edge-emitting Fabry-Perot devices with comparable dimensions was typically smaller by a factor of ~ 2. The bottom panel of Fig. 4-34 also shows the measured cw voltage versus current (V-I) characteristics for the grating device. The negative differential-resistance (NDR) occurs at ~ 15 V beyond which the device stops lasing. Removal of the highly doped GaAs contact layer on top of the active region stack causes an extra 2–3 V drop to occur at the top metal-semiconductor Schottky contact, causing extra Joule heating in the grating devices and thus limiting their cw performance. However, the lasing dynamic range

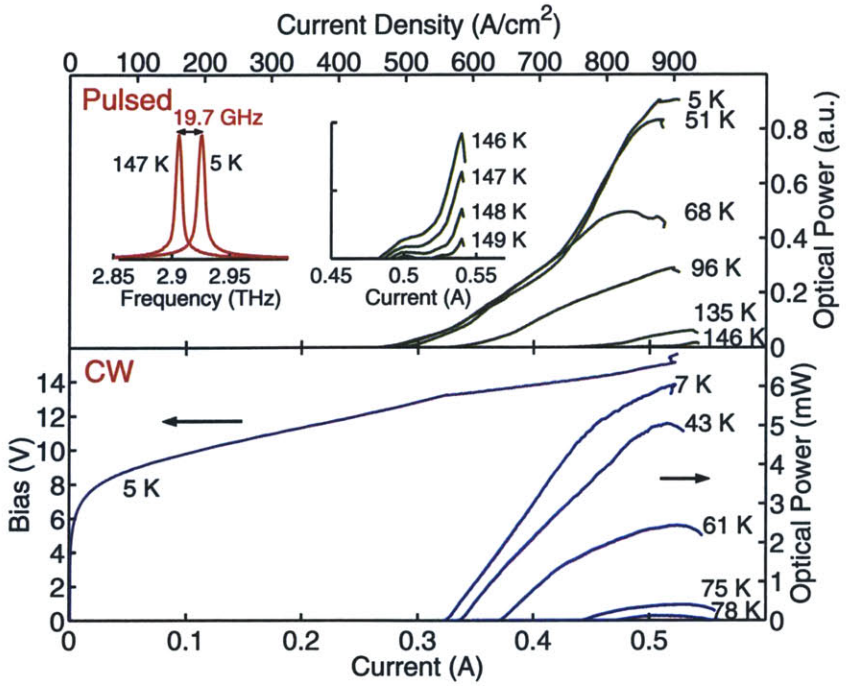


Figure 4-34: Pulsed (top) and cw (bottom) *L-I* characteristics measured from the $\Lambda = 30$ μm grating device (ridge-size: 65 μm $\times$ 0.90 mm). The cw *I-V* measured at 5 K is also plotted. The pulsed data is taken using 200-ns pulses repeated at 10 kHz with a He-cooled Ge:Ga photo-detector, while the cw data is measured with room-temperature pyroelectric detector. The upper panel inset on the left shows temperature tuning of the single-mode spectrum due to change in the active region refractive index from 5 K to 147 K, while the inset on the right shows an expanded version of the *L-I*s at higher temperatures.

in current and hence the pulsed $T_{\max}$ remain unaffected.

L-I characteristics for the $\Lambda = 32$ μm grating device could not be measured since it was lasing right on a strong atmospheric water absorption line. While the $\Lambda = 30$ μm grating device lased up to 149 K in pulsed mode, the $\Lambda = 28$ μm grating device lasing further away from the peak frequency of the gain spectrum (as deduced from the Fabry-Perot mode-spectrum of Fig. 4-30c) lased up to 141 K. For comparison, edge-emitting Fabry-Perot ridge lasers with metal on the sidewalls, measured from the same die as that of the grating devices, lased up to 153 K. It is likely that temperature performance of grating devices is slightly degraded due to non-uniform current injection on top because of metal gratings, and the lasing frequency is away from the gain peak. This suggests that the surface losses induced by the gratings for

the parameters used in the designs are comparable with the facet loss of the Fabry-Perot devices. Further adjustment of grating parameters, such as duty-cycles, phase of reflection at the facets, and central grating phase-shifts, could lead to much higher output power levels.

4.5.5 Selective excitation of the fundamental lateral mode using exposed doped contact layer on top

The method of using metal on the sidewalls to selectively eliminate the higher order lateral modes as described in section 4.5.4 is bound to be less effective for wider MM waveguides that can support a multitude of different lateral modes. DFB waveguides of the type discussed in section 4.5.4 with a width greater than 65 μm were not tried, and thus an estimate of the widest width (or more appropriately, the largest width/λ ratio) for which that technique can be robustly utilized is not yet available.

This section describes surface-emitting DFB designs that can be potentially implemented in much wider MM waveguides. Such waveguides are useful to obtain narrow beam emission in both the longitudinal and the lateral dimensions as compared to the DFBs of section 4.5.4, for which the beam was divergent in the lateral (x) direction due to the narrow width of the grating apertures (Fig. 4-33).

If the thin highly doped contact layer on the top of the active region (section 3.3.3) is left exposed instead of covering it with metal, the propagation losses of the double surface-plasmon mode in the resultant waveguide increase by multiple orders of magnitude, thereby effectively quenching the mode within a distance of few micrometers. This scheme was first utilized to realize anti-reflecting regions at the longitudinal ends of surface-emitting DFB structures in terahertz QCLs with SISP waveguides [41]. A similar method has been applied, albeit with a very different operating principle, to surface-emitting DFB structures in wider ($\gtrsim$ 100 μm) MM waveguides, as shown for the "design B" in Fig. 4-35. In this case, narrow stripes of doped contact layer are left exposed near the sidewalls on the top of the waveguide to implement *lateral loss sections*, which help make the higher order lateral modes more lossy, as is shown

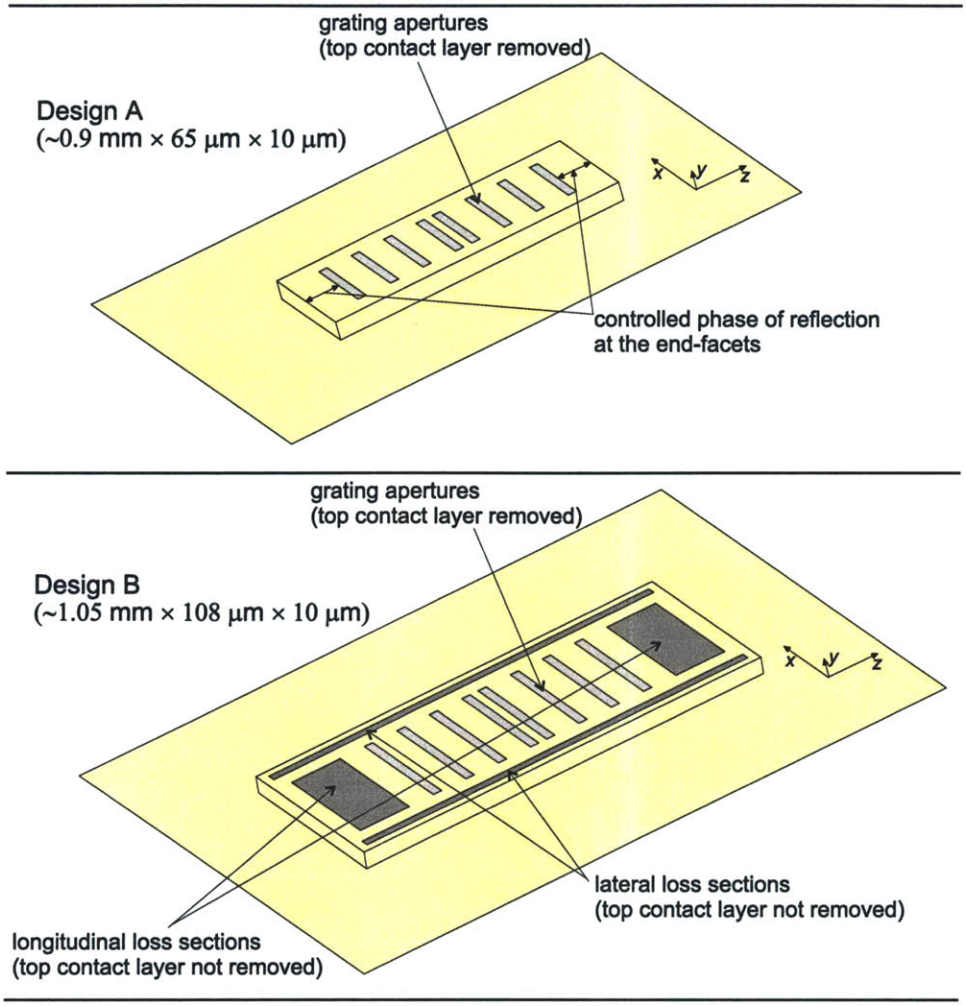


Figure 4-35: Schematics of two surface-emitting DFB designs for terahertz QCLs in MM waveguides. Design A is for narrower waveguides (width $\lesssim 80$ μm), the results of which were presented in section 4.5.4. In this design, the higher order lateral modes are made more lossy due to the metal on the sidewalls, and the phase of reflection at the end-facets is well defined. Design B, which is discussed in this section, is for wider waveguides (width $\gtrsim 100$ μm) and achieves similar effects due to *lateral loss sections* (with exposed doped contact layer on top). The design shown in this schematic also has *longitudinal loss sections*, which eliminate the need for a well defined phase of reflection at the end-facets due to them being close to perfectly absorbing (calculated reflectivity $\sim$ 1% at the either end). However, the scheme of controlled phase of reflection such as that in Design A can also be used, as is discussed in the text.

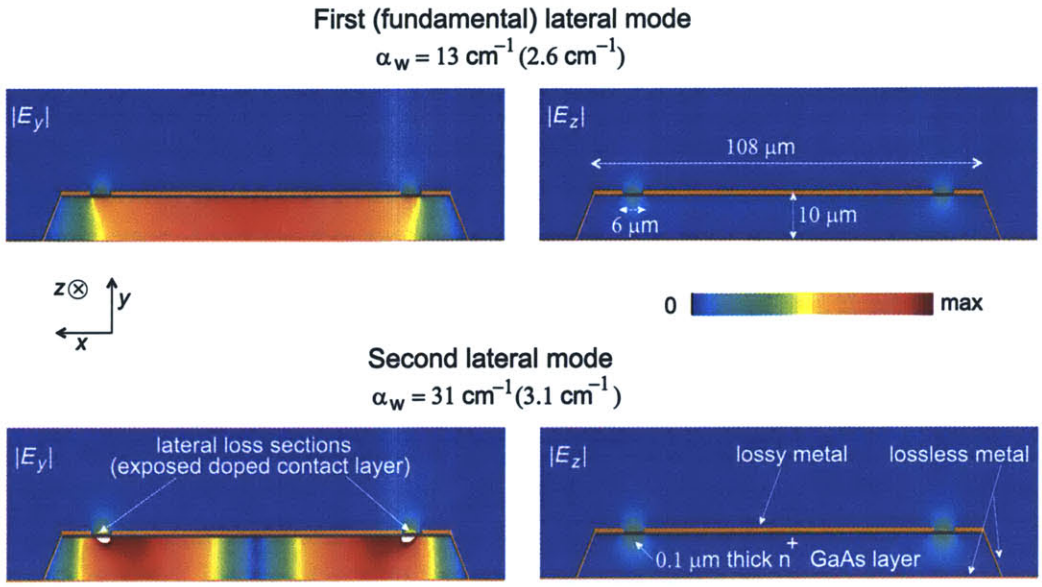


Figure 4-36: Calculated propagation losses for the fundamental and the second order lateral mode in design B of Fig. 4-35 (for an infinite length waveguided). The corresponding mode-shapes are also plotted for the transverse (E_y) and the longitudinal (E_z) components of the electric-field. The values for α_w in the brackets are for the same geometry but without the doped contact layer on the top. These calculations are done for $\nu = 3$ THz. The values of $n = 5 \times 10^{18}$ cm^{-3} and $\tau = 0.1$ ps were used for the Drude parameters of the top n^+ GaAs layer.

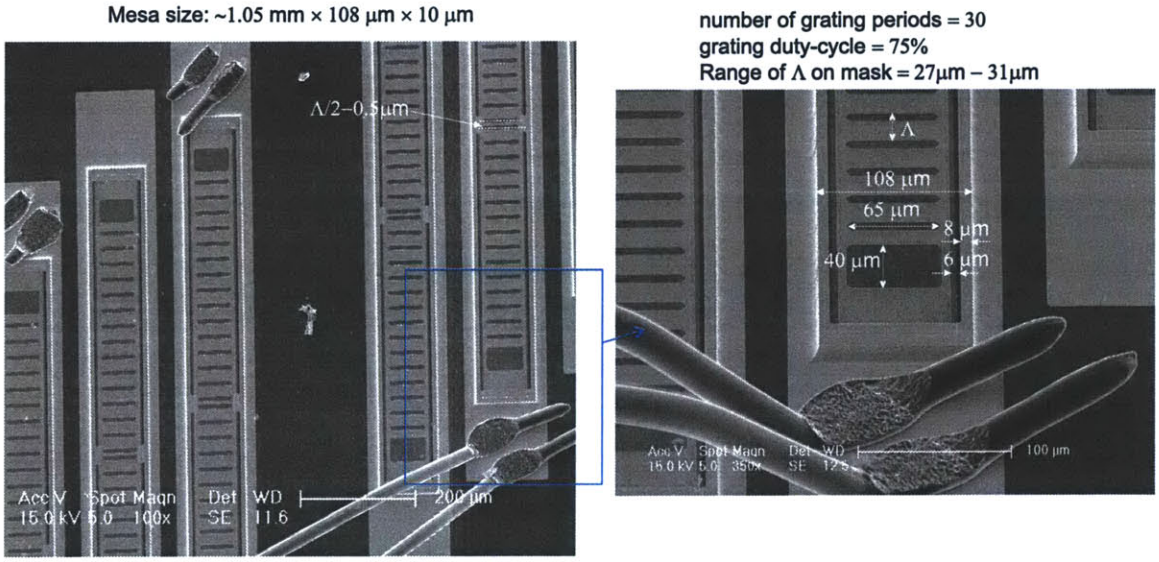


Figure 4-37: SEM images of the design B structures of Fig. 4-35, fabricated with a similar process to that of design A structures in section 4.5.4. An extra mask layer was used to selectively etch the top-contact layer from the grating aperture regions only, to obtain the lateral and longitudinal loss sections respectively.

by finite-element calculations for the propagating mode solutions in such a waveguide in Fig. 4-36. The extra loss for a higher lateral modes is caused by the increased intensity of the longitudinal electric-field component (E_z) inside the loss-sections, since the lateral extent for higher lateral mode is larger, and hence it couples more strongly to the surface-plasmon modes bound to the loss-sections on the either side.

Surface-emitting DFB waveguides with a width of $\sim$ 108 μm were designed and fabricated using the aforementioned technique. In addition to the lateral loss sections, longitudinal loss sections of the type used in Ref. [41] were also implemented. The propagation losses in these longitudinal loss sections is calculated to be $\alpha_w \sim 700/1500/2500$ cm^{-1} at $\nu = 2/3/4$ THz, respectively, and a reflectivity of $\approx 1\%$ (at $\nu \sim 3$ THz) is calculated at the boundaries of the metal-metal region and the longitudinal loss sections. Waveguide fabrication followed steps similar to those outlined in section 4.5.4, except that an extra mask layer was used to remove the top-contact layer in areas under the grating apertures only, as opposed to the devices of section 4.5.4 ("design A" in Fig. 4-35) for which the top-contact layer was removed from everywhere. SEM images for the processed devices are shown in Fig. 4-37. Figure 4-38 shows the experimental results from devices with different grating periods and their corresponding radiation patterns. For most of the devices that were tested, the radiation patterns were also measured with a microbolometer camera [101], from which it could be deduced that none of the devices were lasing in a higher order lateral mode, since strong multiple lobes in the lateral (x) direction were not observed. The camera image in Figure 4-38(d) is significantly narrower in the lateral direction as compared to that for the design A devices of section 4.5.4 (Fig. 4-33), which is one of the primary advantages of having a wider waveguide geometry of design B.

Besides a better radiation pattern, the output power from a wider surface-emitting structure should be much higher, ideally scaled proportionately with its surface-area. However, for the devices discussed in this section, cw values of $\sim 5 \pm 2$ mW were recorded (at 5 K) for most of the devices tested, which is similar to that obtained from the much narrower devices of section 4.5.4 (Fig. 4-34). This is attributed to the overall increase in the waveguide losses even for the fundamental lateral mode due to

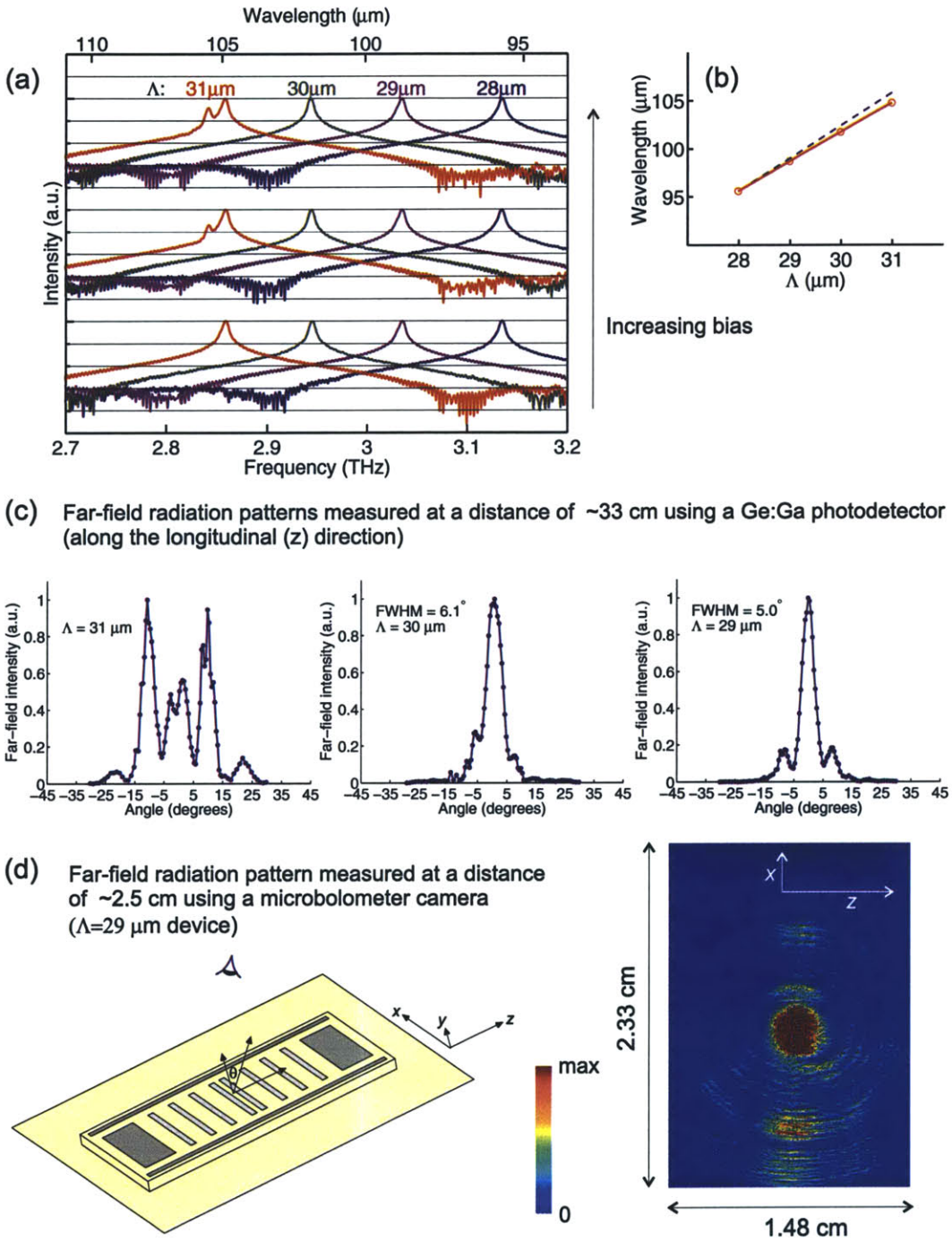


Figure 4-38: 5-K cw spectra for the design B devices with different grating periods and their corresponding radiation patterns. Radiation pattern for $\Lambda = 28$ μm device could not be measured since it was lasing close to a strong atmospheric water absorption line ($\nu = 3.135$ THz). Out of the range of $\Lambda = 27 - 31$ μm devices implemented on the mask, the $\Lambda = 27$ μm devices did not lase, possibly due to the frequency of their grating mode lying outside the gain bandwidth. Plot (b) shows variation of the lasing wavelength (λ_0) with the grating period Λ (in solid red). A line going through origin and corresponding to $n_{\text{eff}} \approx \lambda_0/\Lambda = 3.415$ is also plotted (in dashed blue). The radiation patterns in plots (c) and (d) are taken similarly to the ones in Fig. 4-33.

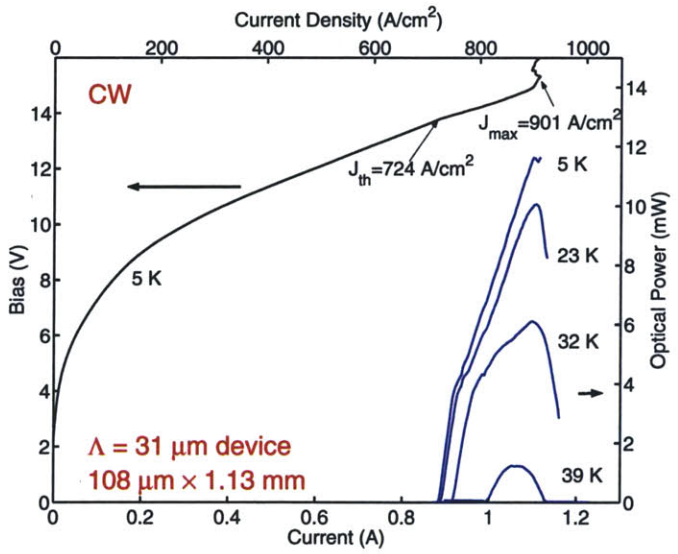


Figure 4-39: CW L-Is and I-V from a $\Lambda = 31$ μm grating device with lateral and longitudinal loss sections (for the waveguide geometry, see design B in Fig. 4-35). Spectra and far-field radiation plots for this device are shown in Fig. 4-38, and finite-element calculations for similar but infinite-length and infinite-width geometries are shown in Fig. 4-36 and Fig. 4-40, respectively. Although pulsed measurements results are not available for this device, $T_{\mathrm{max,pul}} \sim 115$ K was obtained from similar devices.

the introduction of the lateral loss sections, and also due to additional loss caused by the longitudinal loss sections. The CW L-Is and I-V from the $\Lambda = 31$ μm device are shown in Fig. 4-39. A value of $J_{\mathrm{th,5K}} \sim 725$ A/cm^2 was measured for this device as compared to a value of $J_{\mathrm{th,5K}} \sim 560$ A/cm^2 for the design A device in Fig. 4-34, which confirms this argument. As a consequence, the lasing range in current $J_{\mathrm{max}} - J_{\mathrm{th}}$ was smaller, thereby reducing the output optical power.

Unlike the design A devices that had robust single mode operation (section 4.5.4, Fig. 4-32), few of the design B devices lased in two modes such as the $\Lambda = 31$ μm device in Fig. 4-38. This is attributed to longitudinal spatial hole burning in these structures due to a highly non-linear modeshape that arises due to the longitudinal loss sections. Fig. 4-40 shows results from finite-element calculations for an infinite-width waveguide structure with same parameters as those of the $\Lambda = 31$ μm device. The measured radiation pattern for the $\Lambda = 31$ μm device in Fig. 4-38 is similar to what expected by combining the calculated radiation patterns for the first two lowest-loss modes in Fig. 4-40. Moreover, the cw power from that device (11.6 mW,

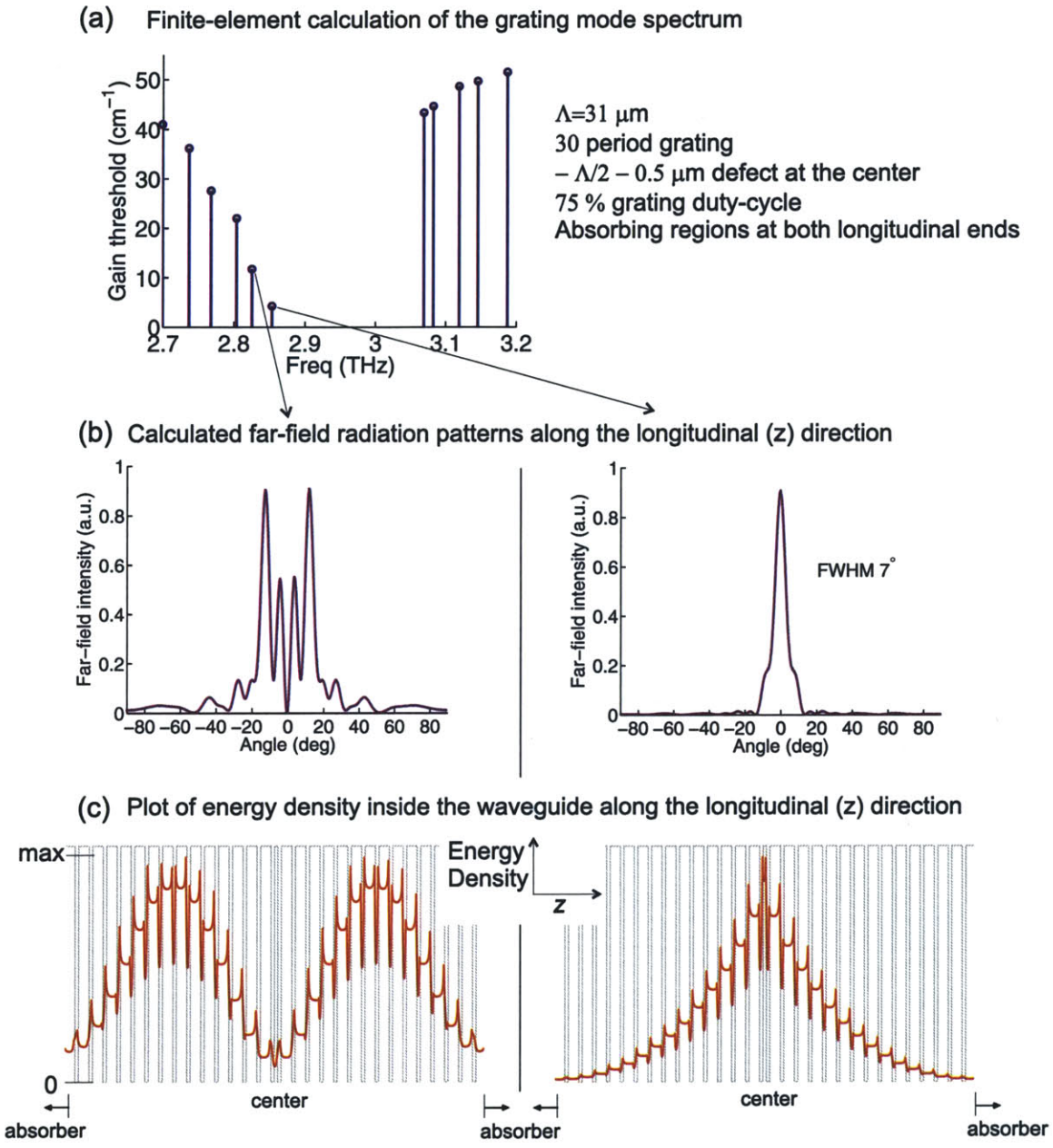


Figure 4-40: Finite-element calculations of the grating mode spectrum, far-field radiation patterns, and mode-shapes for an infinite-width grating structure with longitudinal loss sections similar to that for design B in Fig. 4-35. Note that the threshold gain for the different modes in plot (a) includes a contribution due to the surface out-coupling loss and also due to the loss at the end-absorbing regions (which is a function of the amplitude of the mode at the longitudinal ends and the modeshape along the length of the waveguide). In comparison, the losses calculated for design A in Fig. 4-23, for example, were due to surface-loss only. The loss due to the metal itself is not included in either of these calculations since it is not expected to change significantly from one grating mode to another, and ≈ 5 cm^{-1} (Fig. 4-5).

Fig. 4-39) was approximately twice of that obtained from the single-mode devices, which would be possible if the second lowest-loss mode was lasing since it has a higher surface out-coupling loss as compared to the lowest-loss mode (Fig. 4-40a). These observations suggest that the second lowest-loss grating mode is indeed the one that is excited at the higher bias conditions, which unambiguously confirms the role of spatial hole burning in the spectral behavior of these devices.

Future modifications to design B

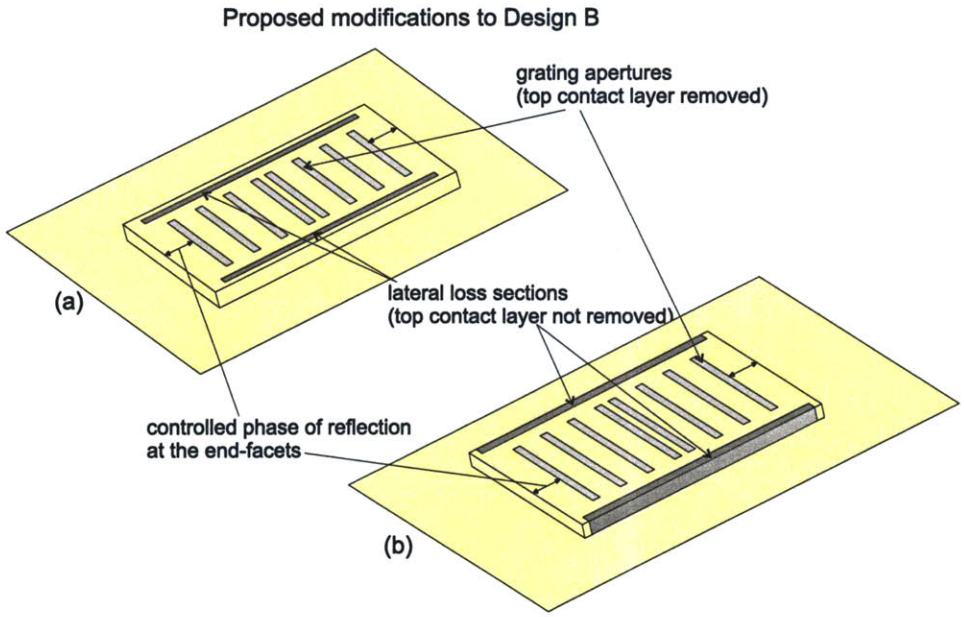


Figure 4-41: Proposed design modifications to the design B of Fig. 4-35 for future implementations.

Even though the design B of Fig. 4-35 suppresses the higher order lateral modes robustly, the devices discussed in this section were the first of their kind that were tried and significant improvements are deemed likely to be possible. In future implementations, the longitudinal loss sections can be replaced by the technique of using controlled phase of reflection at the end-facets, which was described in section 4.5.4. This reduces the over losses since there are no extra losses due to the absorber regions even for waveguides with small cavity lengths. The controlled phase allows the surface loss to be increased controllably, which might lead to greater amount of optical

power from such a design. Moreover, the slope efficiency is likely to be increased due to the more uniform mode-shape in the cavity, which causes the entire cavity to contribute to photon emission rather than only a fraction of it. A more uniform mode shape inside the cavity also greatly reduces the possibility of multiple-mode excitation due to spatial hole burning. The schematics for two of such possible designs are shown in Fig. 4-41. Besides the removal of the longitudinal loss sections, the width of the lateral loss-sections can probably be reduced so as to lower the overall losses in the structure. For example, by narrowing the widths to 3 μm instead of 6 μm for devices discussed in this section, the mode propagation losses in Fig. 4-36 reduce from 13 cm^{-1}/31 cm^{-1} to 3.8 cm^{-1}/6 cm^{-1} for the fundamental/second lateral mode, respectively, whereby the difference in the losses should still be enough to selectively excite the fundamental lateral mode only.

4.6 Conclusions and summary

Terahertz QCLs have been demonstrated with either the metal-metal (MM) waveguides or the semi-insulating surface-plasmon waveguides (SISP). The MM waveguides provide a lower waveguide loss at the terahertz frequencies, which was shown both experimentally and otherwise in section 3.6 of chapter 3.

Besides lower waveguide loss, the MM waveguides also provide strong mode confinement for the terahertz mode, which happens primarily due to subwavelength mode confinement in the vertical dimension. This has some practical advantages in that lasing can be obtained in very small cavities, which makes it relatively easy to obtain robust continuous-wave operation without the need any any special fabrication techniques.

The first part of this chapter described the electromagnetic properties of the MM waveguides at the terahertz frequencies. The surface-plasmon electromagnetic modes that are the underlying mechanism for energy propagation in either of the MM or the SISP waveguides were first discussed in section 4.1. The fact that energy can propagate within a subwavelength dimension for such a mode is the reason why

subwavelength mode propagation is possible in the MM waveguides. However, subwavelength mode confinement causes the modal properties of these waveguides to be significantly different than those for the typical dielectric waveguides that are used at much higher ($\nu \gg 10$ THz, $\lambda \ll 30$ μm) frequencies. These include occurrence of high facet-reflectivities, and relatedly, close to unity mode confinement even in the microcavities (with lateral/longitudinal dimensions that are smaller than the wavelength). These properties were discussed in detail in section 4.3.1. The strong mode confinement causes very low amount of optical power to be out-coupled from the cavities, and resultant radiation patterns are highly divergent. Ring-like diffraction patterns are observed in the output beams of edge-emitting Fabry-Pérot cavity lasers, the experimental evidence and finite-element calculations for which were shown in section 4.3.3.

The second part of this chapter described techniques to implement both first-order and second-order distributed feedback (DFB) gratings in MM waveguides to improve their spectral and radiative properties without significantly affecting the temperature performance (by only fractionally increasing the waveguide losses, if at all). DFB design for these structures is significantly different than that for the dielectric waveguides, since the conventional coupled-mode theory is not directly applicable due to the strong coupling effects that arise because of a subwavelength mode confinement. In this work, finite-element calculations were used as an aid to understand and design the DFB structures.

In section 4.5, the techniques for implementing second-order gratings to realize single-mode surface-emitting operation were presented. Section 4.5.1 discussed the motivation and need to have surface-emitting operation in the MM waveguides. In short, the surface-emitting DFBs provide promise for higher output powers and improved radiation patterns along with single mode operation, without significantly affecting the waveguide losses and hence the temperature performance. However, implementation of second-order distributed feedback met with many challenges during the initial stages of development. For the MM waveguides, strong mode confinement leads to easy excitation of the higher order lateral modes in the rectangular geome-

tries. Moreover, high facet-reflectivities interfere with the distributed feedback of the gratings. To selectively excite the fundamental lateral mode only, two different techniques were developed, as detailed in sections 4.5.4 and 4.5.5 respectively. Also, it was shown in section 4.5.4 that by controlling the phase of reflection at the facets (which is lithographically possible at the terahertz frequencies due to the longer wavelengths), the mode shape inside the waveguide can be engineered to limit the possibility of spatial-hole burning. Robust single-mode operation over a large frequency range was demonstrated using that technique. Additionally, implementation of a central phase-shift in the gratings has yielded single-lobed far-field radiation patterns. The ability to achieve robust-single mode operation while almost maintaining the temperature performance of the Fabry-Pérot cavity lasers, with relatively large amount of optical power and a good beam pattern, makes the surface-emitting MM waveguide DFB lasers ideal for practical applications that are being targeted by the terahertz QCLs.

Bibliography

[1] NASA's Aura mission. http://aura.gsfc.nasa.gov/.

[2] All two- and three-dimensional simulations were carried out in FEMLAB 3.1, a finite-element based partial differential equation software solver from COMSOL, Inc. (http://www.comsol.com).

[3] High-resolution transmission molecular absorption (HITRAN) database, http://www.cfa.harvard.edu/HITRAN.

[4] Optical properties of pure metals and binary alloys. In K.-H Hellwege and O. Madelung, editors, *Landolt-Börnstein*, volume III/15b of *New Series*, chapter 4, pages 210–222. Springer-Verlag, Berlin, Germany, 1985.

[5] The terahertz gap: the generation of far-infrared radiation and its applications. *Phil. Trans. R. Soc. Lond. A*, 362(1815), 2004. Special issue with a compilation of 16 papers.

[6] Photonic terahertz technolgy. *Semicond. Sci. Technol.*, 20(7), 2005. Special issue with a compilation of 22 papers.

[7] Terahertz for military and security application IV. *Proc. SPIE*, 6212, 2006. Special issue with a compilation of 33 papers.

[8] S. Adachi. GaAs, AlAs, and $Al_xGa_{1-x}As$: Material parameters for use in research and device applications. *J. Appl. Phys.*, 58:R1–R28, 1985.

[9] A. J. L. Adam, I. Kašalynas, J. N. Hovenier, T. O. Klaassen, J. R. Gao, E. E. Orlova, B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. Beam patterns of

terahertz quantum cascade lasers with subwavelength cavity dimensions. *Appl. Phys. Lett.*, 88:151105, 2006.

[10] T. Aellen, R. Maulini, R. Terazzi, N. Hoyler, M. Giovannini, J. Faist, S. Blaser, and L. Hvozdara. Direct measurement of the linewidth enhancement factor by optical heterodyning of an amplitude-modulated quantum cascade laser. *Appl. Phys. Lett.*, 89:091121, 2006.

[11] L. Ajili, J. Faist, H. Beere, D. Ritchie, G. Davies, and E. Linfield. Loss-coupled distributed feedback far-infrared quantum cascade lasers. *Electron. Lett.*, 41:53, 2005.

[12] L. Ajili, G. Scalari, J. Faist, H. Beere, E. Linfield, D. Ritchie, and G. Davies. High power quantum cascade lasers operating at $\lambda \cong 87$ and 130 μm. *Appl. Phys. Lett.*, 85:3986, 2004.

[13] J. Alton, S. Barbieri, C. Worral, M. Houghton, H. E. Beere, E. L. Linfield, and D. A. Ritchie. Optimum resonant tunnelling injection and influence of doping density on the performance of THz bound-to-continuum cascade lasers. *Proc. SPIE*, 5727:65, 2005.

[14] T. Ando, A. B. Fowler, and F. Stern. Electronic properties of two-dimensional systems. *Rev. Mod. Phys.*, 54:437, 1982.

[15] A. A. Andronov. From H. Kroemer and B. Lax to modern solid-state (semiconductor) cyclotron resonance masers. *IEEE Trans. Plasma Sci.*, 27:303, 1999.

[16] N. W. Ashcroft and N. D. Mermin. *Solid State Physics.* Harcourt Brace College Publishing, Fort Worth, first edition, 1976.

[17] A. G. Baca, F. Ren, J. C. Zolper, R. D. Briggs, and S. J. Pearton. A survey of ohmic contacts to III-V compound semiconductors. *Thin Solid Films*, 308-309:599, 1997.

[18] F. Banit, S.-C. Lee, A. Knorr, and A. Wacker. Self-consistent theory of the gain linewidth for quantum-cascade lasers. *Appl. Phys. Lett.*, 86:041108, 2005.

[19] W. L. Barnes, A. Dereux, and T. W. Ebbesen. Surface plasmon subwavelength optics. *Nature*, 424:824, 2003.

[20] G. Bastard. Superlattice band structure in the envelope-function approximation. *Phys. Rev. B*, 24:5693, 1981.

[21] M. Beck, D. Hofstetter, T. Aellen, J. Faist, U. Oesterle, M. Ilegems, E. Gini, and H. Melchior. Continuous wave operation of a mid-infrared semiconductor laser at room temperature. *Science*, 295:301, 2002.

[22] A. L. Betz, R. R. Boreiko, B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. Frequency and phase-lock control of a 3 THz quantum cascade laser. *Opt. Lett.*, 30:1837, 2005.

[23] J. S. Blakemore. Semiconducting and other major properties of gallium arsenide. *J. Appl. Phys.*, 53:R123, 1982.

[24] L. L. Bonilla and H. T. Grahn. Non-linear dynamics of semiconductor superlattices. *Rep. Progr. Phys.*, 68:577–683, 2005.

[25] N. Braslau. Alloyed ohmic contacts to GaAs. *J. Vac. Sci. Technol.*, 19:803, 1981.

[26] S. D. Brorson, H. Yokoyama, and E. P. Ippen. Spontaneous emission rate alteration in optical waveguide structures. *IEEE J. Quantum Electron.*, 26:1492, 1990.

[27] M. G. Burt. On the validity and range of applicability of the particle in a box model. *Appl. Phys. Lett.*, 65:717–719, 1994.

[28] H. Callebaut. *Analysis of the electron transport properties in quantum cascade lasers.* PhD dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2006.

[29] H. Callebaut and Q. Hu. Importance of coherence for electron transport in terahertz quantum cascade lasers. *J. Appl. Phys.*, 98:104505, 2005.

[30] H. Callebaut, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno. Analysis of transport properties of terahertz quantum cascade lasers. *Appl. Phys. Lett.*, 83:207, 2003.

[31] H. Callebaut, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno. Importance of electron-impurity scattering for electron transport in terahertz quantum-cascade lasers. *Appl. Phys. Lett.*, 84:645, 2004.

[32] F. Capasso, K. Mohammed, and A. Y. Cho. Resonant tunneling through double barriers, perpendicular quantum transport phenomena in superlattices, and their device applications. *IEEE J. Quantum Electron.*, 22:1853, 1986.

[33] F. Capasso, K. Mohammed, and A. Y. Cho. Sequential resonant tunneling through a multiquantum-well superlattice. *Appl. Phys. Lett.*, 48:478, 1986.

[34] F. Capasso, R. Paiella, R. Martini, R. Colombelli, C. Gmachl, T. L. Myers, M. S. Taubman, R. M. Williams, C. G. Bethea, K. Unterrainer, H. Y. Hwang, D. L. Sivco, A. Y. Cho, A. M. Sergent, H. C. Liu, and E. A. Whittaker. Quantum cascade lasers: ultrahigh-speed operation, optical wireless communication, narrow linewidth, and far-infrared emission. *IEEE J. Quantum Electron.*, 38:511, 2002.

[35] M. Chand and H. Maris. Personal communication.

[36] L. L. Chang, L. Esaki, and R. Tsu. Resonant tunneling in semiconductor double barriers. *Appl. Phys. Lett.*, 24:593, 1974.

[37] C.-Y. Chen, L. Chang, E. Y. Chang, S.-H. Chen, and D.-F. Chang. Thermal stability of Cu/Ta/GaAs multilayers. *Appl. Phys. Lett.*, 77:3367, 2000.

[38] K. N. Chen, A. Fan, C. S. Tan, R. Reif, and C. Y. Yen. Microstructure evolution and abnormal grain growth during copper wafer bonding. *Appl. Phys. Lett.*, 81:3774, 2002.

[39] R. A. Cheville, M. R. Reiten, R. McGowan, and D. R. Grischkowsky. Applications of optically generated terahertz pulses to time domain ranging and scattering. In D. Mittleman, editor, *Sensing with Terahertz Radiation*, pages 237–293. Springer, Berlin, Germany, 2003.

[40] S. L. Chuang. *Physics of Optoelectronic Devices.* Wiley Series in Pure and Applied Optics. John Wiley and Sons, New York, first edition, 1995.

[41] O. Demichel, L. Mahler, T. Losco, C. Mauro, R. Green, A. Tredicucci, J. Xu, F. Beltram, H. E. Beere, D. A. Ritchie, and V. Tamošinuas. Surface plasmon photonic structures in terahertz quantum cascade lasers. *Opt. Express*, 14:5335, 2006.

[42] J. Devenson, R. Teissier, O. Cathabard, and A. N. Baranov. InAs/AlSb quantum cascade lasers emitting below 3 μm. *Appl. Phys. Lett.*, 90:111118, 2007.

[43] L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, M. Lončar, M. Troccoli, and F. Capasso. High-power quantum cascade lasers grown by low-pressure metal organic vapor-phase epitaxy operating in continuous wave above 400 K. *Appl. Phys. Lett.*, 88:201115, 2006.

[44] L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, M. Lončar, M. Troccoli, and F. Capasso. High-temperature continuous wave operation of strain-balanced quantum cascade lasers grown by metal organic vapor-phase epitaxy. *Appl. Phys. Lett.*, 89:081101, 2006.

[45] A. Dobroiu, K. Kawase, Y. Ogawa, M. Yamashita, and C. Otani. THz imaging applications. *Proc. SPIE*, 6038:60380K, 2006.

[46] P. M. Downey, A. D. Jeffries, S. S. Meyer, R. Weiss, F. J. Bachner, J. P. Donnelly, W. T. Lindley, R. W. Mountain, and D. J. Silversmith. Monolithic silicon bolometers. *Appl. Opt.*, 23:910, 1984.

[47] M. S. Dresselhaus. *Solid State Physics, Part II: Optical Properties of Solids*, chapter 10. Massachusetts Institute of Technology, Cambridge, MA. 6.732 course notes, Course VI (Electrical Engineering and Computer Science), MIT.

[48] L. Esaki and R. Tsu. Superlattice and negative differential conductivity in semiconductors. *IBM J. Res. Dev.*, 14:61, 1970.

[49] L. Esaki and R. Tsu. Tunneling in a finite superlattice. *Appl. Phys. Lett.*, 22:562, 1973.

[50] A. Evans, J. Nguyen, S. Slivken, J. S. Yu, S. R. Darvish, and M. Razeghi. Quantum-cascade lasers operating in continuous-wave mode above 90 deg C at $\lambda \sim 5.25$ μm. *Appl. Phys. Lett.*, 88:051105, 2006.

[51] J. Faist, M. Beck, T. Aellen, F. Capasso, and E. Gini. Quantum-cascade lasers based on a bound-to-continuum transition. *Appl. Phys. Lett.*, 78:147, 2001.

[52] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, and A. Y. Cho. Quantum cascade lasers. In H. C. Liu and F. Capasso, editors, *Intersubband Transitions in Quantum Wells: Physics and Device Applications II*, volume 66 of *Semiconductor and Semimetals*, pages 1–83. Academic Press, San Diego, 2000.

[53] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, S. N. G. Chu, and A. Y. Cho. Quantum-well intersub-band electroluminescent diode at $\lambda = 5$ μm. *Electron. Lett.*, 29:2230, 1993.

[54] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho. Quantum cascade laser. *Science*, 264:553, 1994.

[55] N. Finger, W. Schrenk, and E. Gornik. Analysis of tm-polarized dfb laser structures with metal surface gratings. *IEEE J. Quantum Electron.*, 36:780, 2000.

[56] B. A. Foreman. First-principles envelope-function theory for lattice-matched semiconductor heterostructures. *Phys. Rev. B*, 72:165345, 2005.

[57] A. Friedrich, G. Boehm, M.-C. Amann, and G. Scarpa. Quantum-cascade lasers without injector regions operating above room temperature. *Appl. Phys. Lett.*, 86:161114, 2005.

[58] A. Friedrich, C. Huber, G. Boehm, and M.-C. Amann. Low-threshold room-temperature operation of injectorless quantum-cascade lasers: influence of doping density. *Electron. Lett.*, 42:1228, 2006.

[59] J. R. Gao, J. N. Hovenier, Z. Q. Yang, J. J. A. Baselmans, A. Baryshev, M. Hajenius, T. M. Klapwijk, A. J. L. Adam, T. O. Klaassen, B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. A terahertz heterodyne receiver based on a quantum cascade laser and a superconducting bolometer. *Appl. Phys. Lett.*, 86:244104, 2005.

[60] C. Gmachl, F. Capasso, E. E. Narimanov, J. U. Nockel, A. D. Stone, J. Faist, D. L. Sivco, and A. Y. Cho. High-power directional emission from microlasers with chaotic resonators. *Science*, 280:1556, 1998. Also see the cited references.

[61] C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho. Recent progress in quantum cascade lasers and applications. *Rep. Progr. Phys.*, 64:1533–1601, 2001.

[62] C. Gmachl, A. Straub, R. Colombelli, F. Capasso, D. L. Sivco, A. M. Sergent, and A. Y. Cho. Single-mode, tunable distributed-feedback and multiple-wavelength quantum cascade lasers. *IEEE J. Quantum Electron.*, 38:569, 2002.

[63] S. Golka, C. Pflügl, W. Schrenk, and G. Strasser. Quantum cascade lasers with lateral double-sided distributed feedback grating. *Appl. Phys. Lett.*, 86:111103, 2005.

[64] A. Gordon, C. Y. Wang, L. Diehl, A. Belyanin, F. Capasso, F. X. Kärtner, and J. Faist. Multimode dynamics and mode locking phenomena in QCLs. In *The 2nd International Workshop on Quantum Cascade Lasers*, Brindisi, Italy, Sep. 6–9 2006. Presentation available on the conference website: http://www.rle.mit.edu/2006qcl/.

[65] M. K. Gunde and M. Maček. Infrared optical constants and dielectric response functions of silicon nitride and oxynitride films. *Phys. Stat. Sol. (a)*, 183:439, 2001.

[66] P. L. Hagelstein, S. D. Senturia, and T. P. Orlando. *Introductory Applied Quantum and Statistical Mechanics*, chapter 22. Wiley Series in Pure and Applied Optics. John Wiley and Sons, Hoboken, NJ, 2004.

[67] E. E. Haller. Advanced far-infrared detectors. *Infrared Phys.*, 35:127, 1994.

[68] M. Helm, P. England, E. Colas, F. DeRosa, and Jr. S. J. Allen. Intersubband emission from semiconductor superlattices excited by sequential resonant tunneling. *Phys. Rev. Lett.*, 63:74, 1989.

[69] C. M. Herzinger, C. C. Lu, T. A. DeTemple, and W. C. Chew. The semiconductor waveguide facet reflectivity problem. *IEEE J. Quantum Electron.*, 29:2273, 1993.

[70] D. Hofstetter, J. Faist, M. Beck, and U. Oesterle. Surface-emitting 10.1 μm quantum cascade distributed feedback lasers. *Appl. Phys. Lett.*, 75:3769, 1999.

[71] B. B. Hu and M. C. Nuss. Imaging with terahertz waves. *Opt. Lett.*, 20:1717, 1995.

[72] Q. Hu, B. S. Williams, S. Kumar, H. Callebaut, S. Kohen, and J. L. Reno. Resonant-phonon-assisted thz quantum-cascade lasers with metal-metal waveguides. *Semicond. Sci. Technol.*, 20:S228, 2005.

[73] H.-W. Hübers, S. G. Pavlov, A. D. Semenov, R. Köhler, L. Mahler, A. Tredicucci, H. E. Beere, D. A. Ritchie, and E. H. Linfield. Terahertz quantum cascade laser as local oscillator in a heterodyne receiver. *Opt. Express*, 13:5890, 2005.

[74] H.-W. Hübers, S. G. Pavlov, and V. N. Shastin. Terahertz lasers based on germanium and silicon. *Semicond. Sci. Technol.*, 20:S221, 2005. Also see the cited references.

[75] T. Ikegami. Reflectivity of mode at facet and oscillation mode in double-heterostructure injection lasers. *IEEE J. Quantum Electron.*, 8:470, 1972.

[76] M. Inguscio, G. Moruzzi, K. M. Evenson, and D. A. Jennings. A review of frequency measurements of optically pumped lasers from 0.1 to 8 THz. *J. Appl. Phys.*, 60:R161, 1986.

[77] T. Itoh and R. Mittra. Resonance conditions of open resonators at microwave frequencies. *IEEE Trans. Microwave Theory Tech.*, 22:99, 1974.

[78] Y. L. Ivanov and Vasiley. Cyclotron resonance laser in p-Ge. *Sov. Tech. Phys. Lett.*, 9:264, 1983.

[79] R. W. Jackson and D. M. Pozar. Full-wave analysis of microstrip open-end and gap discontinuities. *IEEE Trans. Microwave Theory Tech.*, 33:1036, 1985.

[80] Z. Jiang and X.-C. Zhang. Free-space electro-optic techniques. In D. Mittleman, editor, *Sensing with Terahertz Radiation*, pages 155–192. Springer, Berlin, Germany, 2003.

[81] J. Jin. *The Finite Element Method in Electromagnetics.* Wiley-IEEE Press, second edition, 2002.

[82] K. Kawase, Y. Ogawa, H. Minamide, and H. Ito. Terahertz parametric sources and imaging applications. *Semicond. Sci. Technol.*, 20:S258, 2005.

[83] R. F. Kazarinov and R. A. Suris. Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice. *Sov. Phys. Semicond.*, 5:707, 1971.

[84] K. Kim, K. Hess, and F. Capasso. Monte Carlo study of electron heating and enhanced thermionic emission by hot phonons in heterolayers. *Appl. Phys. Lett.*, 52:1167, 1988.

[85] D. Kleppner. Inhibited spontaneous emission. *Phys. Rev. Lett.*, 47:233, 1981.

[86] H. Kogelnik and C. V. Shank. Coupled-wave theory of distributed feedback lasers. *J. Appl. Phys.*, 43:2327, 1972.

[87] S. Kohen. *Electromagnetic modeling of terahertz quantum cascade laser waveguides and resonators.* MS dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2004.

[88] S. Kohen, B. S. Williams, and Q. Hu. Electromagnetic modeling of terahertz quantum cascade laser waveguides and resonators. *J. Appl. Phys.*, 97:053106, 2005.

[89] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi. Terahertz semiconductor-heterostructure laser. *Nature*, 417:156, 2002.

[90] J. Kröll, J. Darmo, A. Benz, G. Fasching, and K. Unterrainer. Phase resolved stimulated emission from THz QCLs. In *The 2nd International Workshop on Quantum Cascade Lasers*, Brindisi, Italy, Sep. 6–9 2006. Presentation available on the conference website: http://www.rle.mit.edu/2006qcl/.

[91] S. Kumar, Q. Qin, B. S. Williams, Q. Hu, Z. R. Wasilewski, X. Wu, and H. C. Liu. Quantum-cascade lasers with one-well injector operating at 1.59 THz ($\lambda =$ 188.5 μm). *to be presented at the Conference on Lasers and Electro-Optics, Baltimore, Maryland, 6–11 May 2007.*

[92] S. Kumar, B. S. Williams, , Q. Hu, and J. L. Reno. 1.9 THz quantum-cascade lasers with one-well injector. *Appl. Phys. Lett.*, 88:121123, 2006.

[93] S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno. First-order edge-emitting and second-order surface-emitting distributed feedback terahertz quantum cascade lasers. *presented at the 8th International Conference on Intersubband Transitions in Quantum Wells, Cape Cod, Massachusetts, 11–16 Sept. 2005.*

[94] S. Kumar, B. S. Williams, S. Kohen, Q. Hu, and J. L. Reno. Continuous-wave operation of terahertz quantum-cascade lasers above liquid-nitrogen temperature. *Appl. Phys. Lett.*, 84:2494, 2004.

[95] S. Kumar, B. S. Williams, Q. Qin, A. W. M. Lee, Q. Hu, and J. L. Reno. Surface-emitting distributed feedback terahertz quantum-cascade lasers in metal-metal waveguides. *Opt. Express*, 15:113, 2007.

[96] B. Lax. [111] direct transition exciton and magnetoreflection in germanium. *Phys. Rev. Lett.*, 4:511, 1960.

[97] R. P. Leavitt. Empirical two-band model for quantum wells and superlattices in an electric field. *Phys. Rev. B*, 44:11270, 1991.

[98] A. W. M. Lee, Q. Qin, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno. Unpublished results.

[99] A.W Lee and Q. Hu. Real-time, continuous-wave terahertz imaging by use of a microbolometer focal-plane array. *Opt. Lett.*, 30:2563, 2005.

[100] A.W.M. Lee, Q. Qin, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno. Real-time terahertz imaging over a standoff distance (> 25 m). *Appl. Phys. Lett.*, 89:141125, 2006.

[101] A.W.M. Lee, B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. Real-time imaging using a 4.3-thz quantum cascade laser and a 320 x 240 microbolometer focal-plane array. *IEEE Photon. Technol. Lett.*, 18:1415, 2006.

[102] C. C. Lee, C. Y. Wang, and G. Matijasevic. Au-In bonding below the eutectic temperature. *IEEE Trans. Comp, Hybrids, Manuf. Technol.*, 16:311, 1993.

[103] S.-C. Lee, F. Banit, M. Woerner, and A. Wacker. Quantum mechanical wavepacket transport in quantum cascade laser structures. *Phys. Rev. B*, 73:245320, 2006.

[104] S.-C. Lee and A. Wacker. Nonequilibrium green's function theory for transport and gain properties of quantum cascade structures. *Phys. Rev. B*, 66:245314, 2002.

[105] S. Li, G. Witjaksono, S. Macomber, and D. Botez. Analysis of surface-emitting second-order distributed feedback lasers with central grating phaseshift. *IEEE J. Sel. Topics Quantum Electron.*, 9:1153, 2003.

[106] H. C. Liu, M. Buchanan, and Z. R. Wasilewski. How good is the polarization selection rule for intersubband transitions? *Appl. Phys. Lett.*, 72:1682, 1998.

[107] H. C. Liu, M. Wächter, D. Ban, Z. R. Wasilewski, M. Buchanan, G. C. Aers, J. C. Cao, S. L. Feng, B. S. Williams, and Q. Hu. Effect of doping concentration on the performance of terahertz quantum-cascade lasers. *Appl. Phys. Lett.*, 87:111104, 2005.

[108] P. Lugli, C. Jacoboni, L. Reggiani, and P. Kocevar. Monte Carlo algorithm for hot phonons in polar semiconductors. *Appl. Phys. Lett.*, 50:1251, 1987.

[109] M. S. Lundstrom, M. A. Stettler, T. T. Herman, and P. E. Dodd. *FISH1D User's Manual.* Purdue University, Layfayette, 1990.

[110] Jr. M. F. Pereira, S.-C. Lee, and A. Wacker. Controlling many-body effects in the midinfrared gain and terahertz absorption of quantum cascade laser structures. *Phys. Rev. B*, 69:205310, 2004.

[111] L. Mahler, R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, D. A. Ritchie, and A. G. Davies. Single-mode operation of terahertz quantum cascade lasers with distributed feedback resonators. *Appl. Phys. Lett.*, 84:5446, 2004.

[112] R. Maulini, A. Mohan, M. Giovannini, J. Faist, and E. Gini. External cavity quantum-cascade laser tunable from 8.2 to 10.4 μm using a gain element with a heterogeneous cascade. *Appl. Phys. Lett.*, 88:201113, 2006.

[113] M. McLennan and S. Datta. *SEQUAL User's Manual.* Purdue University, Layfayette, 1989.

[114] C. A. Mead and W. G. Spitzer. Fermi level position at metal-semiconductor interfaces. *Phys. Rev.*, 134:A713, 1964.

[115] I. Mehdi. THz local oscillator technology. *Proc. SPIE*, 5498:103, 2004.

[116] D. Mittleman. Terahertz imaging. In D. Mittleman, editor, *Sensing with Terahertz Radiation*, pages 117–153. Springer, Berlin, Germany, 2003.

[117] W. J. Moore and R. T. Holm. Infrared dielectric constant of gallium arsenide. *J. Appl. Phys.*, 80:6939, 1996.

[118] E. R. Mueller, Jr. W. E. Robotham, R. P. Meisner, R. A. Hart, J. Kennedy, and L. A. Newman. 2.5 THz laser local oscillator for the EOS CHEM 1 satellite. In *Proceedings of the Ninth International Space Terahertz Technology Symposium*, pages 563–572, Pasadena, CA, Mar. 17–19 1998.

[119] S. Mukamel. *Principles of nonlinear optical spectroscopy.* Oxford University Press, New York, 1995.

[120] N. Mustafa, L. Pesquera, C. Y. L. Cheung, and K. A. Shore. Terahertz bandwidth prediction for amplitude modulation response of unipolar intersubband semiconductor lasers. *IEEE Photon. Technol. Lett.*, 11:527, 1999.

[121] D. F. Nelson, R. C. Miller, and D. A. Kleinman. Band nonparabolicity effects in semiconductor quantum wells. *Phys. Rev. B*, 35:7770, 1987.

[122] R. J. Noll and S. H. Macomber. Analysis of grating surface emitting lasers. *IEEE J. Quantum Electron.*, 26:456, 1990.

[123] M. A. Ordal, R. J. Bell, R. W. Alexander-Jr., L. L. Long, and M. R. Querry. Optical properties of fourteen metals in the infrared and far infrared: Al, co, cu, au, fe, pb, mo, ni, pd, pt, ag, ti, v, and w. *Appl. Opt.*, 24:4493, 1985.

[124] E. E. Orlova, J. N. Hovenier, T. O. Klaassen, I. Kašalynas, A. J. L. Adam, J. R. Gao, T. M. Klapwijk, B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. Antenna model for wire lasers. *Phys. Rev. Lett.*, 96:173904, 2006.

[125] R. Paiella, R. Martini, F. Capasso, C. Gmachl, H. Y. Hwang, D. L. Sivco, J. N. Baillargeon, A. Y. Cho, E. A. Whittaker, and H. C. Liu. High-frequency modulation without the relaxation oscillation resonance in quantum cascade lasers. *Appl. Phys. Lett.*, 79:2526, 2001.

[126] E. D. Palik. Gallium arsenide. In E. D. Palik, editor, *Handbook of optical constants of solids*, Academic Press Handbook Series, pages 429–443. Academic Press, Orlando, 1985.

[127] 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.

[128] C. Pflügl, M. Austerer, W. Schrenk, S. Golka, G. Strasser, R. P. Green, L. R. Wilson, J. W. Cockburn, A. B. Krysa, and J. S. Roberts. Single-mode surface-emitting quantum-cascade lasers. *Appl. Phys. Lett.*, 86:211102, 2005.

[129] E. M. Purcell. Spontaneous emission probabilities at radio frequencies. *Phys. Rev.*, 69:681, 1946.

[130] H. Raether. *Surface Plasmons on Smooth and Rough Surfaces and on Gratings.* Springer, Berlin, 1988.

[131] F. Rana and R. J. Ram. Current noise and photon noise in quantum cascade lasers. *Phys. Rev. B*, 65:125313, 2002.

[132] D. G. Revin, J. W. Cockburn, M. J. Steer, R. J. Airey, M. Hopkinson, A. B. Krysa, L. R. Wilson, and S. Menzel. InGaAs/AlAsSb/InP quantum cascade lasers operating at wavelengths close to 3 μm. *Appl. Phys. Lett.*, 90:021108, 2007.

[133] P. L. Richards. Bolometers for infrared and millimeter waves. *J. Appl. Phys.*, 76:1, 1994.

[134] 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.

[135] M. Rochat, J. Faist, M. Beck, U. Oesterle, and M. Ilegems. Far-infrared ($\lambda = 88\ \mu$m) electroluminescence in a quantum cascade structure. *Appl. Phys. Lett.*, 73:3724, 1998.

[136] A. Rogalski. Infrared detectors: status and trends. *Prog. Quantum Electron.*, 27:59–210, 2003.

[137] G. Scalari, L. Ajili, J. Faist, H. Beere, E. Linfield, D. Ritchie, and G. Davies. Far-infrared ($\lambda \cong 87\mu$m) bound-to-continuum quantum- cascade lasers operating up to 90 K. *Appl. Phys. Lett.*, 82:3165, 2003.

[138] G. Scalari, S. Blaser, J. Faist, H. Beere, E. Linfield, D. Ritchie, and G. Davies. Terahertz emission from quantum cascade lasers in the quantum hall regime: Evidence for many body resonances and localization effects. *Phys. Rev. Lett.*, 93:237403, 2004.

[139] G. Scalari, C. Walther, J. Faist, H. Beere, and D. Ritchie. Electrically switchable, two-color quantum cascade laser emitting at 1.39 and 2.3 THz. *Appl. Phys. Lett.*, 88:141102, 2006.

[140] M. Schubert and F. Rana. Analysis of terahertz surface emitting quantum-cascade lasers. *IEEE J. Quantum Electron.*, 42:257, 2006.

[141] J. N. Schulman. Extension of Tsu-Esaki model for effective mass effects in resonant tunneling. *Appl. Phys. Lett.*, 72:2829, 1998.

[142] M. P. Semtsiv, M. Wienold, S. Dressler, and W. T. Masselink. Short-wavelength ($\lambda \sim 3.05\ \mu$m) InP-based strain-compensated quantum-cascade laser. *Appl. Phys. Lett.*, 90:51111, 2007.

[143] K. Shimoda. *Introduction to Laser Physics*, chapter 7. Springer-Verlag, Berlin, second edition, 1983.

[144] N. Shtinkov and S. J. Vlaev. Intersubband absorption in n-type GaAs/AlGaAs (001) quantum wells: A tight-binding study. *Phys. Stat. Sol. (b)*, 241:R11, 2004.

[145] P. H. Siegel. Terahertz technology. *IEEE Trans. Microwave Theory Tech.*, 50:910, 2002.

[146] P. H. Siegel. Terahertz technology in biology and medicine. *IEEE Trans. Microwave Theory Tech.*, 52:2438, 2004.

[147] 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, 1994.

[148] C. Sirtori, F. Capasso, J. Faist, A. L. Hutchinson, D. L. Sivco, and A. Y. Cho. Resonant tunneling in quantum cascade lasers. *IEEE J. Quantum Electron.*, 34:1722, 1998.

[149] C. Sirtori, J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, S. N. G. Chu, and A. Y. Cho. Continuous wave operation of midinfrared (7.4-8.6 μm) quantum cascade lasers up to 110 K temperature. *Appl. Phys. Lett.*, 68:1745, 1995.

[150] C. Sirtori, H. Page, C. Becker, and V. Ortiz. GaAs-AlGaAs quantum cascade lasers: physics, technology and prospects. *IEEE J. Quantum Electron.*, 38:547, 2002.

[151] 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.

[152] 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.

[153] M. Tacke. Lead-salt lasers. *Phil. Trans. R. Soc. Lond. A*, 359:547, 2001.

[154] A. Tredicucci, F. Capasso, C. Gmachl, D. Sivco, A. L. Hutchinson, and A. Y. Cho. High performance interminiband quantum cascade lasers with graded superlattices. *Appl. Phys. Lett.*, 73:2101, 1998.

[155] J. Ulrich, R. Zobl, N. Finger, K. Unterrainer, G. Strasser, and E. Gornik. Terahertz quantum cascade structures: Intra- versus interwell transition. *Appl. Phys. Lett.*, 77:1928, 2000.

[156] J. Ulrich, R. Zobl, K. Unterrainer, G. Strasser, and E. Gornik. Magnetic-field-enhanced quantum-cascade emission. *Appl. Phys. Lett.*, 76:19, 2000.

[157] M. S. Vitiello, G. Scamarcio, V. Spagnolo, B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. Measurement of subband electronic temperatures and population inversion in THz quantum-cascade lasers. *Appl. Phys. Lett.*, 86:111115, 2005.

[158] I. Vurgaftman and J. R. Meyer. TE- and TM-polarized roughness-assisted free-carrier absorption at midinfrared and terahertz wavelengths. *Phys. Rev. B*, 60:14294, 1999.

[159] I. Vurgaftman and J. R. Meyer. Photonic-crystal distributed-feedback quantum cascade lasers. *IEEE J. Quantum Electron.*, 38:592, 2002.

[160] Jr. W. P. Harokopus, L. P. B. Katehi, W. Y. Ali-Ahmad, and G. M. Rebeiz. Surface wave excitation from open microstrip discontinuities. *IEEE Trans. Microwave Theory Tech.*, 39:1098, 1991.

[161] C. Walther, G. Scalari, J. Faist, H. Beere, and D. Ritchie. Low frequency terahertz quantum cascade laser operating from 1.6 to 1.8 THz. *Appl. Phys. Lett.*, 89:231121, 2006.

[162] L. C. Wang, X. Z. Wang, S. S. Lau, T. Sands, W. K. Chen, and T. F. Kuech. Stable and shallow PdIn ohmic contacts to *n*-GaAs. *Appl. Phys. Lett.*, 56:2129, 1990.

[163] T. B. Wang, Z. Z. Shen, R. Q. Ye, X. M. Xie, F. Stubhan, and J. Freytag. Die bonding with Au/In isothermal solidification technique. *J. Electron. Mat.*, 29:443, 2000.

[164] J. Ward, E. Schlecht, G. Chattopadhyay, A. Maestrini, J. Gill, F. Maiwald, H. Javadi, and I. Mehdi. Capability of THz sources based on Schottky diode frequency multiplier chains. *IEEE MTT-S International Microwave Symposium Digest*, 3:1587–111, 2004.

[165] L. C. West and S. J. Eglash. First observation of an extremely large-dipole infrared transition within the conduction band of a GaAs quantum well. *Appl. Phys. Lett.*, 46:1156–1158, 1985.

[166] B. S. Williams. *Terahertz quantum cascade lasers.* PhD dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2003.

[167] B. S. Williams, H. Callebaut, Q. Hu, and J. L. Reno. Magnetotunneling spectroscopy of resonant anticrossing in terahertz intersubband emitters. *Appl. Phys. Lett.*, 79:4444, 2001.

[168] B. S. Williams, H. Callebaut, S. Kumar, Q. Hu, and J. L. Reno. 3.4-THz quantum cascade laser based on longitudinal-optical-phonon scattering for depopulation. *Appl. Phys. Lett.*, 82:1015, 2003.

[169] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno. 3.4 THz quantum cascade laser operating above liquid nitrogen temperature. *Electron. Lett.*, 39:915, 2003.

[170] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno. Terahertz quantum-cascade laser at $\lambda \approx 100$ μm using metal waveguide for mode confinement. *Appl. Phys. Lett.*, 83:2124, 2003.

[171] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno. Terahertz quantum-cascade laser operating up to 137 K. *Appl. Phys. Lett.*, 83:5142, 2003.

[172] B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. Resonant-phonon terahertz quantum-cascade laser operating at 2.1 THz ($\lambda \cong 141$ μm). *Electron. Lett.*, 40:431, 2004.

[173] B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. Distributed-feedback terahertz quantum-cascade lasers with laterally corrugated metal waveguides. *Opt. Lett.*, 30:2909, 2005.

[174] B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. Operation of terahertz quantum-cascade lasers at 164 K in pulsed mode and at 117 K in continuous-wave mode. *Opt. Express*, 13:3331, 2005.

[175] B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno. High-power terahertz quantum-cascade lasers. *Electron. Lett.*, 42:89, 2006.

[176] B. S. Williams, S. Kumar, Q. Qin, Q. Hu, and J. L. Reno. Terahertz quantum cascade lasers with double-resonant-phonon depopulation. *Appl. Phys. Lett.*, 88:261101, 2006.

[177] B. S. Williams, B. Xu, Q. Hu, and M. R. Melloch. Narrow-linewidth terahertz intersubband emission from three-level systems. *Appl. Phys. Lett.*, 75:2927, 1999.

[178] R. Williams. *Modern GaAs processing methods*, chapter 5, page 97. Artech House, Boston, second edition, 1990.

[179] J. M. Woodall, J. L. Freeouf, G. D. Pettit, T. Jackson, and P. Kirchner. Ohmic contacts to n-GaAs using graded band gap layers of $Ga_{1-x}In_xAs$ grown by molecular beam epitaxy. *J. Vac. Sci. Technol.*, 19:626, 1981.

[180] C. Worral, J. Alton, M. Houghton, S. Barbieri, H. E. Beere, D. Ritchie, and C. Sirtori. Continuous wave operation of a superlattice quantum cascade laser emitting at 2 thz. *Opt. Express*, 14:171, 2006.

[181] D. C. F. Wu, R. C. Rudduck, and E. L. Pelton. Application of a surface integration technique to parallel-plate waveguide radiation-pattern analysis. *IEEE Trans. Antennas Propagat.*, 17:280, 1969.

[182] B. Xu. *Development of intersubband terahertz lasers using multiple quantum well structures.* PhD dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1998.

[183] B. Xu, Q. Hu, and M. R. Melloch. Electrically pumped tunable terahertz emitter based on intersubband transition. *Appl. Phys. Lett.*, 71:440, 1997.

[184] R. Q. Yang, J. M. Xu, and M. Sweeny. Selection rules of intersubband transitions in conduction-band quantum wells. *Phys. Rev. B*, 50:7474, 1994.

[185] A. Yariv. *Quantum Electronics*, chapter 8. John Wiley & Sons, New York, second edition, 1975.

[186] A. Yariv. *Quantum Electronics.* John Wiley & Sons, New York, second edition, 1979.

[187] A. Yariv. *Quantum Electronics*, chapter 20. John Wiley & Sons, New York, third edition, 1989.

[188] J. S. Yu, S. Slivken, A. Evans, S. R. Darvish, J. Nguyen, and M. Razeghi. High-power $\lambda \sim 9.5$ μm quantum-cascade lasers operating above room temperature in continuous-wave mode. *Appl. Phys. Lett.*, 88:091113, 2006.

[189] X.-C. Zhang. THz wave imaging. In *ICIS '06: 30th International Congress of Imaging Science*, page 498, Rochester, NY, May 2006.