UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

HIGH-PERFORMANCE INAS-BASED INTERBAND CASCADE LASERS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

YUCHAO JIANG
Norman, Oklahoma
2016

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/34604

HIGH-PERFORMANCE INAS-BASED INTERBAND CASCADE LASERS

A DISSERTATION APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

_____
Dr. Rui Q. Yang, Chair

_____
Dr. Michael B. Santos

_____
Dr. Zhisheng Shi

_____
Dr. Matthew B. Johnson

_____
Dr. James Sluss

© Copyright by YUCHAO JIANG 2016
All Rights Reserved.

To my mother, Zi-Zhen Huang

Acknowledgements

Life is like a cascade of transitions: in each stage, I accumulate my energy with the help of others and make the final jump, emitting the wave that resonates with my soul. Each "degree" is a new challenge. In retrospect, it was in overcoming various barriers that I obtained a deeper insight into the world. I like what Albert Einstein said, "a hundred times a day I remind myself that my inner and outer life depend on the labors of other men, living and dead, and that I must exert myself in order to give in the measure as I have received and am still receiving."

First and foremost, I would like to thank my advisor, Prof. Rui Q. Yang. He provided me a precious opportunity to learn how an original concept was developed into practical devices. He was generous to offer his knowledge and experience for me to complete various daunting projects. I have acquired both theoretical understanding and experimental skills that are crucial to my future career. His foresight and high standard have driven me out of a student's comfort zone to think more thoroughly and act more professionally. I also thank him for giving me freedoms to explore my thoughts and ideas, which is the most intriguing part of scientific research.

I am very appreciative of the efforts made by my group members and collaborators. I would like to thank Lu Li and Hao Ye, who performed the material growth and fabrication, for providing reliable devices presented in this work. I learned a lot from their responsible attitudes and attentions to detail. I would like to thank Hossein Lotfi for his insightful questions. I was inspired by his carefulness when we did some experiments together. I would also like to thank Zhaobing Tian, Robert Hinkey,

Zhuowei Yin, Lin Lei, SM Shazzad Rassel, Tetsuya Mishima, James Gupta, David Jui-Yang Feng and Prof. Fow-Sen Choa for their help in the research.

I would like to thank Prof. Matthew Johnson and Prof. Michael Santos for constructive conversations and for providing resources and expertise in device fabrication and material growth. I also thank Prof. Zhisheng Shi and Prof. James Sluss for serving on my committee.

The last miles are always the toughest. I am deeply thankful to my friends for illuminating me to walk through the darkest time in this final year. I would like to thank my running pals Echo and Holly for their enthusiastic encouragements. I enjoyed every weekend training and three marathon races, which blew away my worries and recharged my spirit. I would like to thank Dr. Michele Eodice and William Dower for organizing a writing group to improve my writing. Special thanks go to Ying Lan for setting a writing camp, a neat and undisturbed heaven for me to concentrate on my dissertation and nourish my thoughts.

I want to express my deepest gratitude to Prof. Chang-Zhi Guo. Before I came to U.S., he planted a seed of curiosity of laser research in my mind during my master's study. He was so patient to make detailed comments on each of my misunderstanding in the homework and was always encouraging when I stepped further. I learned the basic modeling skills from him, and he was kind enough to show me his Matlab codes at his apartment. I also thank Prof. Gang Yu for encouraging me to study abroad.

Lastly, I would like to thank my family for their unconditional love and support over the years.

Table of Contents

Acknowledgements iv

List of Tables xi

List of Figures xiv

Abstract xix

Chapter 1 Introduction 1

1.1 Lasers 1

1.2 Semiconductor lasers 3

1.3 Mid-IR semiconductor lasers 5

1.4 Brief history of IC lasers 8

Chapter 2 Fundamentals of interband cascade lasers 13

2.1 Basis in laser operation 13

2.1.1 Gain analysis 13

2.1.2 Rate equation 17

2.2 Gain coefficient in semiconductor materials 22

2.2.1 Probability inversion 22

2.2.2 Quasi-Fermi levels in bulk and QW material 23

2.2.3 Gain coefficient 27

2.3 Active region material and band structure 29

2.3.1 Type-II QW materials 30

2.3.2 Band structure modeling 31

2.3.3 Two-band model and formalism 35

2.3.4 Kane's two-band model 37

2.4 Optical wave modeling 38
2.4.1 Optical wave equations 38
2.4.2 Light propagating in a multilayer structure 40
2.4.3 Waveguide mode 45
2.4.4 Far-field patterns 47
Chapter 3 Material growth, fabrication and laser characterizations 49
3.1 Material growth 49
3.2 Device fabrication 51
3.3 Laser characterizations 52
3.3.1 Light-current-voltage characteristic 52
3.3.2 Double-modulation technique in pulse spectra 55
3.3.3 Pulse power calibration 59
3.3.4 Gain spectrum 61
3.3.5 Thermal resistance 65
Chapter 4 Far-field patterns and beam qualities of plasmon-enhanced waveguide interband cascade lasers 70
4.1 Waveguide modeling 70
4.1.1 Plasmon-enhanced waveguide 70
4.1.2 Calculation of refractive index 73
4.1.3 Device structure and waveguide simulation 79
4.2 Far-field pattern measurement 83
4.2.1 Experimental setup 83
4.2.2 Experimental and simulated results 84

4.3 Beam quality 89
4.3.1 Gaussian beam 90
4.3.2 Brightness 92
4.3.3 Beam quality factors for IC lasers 93
Chapter 5 InAs-based interband cascade lasers above room temperature 97
5.1 InAs-based IC lasers with asymmetric cladding layers 97
5.1.1 Device structure and waveguide modeling 97
5.1.2 Experimental results 99
5.2 InAs-based IC lasers with heavily-doped injector 102
5.2.1 Carrier rebalancing 102
5.2.2 Band structure and experimental results 107
5.3 InAs-based IC lasers above room temperature in CW mode 112
5.3.1 SL as the inner cladding layer 112
5.3.2 Experimental results 116
Chapter 6 Long-wavelength InAs-based interband cascade lasers 120
6.1 Challenges of long-wavelength interband cascade lasers 120
6.2 Interband cascade lasers at wavelength beyond 11 μm 122
6.2.1 Design considerations and simulations 122
6.2.2 Experimental results 125
6.2.3 Negative differential resistance 126
6.2.4 Amplified spontaneous emission around 13 μm 128
6.3 Room temperature IC lasers beyond 6 μm 131
6.3.1 Active region design 131

6.3.2 Waveguide simulation 133
6.3.3 Experimental results 135
6.3.4 Further optimization and current status 137
Chapter 7 Electrically widely-tunable interband cascade lasers 140
7.1 Tuning mechanisms 140
7.2 Tunable IC laser structures and modeling 144
7.3 Experimental results 147
Chapter 8 InAs-based single-mode distributed feedback interband cascade lasers 152
8.1 Introduction 152
8.2 Laser structure and fabrication 153
8.3 Experimental results and discussion 156
Chapter 9 Conclusions and future work 165
9.1 Dissertation summary 165
9.2 Future work 167
References 172
Appendix A: Fourier transform in lock-in amplifier 194
Appendix B: Light collecting efficiency due to beam divergence 197
Appendix C: Plasmonic waveguide 203
References 208
Appendix D: Coupling coefficient in DFB lasers 210
References 212
Appendix E: Publication list 213

List of Tables

Table 3-1. Thermal resistance of each layer from cascade region to bonding area 67

Table 7-1. Tuning ranges and rates at 80 K..... 151

List of Figures

Figure 1-1. Stimulated emission [5] 2

Figure 1-2. Absorption spectra for some invisible gas molecules [21] 6

Figure 1-3. Conventional interband laser vs. quantum cascade laser [32] 7

Figure 1-4. Band-edge diagram of one cascade stage in an IC laser design 10

Figure 1-5. Evolving of the threshold current densities in IC lasers over the years..... 11

Figure 2-1. Modal gain as a function of injection 15

Figure 2-2. Pumping and carrier lifetime in a three-level laser [74] 16

Figure 2-3. Dynamic process in a semiconductor laser before the steady state. 21

Figure 2-4. Quasi-Fermi levels as a function of carrier density in the bulk material. 25

Figure 2-5. Comparison of quasi-Fermi levels between bulk and QW materials 26

Figure 2-6. Gain spectra and spontaneous emission spectra for GaAs bulk and QW 29

Figure 2-7. Band alignment for (a) type-I QW; (b) type-II QW [81]..... 30

Figure 2-8. Propagation of light through a multilayer structure..... 42

Figure 3-1. Double modulation setup..... 57

Figure 3-2. Experimental setup for pulse power measurement..... 59

Figure 3-3. Gain spectra for an IC laser 65

Figure 4-1. Two types of waveguides for IC lasers..... 71

Figure 4-2. Comparison of optical intensity and index profiles for two waveguides 72

Figure 4-3. Energy E(k) and effective mass m*(k) as a function of wavevector..... 76

Figure 4-4. Complex refractive index of InAs 78

Figure 4-5. Device structure for narrow-ridge IC laser. Device A has a one metal contact trench in the center; Devices B and C have two metal contact trenches near the edges. 81
Figure 4-6. One-dimensional optical wave and refractive index profile 82
Figure 4-7. Two experimental setups for far-field measurements: (a) translation method. (b) rotation method. 84
Figure 4-8. (a) Measured 2D far-field profile for device A at 18 mA. (b) Far-field profiles along vertical direction (top panel) and lateral direction (bottom panel). 85
Figure 4-9. Example of the effective index method applied to device A. 86
Figure 4-10. Lateral far-field profiles of device A at different currents. Inset: far-field (dark black line) consists of 50% fundamental mode and 50% first mode 87
Figure 4-11. Vertical and lateral far-field profiles of four devices 88
Figure 4-12. Spread of the Gaussian beam [155] 90
Figure 4-13. (a) Divergence of a plane wave due to diffraction ;(b) Intensity distribution in the focal plane of a lens for a divergent beam. 92
Figure 4-14. Beam quality factors for devices from four representative wafers 96
Figure 5-1. Optical mode and refractive index profiles along the growth direction 99
Figure 5-2. (a) Temperature dependent threshold current densities in pulsed mode for lasers from 4 wafers; (b) Power-current curves for lasers from 3 wafers. 101
Figure 5-3. Carrier density dependent occupation probability for active QWs with injectors undoped and heavily doped 104
Figure 5-4. Normalized differential gain in InAs/GaSb QWs with injectors undoped and heavily doped 105

Figure 5-5. Band diagrams of one complete cascade stage and eigenenergy alignments for (a) R083 (b) reference structure in [51] 108
Figure 5-6. Characterizations for device R083BA_3H. (a) pulse spectra; (b)CW spectra; (c) IVP and differential resistance in CW operation; (d) temperature dependence of lasing wavelengths and threshold current densities..... 110
Figure 5-7. CW spectra for device R083NR_6C..... 111
Figure 5-8. Complex refractive index of InAs and SL..... 112
Figure 5-9. Three types of waveguides in IC lasers: GaSb spacer/SL cladding (Black solid), InAs spacer/SL inner cladding/n^{++} InAs outer cladding (blue dotted), and InAs spacer/ n^{++} InAs cladding (red dashed) 116
Figure 5-10. (a) Threshold current density vs. heat-sink temperature (b) Current-voltage-light characteristics for a narrow-ridge device in CW operation. Two insets show CW and pulsed lasing spectra, respectively..... 118
Figure 6-1. Calculated threshold current densities as a function of wavelength. 121
Figure 6-2. Band diagrams of one complete cascade stage and eigenenergy alignment for (a) R069 (b)R110 122
Figure 6-3. Waveguide comparison between R069 (black solid) and R110 (red dotted). Complex refractive indices at 80K are corrected by temperature coefficients. 124
Figure 6-4. (a) Light-current curve for a broad-area IC laser in pulsed mode. The arrows point to threshold currents where spectra were taken. Inset: pulse spectra. (b) Current-voltage-light characteristics in CW operation. Insets: CW lasing spectra at 80 K to 97 K; emission spectra at 80 K with different injection currents. 126

Figure 6-5. (a) Differential resistance as a function of dc current at temperatures of 80 K to 220 K.(b) Injected currents and voltages corresponding to the negative differential resistance at different temperatures. 128
Figure 6-6. Amplified spontaneous emissions at 80K: (a) R111 (b) R112 131
Figure 6-7. (a) Transition energy and overlap integral as a function of InAs QW thickness; (b) The intensity of overlap integral as a function of transition energy. 132
Figure 6-8. Waveguide structure and optical profile for R139. 135
Figure 6-9. Characteristics for R139: (a) pulsed spectra; (b) temperature dependence of wavelength and threshold current densities. 136
Figure 6-10. Temperature dependence of wavelength and threshold current densities for R139 and R141 137
Figure 6-11. Waveguide optimization for 6.5 μm InAs-based IC lasers. 138
Figure 6-12. Maximum pulsed operating temperatures for IC lasers beyond 6 μm..... 139
Figure 7-1. Transition dipole and wavefunction overlap for active regions with one QW(a) and two QWs(b) 143
Figure 7-2. Fundamental electron and hole wavefunctions in the active region of structure R64 at different electric fields. The band alignment and the diagonal interband transition (dashed arrow) are plotted at F=75 kV/cm. 145
Figure 7-3. Calculated fundamental transition energy and wavefunction overlap as a function of electric field for active regions with different numbers of QWs. 146
Figure 7-4. Light-current-voltage characteristics of an IC laser (20 μm × 2.1 mm) from R064 at different heat-sink temperatures in CW mode. 148
Figure 7-5. CW tuning spectra of an IC laser (15 μm × 1.8 mm) from R64 at 80 K. .. 149

Figure 7-6. Tunable lasing spectra of an IC laser in both CW and pulsed modes. 150
Figure 8-1. SEM images of IC laser structures. (a) a top view of the DFB gratings covered with metal and without metal; (b) zoom-in of the DFB grating; (c) cross-sectional view of a laser; (d) the top view of a laser mesa 156
Figure 8-2. Emission spectra from two IC lasers at 80 K for several currents using (a) device C; and (b) device H, with a log intensity plot inset. 158
Figure 8-3. Sub-threshold emission spectra from a BA DFB IC laser with a grating period of ~640-nm at heat-sink temperatures ranging from 150 to 180 K. 159
Figure 8-4. Current-tuning CW spectra of a DFB IC laser at 160 K, with an SMSR of ~ 30 dB. Inset is a plot of intensity with log scale. 160
Figure 8-5. Single-mode lasing wavelengths of a DFB IC laser as a function of injected currents at T=150-180 K. A total tuning range of 16 nm was achieved for a single device. Only data with spectral SMSR larger than 20 dB are included. 161
Figure 8-6. Current-voltage-light characteristics for a DFB IC laser. Inset: CW spectra at temperatures from 150 to 180 K. 162
Figure 8-7. Comparison of FP and DFB lasers at various operating temperatures: (a) CW lasing wavelengths and (b) threshold current densities 164
Figure 9-1.Calculated temperature dependency of the threshold current density 171

Abstract

Currently, there are only two types of mid-infrared lasers that are capable of continuous-wave (CW) operation above room temperature: quantum cascade (QC) lasers and interband cascade (IC) lasers. Both of them share the cascade feature for carrier recycling. The most successful QC lasers, based on the inter-subband transition and the well-established InGaAs/InAlAs/InP material system, are able to deliver several watts of optical power. In contrast, IC lasers, based on the interband transition and the unique InAs/GaSb/AlSb type-II broken-bandgap material system, have the threshold power density more than an order of magnitude lower than that of QC lasers (*e.g.*, 0.3 kW/cm^2 vs. 11 kW/cm^2). As a result, IC lasers become a better solution for low-power applications in the mid-infrared region.

GaSb-based IC lasers have achieved the best performance around 3.7 μm with a threshold current density as low as 100 A/cm^2 at 300 K. However, their waveguide cladding layers, consisting of thick InAs/AlSb superlattice, have a low thermal conductivity and are challenging to grow by molecular beam epitaxy. These problems become more severe at longer lasing wavelengths due to the requirement of thicker cladding layers. InAs-based IC lasers, utilizing highly doped InAs as the optical cladding layer, have been developed to address these issues. The goal of this dissertation is to use modeling and experiments to explore several aspects of InAs-based IC lasers, including far-field patterns, high-temperature operation, long-wavelength operation, wide-tunability, and single frequency mode operation.

The beam quality is critical for the laser application. The higher-order spatial modes naturally appear when the laser ridge is wider than the lasing wavelength in the

medium. For InAs-based IC lasers with a thin top cladding layer, the top contact configuration can have a major influence on the spatial modes, which are observed in the measurement of far-field patterns. The physical origin is identified by waveguide modeling based on an effective index method.

Radical design innovations, including "shortened injector" and "carrier rebalancing," have significantly improved the performance of both GaSb-based and InAs-based IC lasers. Furthermore, a hybrid waveguide, consisting of an inner cladding layer with InAs/AlSb superlattice and an outer cladding layer with highly doped InAs, has significantly increased the modal gain of InAs-based IC lasers. As a result, CW operations above room temperature have been achieved at wavelengths of 4.6~4.8 μm. The threshold current density, 247 A/cm^2 at 300 K in pulsed mode, is the lowest ever reported among the mid-infrared semiconductor lasers at similar wavelengths. The pulsed operating temperature is as high as 377 K.

Long-wavelength operations are vigorously explored. With the hybrid waveguide mentioned above, the lasing temperature reaches 324 K at a wavelength of 6.4 μm. Further design improvement and optimization are presented. In addition, the lasing wavelength is extended to 11.2 μm at 130 K. Several things are found to hinder the progress. The waveguide loss is dramatically increased, mainly because the lasing wavelength approaches the plasmon wavelength of the heavily-doped InAs. The negative differential resistance is observed and may be related to the unexpected high threshold.

A wide tuning range is highly desirable for many applications such as spectroscopy and biochemical analysis. A repeatable, large electrical tunable range of

180 cm^{-1} (or 900 nm in wavelength near $\lambda \sim 7$ μm) is achieved by a novel active region consisting of three InAs quantum wells. This challenges the conventional idea that the carrier density pinning at the threshold level would not allow a significant tuning by Stark effect. The gain analysis, based on the calculation of the field dependent wavefunction overlap, well explains the physical mechanism. This strategy is very useful for the design of tunable lasers.

For sensitive detection of important gas molecules such as carbonyl sulfide/COS (4.5 μm), single-mode distributed feedback IC lasers are highly desirable for tunable laser absorption spectroscopy. A grating is patterned using interference lithography to etch through the thin top cladding into the top spacer layer of the IC laser structure. Single-mode emission with a side mode suppression ratio of 30 dB is obtained in continuous wave operation at temperatures up to 180 K near 4.5 μm. A total tuning range of 16 nm is achieved for a single device, with a temperature-tuning rate of 0.4 nm/K and a current-tuning rate of 0.016 nm/mA. The impact of the grating on device performance is evaluated and discussed in comparison with Fabry–Perot lasers.

Chapter 1 Introduction

This chapter will briefly review the developments of lasers, semiconductor lasers, mid-infrared semiconductor lasers and interband cascade lasers. Each section is a subset of the previous one.

1.1 Lasers

The Nobel Prize in Physics in 1964 was awarded to Charles Townes, Nikolay Basov, and Alexander Prokhorov, "for fundamental work in the field of quantum electronics, which has led to the construction of oscillators and amplifiers based on the maser-laser principle." [1] These pioneers built the first masers, which are laser's counterpart in the microwave range. For example, Townes's maser was made in 1954 by isolating excited ammonia molecules into a cavity resonant at 24 GHz, or 1.25 cm in wavelength [2] [3].

It is well known that laser is an acronym for "Light **Amplification** by Stimulated Emission of Radiation." The concept of **stimulated emission** was first proposed by Albert Einstein in 1916 [4]. From the original concept to the realization by Townes' work, it took scientists 38 years to realize how to achieve population inversion for stimulated emission. It is less well known that a laser itself is essentially an **oscillator** (which makes it a "loser") because it generates its own beam internally and sets its own wavelength by the laser material and cavity. In contrast, an optical amplifier, also called as laser amplifier, merely amplifies the new incoming light through a single pass without a resonant cavity.

Stimulated emission can be illustrated in Figure 1-1. In a two-level system, the energy difference between the upper energy level E_2 and the lower energy level E_1 is hν. An incoming photon with the same energy hv can cause a particle in the upper level to drop to the lower level, and the lost energy is transferred to a new photon with the identical phase, frequency, polarization and direction of the travel as the incoming photon. This is in contrast to spontaneous emission, which generates photons at random intervals irrelevant to the incoming photon.

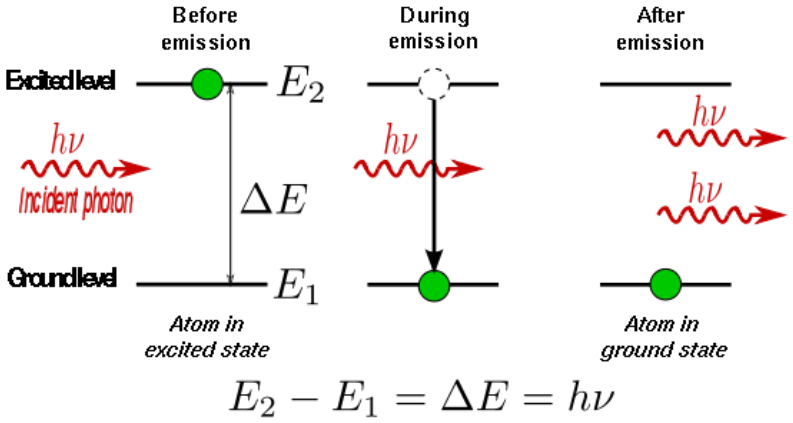


Figure 1-1. Stimulated emission [5]

In 1960, Theodore Maiman demonstrated the first laser in human history [6]. He chose a synthetic chromium-doped ruby crystal as the laser material, instead of alkali vapor suggested by Townes [7]. Pumped by a high-power flash lamp, this ruby laser with two facets coated by silver had an emission wavelength of 694.3 nm at room temperature.

Since this breakthrough, researchers have been making marvelous progress on lasers, and greatly expanded the wavelength range by trying a wide variety of materials and novel structures. The performances of lasers have been sufficiently improved for commercial applications and now exist everywhere in our daily life. Examples are bar

code scanners, CD and DVD players, laser printers, laser pointers, and laser surgery. The fiber-optic communications, which have an unprecedented impact on modern civilization, use InGaAsP lasers as a source and erbium fiber as an optical amplifier.

1.2 Semiconductor lasers

Whether a material is suitable to construct a laser core depends on its capability to sustain a population inversion under external excitation. Regarding light-emitting medium, lasers can be classified into several types: a gas laser, a liquid laser (organic dye laser), a solid-state laser, and a semiconductor laser.

Although the semiconductor lasers are technically solid-state devices, they are entirely different from the solid-state lasers, because they operate in a fundamentally different way. In solid-state lasers, the materials are dielectric solids, so the only way to inject energy is with an optical pump. Typical examples of solid-state lasers are the ruby laser which uses sapphire as the host material and chromium as the active species, and the fiber laser which uses glass as the host material and erbium as the active species. The emission wavelengths of solid-state lasers are mainly determined by the active species dispersed in the host materials, leading to sharp emission peaks. In contrast, the semiconductor lasers can be electrically pumped, and their emission wavelengths rely on the energy bands formed by the host material (semiconductor). As a result, the gain spectra in semiconductor lasers are intrinsically broader, and the energy efficiencies are higher than that in solid-state lasers. Due to their many advantages, such as being compact, power-efficient and low-cost, the semiconductor lasers have been occupying about half of the market share for decades.

A semiconductor laser is also known as a laser **diode**. This can trace back to its early development when the active medium is formed by a p-n junction, similar to a semiconductor diode used in the electric circuit. Most laser diodes are composed of the "direct bandgap" materials such as GaAs and InP. Without bias, a built-in electric field is formed near the p-n junction, leading to a depletion region. A forward electrical bias can cancel the built-in field and make electrons and holes recombine at the junction, generating photons with energy equal to the bandgap. The p-n junction, also known as a homojunction, is created by p-type and n-type doping in the same semiconductor material. This is in contrast to a heterojunction, also called as a heterostructure, which is formed by two materials of different bandgaps. In 1963 Herbert Kroemer [8] and Zhore I. Alferov [9] independently proposed double heterostructure (DH) lasers, which significantly enhanced the confinements for both carriers and photons in the active region. The first continuous-wave (CW) room-temperature semiconductor lasers were achieved based on GaAs/AlGaAs DH, with threshold current density around 1000 A/cm^2 at 300 K [10 ,11]. The Nobel Prize in Physics in 2000 was awarded to Kroemer and Alferov for their contribution in DH lasers.

According to de Broglie hypothesis, all matter can exhibit wave-like behavior. The wavelength of an electron in semiconductor can be related to its momentum by

$$\lambda = \frac{h}{p} = \frac{2\pi\hbar}{\sqrt{2E_p m_0}} \approx 25\ nm \qquad (1.2.1)$$

where the band structure parameter E_p is 21~29 eV for most binary semiconductors and will be further explained in section 2.2.3. If the core layer of DH is made sufficiently thin, *e.g.*, less than 25 nm, the quantum effect will manifest in that the electron energy is quantized in the growth direction. Thus, a DH laser becomes a quantum well (QW)

laser and the corresponding densities of states are stair-like. Consequently, the differential gain is increased, and the lasing threshold can be reduced. Although the threshold current density for the first QW lasers was about 3000 A/cm^2 [12], continuous improvements, such as graded-index separate confinement structure [13], managed to bring it down to 160 A/cm^2 [14]. Extremely low threshold current densities of 65 A/cm^2 [15] and 56 A/cm^2 [16] were achieved by lowering the valence band effective mass via compressive strain [17]. Nowadays, most commercial semiconductor lasers are based on the QWs due to their superior performances.

1.3 Mid-IR semiconductor lasers

The III-V direct bandgap semiconductor lasers have successfully covered broad spectra ranging from blue ray to near infrared (IR) with high performance. However, in the mid-wavelength IR (*i.e.*, 3-8 μm) and long-wavelength IR (*i.e.*, 8-15 μm) region, no laser had achieved CW operation above room temperature until 2002 [18]. Coherent mid-IR light sources are very useful for many applications. The top application is gas-sensing. Many invisible gas molecules have their most sensitive absorption within this range, as shown in Figure 1-2. These molecules can be detected by tunable diode laser absorption spectroscopy (TDLAS) with detection limit around 1 ppb, which is about three orders of magnitude more sensitive than that in near-infrared. Other applications for mid-IR lasers include stand-off detection [19] and free space optical communication [20].

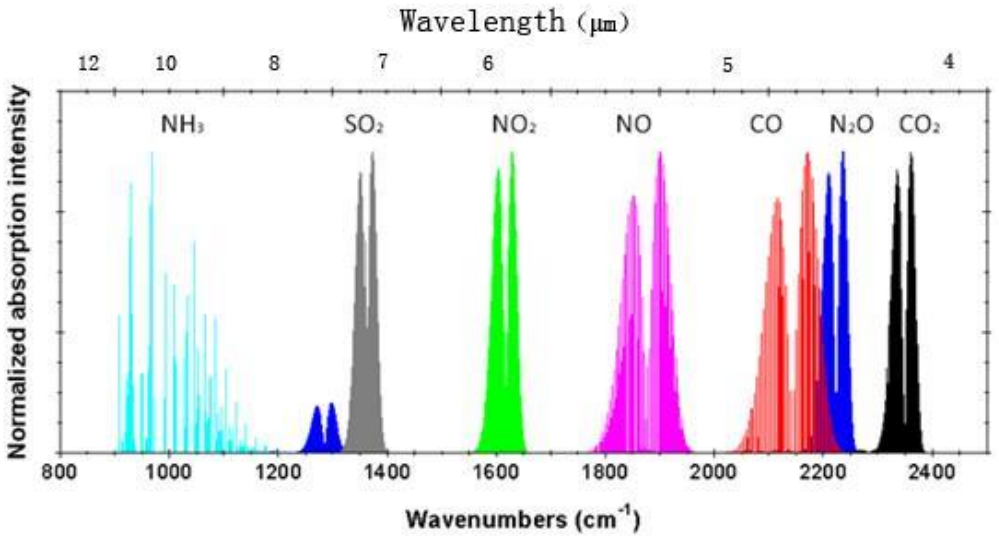


Figure 1-2. Absorption spectra for some invisible gas molecules [21]

For emission at mid-IR wavelengths, traditional type-I interband laser diodes are based on narrow bandgap IV-VI materials (also called lead salt materials) [22]. The Auger coefficients in IV-VI materials are $1\sim8\times10^{-28}$ cm^6/s [23,24], comparable to $2\sim3\times10^{-28}$ cm^6/s in type-II InAs/GaSb QWs [25]. However, the maximum CW operating temperature, achieved by optically pumped PbSrSe/PbSe QW disk lasers at 4.3 μm, is only 275 K [26]. Further improvement on performance is impeded by the relatively low thermal conductivity of the material and inferior epitaxial material quality [27].

In 1994, the demonstration of quantum cascade (QC) lasers by Capasso's group [28] had radically changed the landscape of mid-IR semiconductor lasers. The photon generation in a QC laser is achieved through the use of intersubband transitions in a repeated stack of multiple QWs. The photon energy is determined by the quantized states within the conduction band and can be tailored over a wide range from mid-IR to terahertz region. Although the fast phonon scattering (on the order of 1 ps) was

detrimental to establishing a population inversion, the researchers overcame this difficulty by innovative active region designs such as "bound-to-continuum transition" [29] and "double-phonon resonance" [30]. Continuous wave operation above room temperature was achieved in 2002, at an emission wavelength of 9.1 μm and with a pulsed threshold current density of 3100 A/cm^2 at 292 K [18]. Tremendous research efforts on the QC laser have been creating every new record. The state-of-the-art QC lasers have a wall-plug-efficiency (WPG) of 21% with a maximum output power of 5.1 W in CW operation at room temperature [31].

Figure 1-3 shows the difference between conventional interband laser, where an electron and a hole recombines to generate a photon, and QC laser, where only electrons participate in the inter-subband transition and each electron can produce multiple photons in the cascade structure.

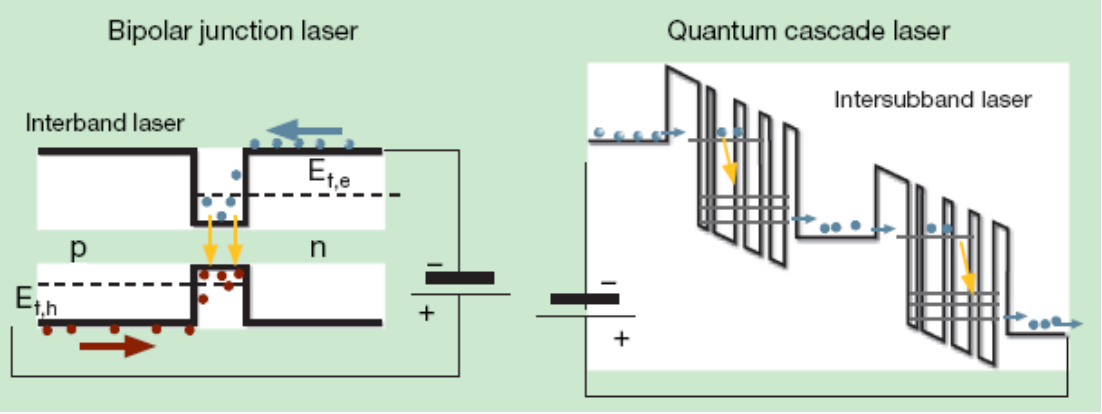


Figure 1-3. Conventional interband laser vs. quantum cascade laser [32]

Another competitive contender in the mid-IR region is interband cascade (IC) laser [33,34], which has combined the advantages of both conventional interband laser and QC laser. On the one hand, the interband transition has a longer carrier lifetime than intersubband transition (*e.g.*, 1 ns vs. 1 ps), resulting in a much lower threshold current

density. On the other hand, the cascade connection, equivalent to an in-series circuit [35], has a significant advantage over the non-cascade interband laser in terms of Auger recombination. The series connection trades low threshold voltage for low threshold current. With fewer injected carriers, the Auger recombination is significantly reduced. Another advantage is similar to power transmission line, the higher voltage and lower current can reduce the ohmic loss in the device.

A brief review of the development of IC lasers will be given in next section. An interesting note is that the cascade concept is also applicable to interband lasers at shorter wavelengths. As early as 1984, van der Ziel and Tsang connected three GaAs laser cores in series by highly doped tunnel junctions (Esaki junction) for the purpose of higher output [36]. The so-called "bipolar cascade" lasers with 2 or 3 stages were reported at wavelengths ranging from 1 to 2 μm [37-41]. The active regions in these lasers were also connected by highly-doped tunnel junctions. But the tunnel junctions introduced so much optical loss that they didn't outperform their counterparts with one stage. In addition, GaN light emitting diodes in a cascade style were claimed to overcome efficiency droop due to a reduction of threshold current [42].

1.4 Brief history of IC lasers

The prototype of IC laser, proposed by Rui Q. Yang in 1994 [43,44], is an alternative infrared laser to quantum cascade laser [28]. The backbone of the IC laser is the type-II broken-bandgap heterostructure, InAs/Ga(In)Sb QW, which facilitates the interband transition and enables the recycling of carriers. In 1997, the first demonstration of an IC laser was reported with a wavelength of 3.8 μm [45]. Soon a

hole injector consisting of one GaInSb QW was proposed to prevent carrier tunneling leakage [46].

Further development was carried out by Yang and his collaborators at several institutions, as well as by groups at Naval Research Laboratory and the University of Wurzburg. On the one hand, InAs/GaSb/AlSb material system is less established than InP and GaAs, so related technologies are explored to overcome the practical difficulties in material growth and device fabrication. On the other hand, several key changes on the design have significantly improved the laser performance.

Figure 1-4 shows a modern-day style of the cascade structure, which is divided into three parts: an active region, a hole injector, and an electron injector. For the active region, two InAs QWs consisting of a "W" shape, previously used in type-II QW lasers [47], was introduced to IC lasers to improve the wavefunction overlap as well as the optical gain [48]. For the hole injector, a second hole QW was added to suppress the leakage current further [49]. For the electron injector, two strategies are have contributed to substantial improvements. The first one is called "shortened electron injector" [50], which reduces the thicknesses of InAs QWs in the electron injector, leading to an enhanced optical mode intensity in the active region and a higher voltage efficiency. The second one is called "carrier rebalancing" [51], which increases the doping concentration in the electron injector by one order of magnitude, resulting in a significantly reduced threshold current density. However, the physical mechanism is not conclusively known. This will be further discussed in section 5.2.1.

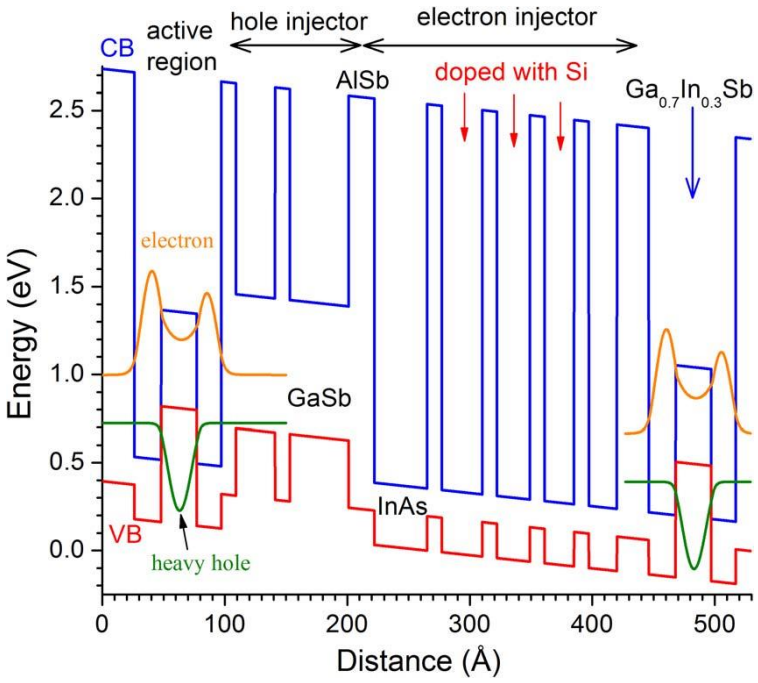


Figure 1-4. Band-edge diagram of one cascade stage in an IC laser design

Over the last decade, with the improvement of theoretical understanding, material quality, and device fabrication, the pulsed threshold current densities of room temperature (300 K) IC lasers at wavelengths below 4 μm had been reduced by more than one order of magnitude, as shown in Figure 1-5. The first room temperature IC laser was reported in 2002, with a rather high threshold current density of 6900 A/cm^2 [52]. Soon it was reduced to 1000 A/cm^2 [53], which was already lower than the best QC laser at similar wavelengths [54,55]. The reduction to about 600 A/cm^2 [56] enabled the IC lasers to work in CW operation at 264 K, accessible by a simple one-stage thermoelectric cooler (TEC) [57]. Then the distributed feedback (DFB) IC lasers were developed [58] and delivered to 2007 NASA flight project and installed in the Mars rover Curiosity for methane detection [59,60].

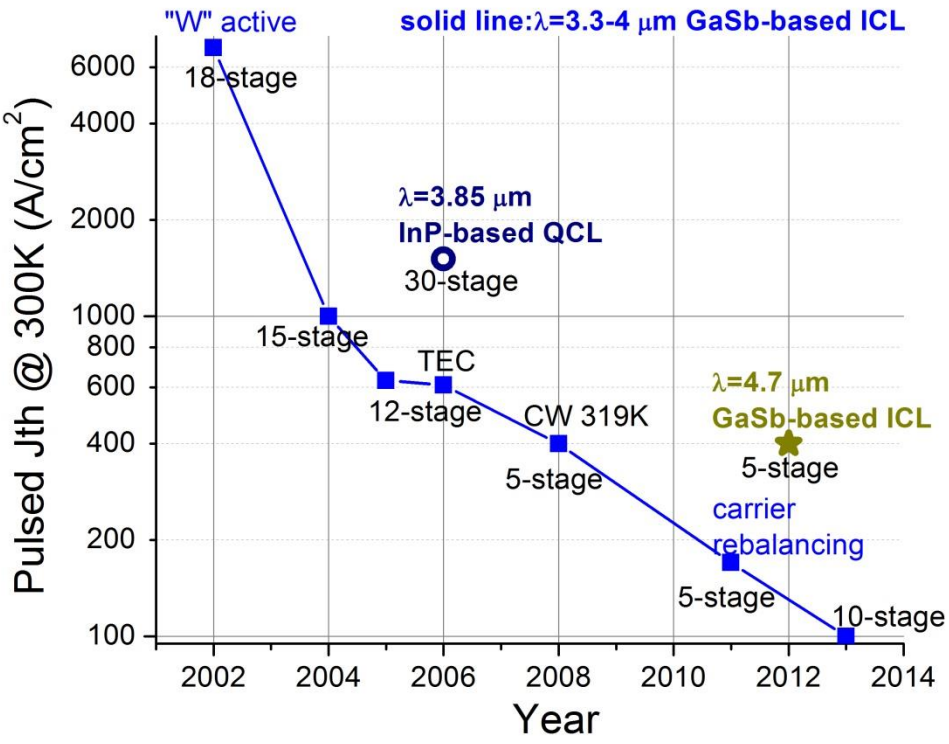


Figure 1-5. Evolving of the threshold current densities in IC lasers over the years

In 2008, the major breakthrough, CW operation at room temperature, was achieved with a threshold current density of 400 A/cm^2 [61]. The "carrier rebalancing" design and other further refinement brought the threshold down to 170 A/cm^2 [51] and ~100 A/cm^2 [62]. The low threshold current and thus a remarkably low power consumption, *e.g.*, 29 mW [51], are highly desirable for battery powered applications. The state-of-the-art IC lasers are able to operate in CW mode at temperatures up to 388 K [63], deliver high output power up to 592 mW at room temperature [64], and have wall plug efficiency (WPE) as high as 18% at room temperature [65]. This WPE value is comparable with the best value, 21%, obtained by QC lasers in CW mode [31].

The above IC lasers, with wavelengths of 3.3~3.9 μm, are all grown on the GaSb substrates. Their waveguide cladding layers, consisting of thick InAs/AlSb

superlattices (SL), have a low thermal conductivity and are challenging to grow by molecular beam epitaxy (MBE). These problems become more severe at longer wavelengths because of the requirement of a thicker cladding layer. InAs-based IC lasers [66], utilizing highly doped InAs as the optical cladding layer, have been developed to alleviate the issues. A coverage of wavelength from 3.3 μm [67] to 11 μm [68] was demonstrated. Pulsed operations at room temperature were obtained beyond 5 μm [69], 6 μm [70] and 7 μm [71]. Very recently, CW operation of InAs-based IC lasers above room temperature was achieved at wavelength ~ 4.8 μm [72,73]. The threshold current density of 247 A/cm^2 is the lowest value ever reported among semiconductor mid-infrared lasers at similar wavelengths.

This dissertation is organized as follows. Chapter 2 discusses the fundamentals and physical models of IC lasers that are crucial for laser design. Chapter 3 describes the material growth, the fabrication, and the laser characterizations. Chapter 4 analyzes the far-field patterns and the beam qualities of InAs-based IC lasers. Chapter 5 presents the performance improvement in InAs-based IC lasers, including the breakthrough of continuous wave operation above room temperature. Chapter 6 shows long-wavelength InAs-based IC lasers, including lasing wavelength up to 11 μm, and the room temperature operation at 6 μm. Chapter 7 investigates the electrical tunability of IC lasers. Chapter 8 describes InAs-based distributed-feedback IC lasers with a thin top cladding layer. A research summary and future outlook are provided in Chapter 9.

Chapter 2 Fundamentals of interband cascade lasers

This chapter presents the basic physics and some physical models of interband cascade lasers. Section 2.1 analyses the gain level in the laser operation and the dynamic process with the help of rate equation. Section 2.2 further quantitatively discuss the transparency condition, threshold condition and gain coefficient in semiconductor lasers. Section 2.3 introduces the Kane model, including both 8-band model and 2-band model, for calculation of the band structure. Section 2.4 presents models for the optical wave propagation, the optical waveguide, and the far-field pattern.

2.1 Basis in laser operation

2.1.1 Gain analysis

In thermal equilibrium, the distribution of classical particles can be described by the Boltzmann distribution:

$$\frac{n_2}{n_1} \propto \exp(-\frac{E_2 - E_1}{k_B T}) \tag{2.1.1}$$

where n_2 and n_1 [cm^{-3}] are carrier concentration in the upper energy level E_2 and lower energy level E_1, respectively. k_B=0.0086 meV/K is the Boltzmann constant. T [K] is the temperature. According to the Boltzmann distribution, the population in the upper lever is always smaller than that in the lower level, unless the temperature is "negative" [6].

However, a laser doesn't work in thermal equilibrium. Instead, a laser requires an external pumping source to establish population inversion to obtain the gain.The gain coefficient g [cm^{-1}] is proportional to the population inversion [74]:

$$g(\nu) = \sigma(\nu)(n_2 - \frac{g_2}{g_1} n_1) \qquad (2.1.2)$$

where g_i is degeneracy, the number of ways that a carrier can have in the energy level. The **stimulated emission cross section** $\sigma(\nu)$ [cm^2] is actually a cluster of several physical quantities:

$$\sigma(\nu) = A_{21} \frac{\lambda^2}{8\pi n^2} f(\nu) \qquad (2.1.3)$$

where A_{21} [s^{-1}] is the spontaneous emission coefficient, λ is the wavelength, n is the refractive index, and f(*ν*) [s] is the line shape function which describes the relative distribution at frequency *ν*. The stimulated emission cross section $\sigma(\nu)$ has a typical value of 10^{-16} cm^2, but may vary depending on the system property.

Figure 2-1 shows the modal gain Γg as a function of injection. Note that the material gain g is not readily measurable, because only a portion of the light will interact with the gain medium. This portion is usually indicated by the confinement factor Γ. The injection will first reach the **transparency condition** $(n_2 = \frac{g_2}{g_1} n_1)$, where the gain medium starts to have a positive gain. The laser serves as a light emitting diode (LED) and has amplified spontaneous emission. The spontaneous emission generates broad-spectrum photons which are the initial "seeds" to be amplified. However, such light is not coherent due to insufficient optical feedback inside the cavity. In other words, all the initial "seeds" will eventually decay completely during several round trips in the cavity; even their amplified "copies" will not survive.

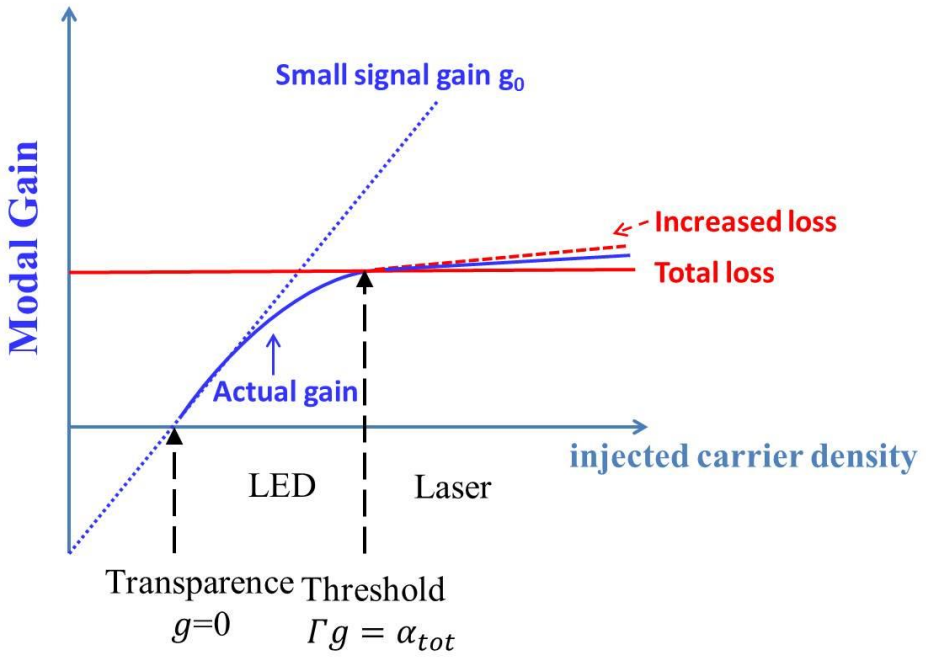


Figure 2-1. Modal gain as a function of injection

With more injection, the modal gain Γg can be increased to compensate the total loss, which is called the **threshold condition**:

$$\Gamma g = \alpha_{tot} \tag{2.1.4}$$

where Γ is the confinement factor describing the portion of the light confined in the active region.

However, the gain is not going to increase with further injection. Although somewhat counterintuitive, this can be quantitatively explained by a three-level system, as shown in Figure 2-2. The pumping rates from ground level to level 2 and level 1 are R_2 and R_1, respectively. The lifetime of an electron at level 2 and level 1 are τ_2 and τ_1, respectively. The lifetime of spontaneous transition from level 2 to level 1 is τ_{21}.

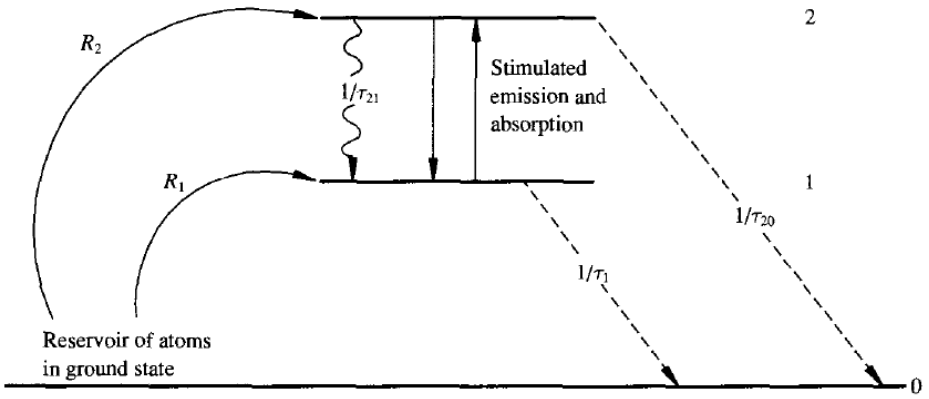


Figure 2-2. Pumping and carrier lifetime in a three-level laser [74]

The population inversion and the gain coefficient in a steady state can be expressed as:

$$n_2 - n_1 = \frac{R_2\tau_2(1-\frac{\tau_1}{\tau_{21}}) - R_1\tau_1}{1+(\tau_1+\tau_2-\frac{\tau_1\tau_2}{\tau_{21}})(\frac{\sigma I_\nu}{h\nu})} \tag{2.1.5}$$

$$g(\nu) = \frac{g_0(\nu)}{1+I_\nu / I_s} \tag{2.1.6}$$

$$g_0(\nu) \equiv \sigma(\nu)\left[R_2\tau_2(1-\frac{\tau_1}{\tau_{21}}) - R_1\tau_1\right] \tag{2.1.7}$$

$$I_s \equiv \frac{h\nu}{\sigma} / (\tau_1+\tau_2-\frac{\tau_1\tau_2}{\tau_{21}}) \tag{2.1.8}$$

The gain coefficient in equation (2.1.6) is decomposed into a small-signal gain coefficient g_0 defined by equation (2.1.7), and saturation intensity I_s [w/cm^2] defined by equation (2.1.8). The small-signal gain coefficient increases with the pumping, regardless of the photon intensity I_v.

As shown in Figure 2-1, when a laser approaches its threshold, the gain will more deviate from the small-signal gain because the ratio I_ν / I_s is significant. If the

modal gain is larger than the total loss, the photon intensity will continue to increase until the modal gain falls below the total loss. This dynamic process is more accurately described by the rate equation, which will be presented in section 2.1.2. As a result, the population inversion and the gain are clamped to a constant value. Further pumping only increases the photon intensity, but the steady-state gain coefficient will remain the same (*i.e.*, $g = \alpha_{tot} / \Gamma$), preventing the photon intensity from growing indefinitely through more and more cavity transits.

However, in reality, the total loss may not be constant. The large injection may: (1) increase the temperature in the laser active core because the power conversion efficiency is far less than one; (2) increase the optical loss, including the free carrier absorption loss, the inter-valence band absorption loss and facet scattering loss; (3) increase the carrier loss such as Auger recombination loss, carrier leakage loss, tunneling loss, *etc*. All of these factors together will ultimately lead to "power rollover" that the laser output power ceases to increase, even in the pulsed operation.

2.1.2 Rate equation

Before a laser reaches the steady state, the carrier and photon density in the cavity undergo some oscillations until the balance between gain and loss is achieved. This dynamic process can be described by a simple rate equation model [75]. For the excess carriers (either electrons or holes), the rate equation is:

$$\frac{dn}{dt} = \frac{\eta_{inj} j}{e d_a} - \frac{n}{\tau_{sp}} - v_g g s \qquad (2.1.9)$$

where n[cm^{-3}] is the carrier density at the excited state, s [cm^{-3}] is the photon density, η_{inj} is the current injection efficiency, j [A/cm^2] is the current density, e [C] is the charge of a single electron, d_a is the thickness of active region, $v_g = \frac{c}{n_g}$ [cm/s] is the group velocity of photons in the cavity, and τ_{sp} [s] is the lifetime of spontaneous recombination, which can be generally decomposed into three types of carrier recombination:

$$\frac{n}{\tau_{sp}} = \frac{n}{\tau_{SRH}} + \frac{n}{\tau_r} + \frac{n}{\tau_{Auger}} = An + B(T)np + C(T)np^2 \qquad (2.1.10)$$

where τ_{SRH}, τ_r and τ_{Auger} are the lifetime of Shockley-Read-Hall recombination, spontaneous radiative recombination and Auger recombination, respectively. A, B and C are the corresponding recombination coefficients. They are all temperature dependent.

To simplify the discussion, the gain coefficient g in equation (2.1.2) is assumed to have a linear relation with the injected carrier density:

$$g = \sigma(n - n_e) \qquad (2.1.11)$$

where n_e is the transparency carrier density. Further discussion of the gain coefficient will be given in section 2.2.3 and section 5.2.1 with a more accurate equation.

For photons in single mode, the rate equation is:

$$\frac{ds}{dt} = \left(\Gamma g - \alpha_{tot}\right) v_g s + \Gamma \gamma' \frac{n}{\tau_{sp}} \qquad (2.1.12)$$

where γ' is the ratio of single-mode photons to the carriers participating in the spontaneous recombination. At a steady state $\frac{ds}{dt}$ =0, the modal gain is only slightly smaller than the total loss.

For a more general expression, the carrier density, the current density, and the photon density are normalized:

$$N = \frac{n}{n_c}, \Gamma v_g \sigma (n_c - n_e) = v_g \alpha_{tot} = \frac{1}{\tau_{ph}} \tag{2.1.13}$$

$$J = \frac{j}{j_r}, j_r = (1-\gamma')\frac{q_e d_a}{\eta_i}\frac{n_c}{\tau_{sp}} \tag{2.1.14}$$

$$S = \frac{s}{s_r}, s_r = \Gamma(1-\gamma')\frac{n_c}{\tau_{sp} v_g \alpha_{tot}} \tag{2.1.15}$$

where n_c [cm^{-3}] is the critical carrier density that its modal gain is equal to the total loss, τ_{ph} [s] is the cavity lifetime of the photon), j_r [A/cm^2] is the threshold current density, and s_r [cm^{-3}] is the photon density when the injected current is twice the threshold value.

The normalized time gain G is defined as:

$$G = \frac{g(n)}{g(n_c)} = \Gamma \tau_{ph} v_g g(n) = \Gamma \tau_{ph} v_g n_c (N - N_e) = A(N - N_e) \tag{2.1.16}$$

The rate equations (2.1.9) and (2.1.12) can be transformed to:

$$\tau_{sp}\frac{dN}{dt} = (1-\gamma')J - N - (1-\gamma')GS \tag{2.1.17}$$

$$\tau_{ph}\frac{dS}{dt} = (G-1)S + \frac{\gamma'}{1-\gamma'}N \tag{2.1.18}$$

At the steady state, $\frac{dN}{dt} = 0, \frac{dS}{dt} = 0$, the solution is :

$$N = \frac{1}{2}\left\{\left(J + \frac{1-\gamma' N_e}{1-\gamma'}\right) - \sqrt{\left(J + \frac{1-\gamma' N_e}{1-\gamma'}\right)^2 - 4J}\right\} \approx \frac{1}{2}(J + 1 - |J-1|) \qquad (2.1.19)$$

$$S = \frac{1}{2}\left\{\left(J - \frac{1-\gamma' N_e}{1-\gamma'}\right) + \sqrt{\left(J + \frac{1-\gamma' N_e}{1-\gamma'}\right)^2 - 4J}\right\} \approx \frac{1}{2}(J - 1 + |J-1|) \qquad (2.1.20)$$

where the approximation holds when $\gamma' << 1$.

These two equations show the expected results. For most semiconductor lasers, an only slight portion of the spontaneous recombined carriers turn into single-mode photons, *i.e.*, $\gamma' << 1$. Whether the injected current is above threshold results in two unique results. When the injected current is below the threshold, J<1, the injected current becomes the excited carrier, N=J, and there's no laser light, S=0. When the injected current is above the threshold, J>1 , the excited carrier density is pinned at the threshold level, N=1, and the photon density increases linearly with the injected current, S=J-1. Since the carrier density is pinned, the gain is also saturated, which is roughly equal to α_{tot} / Γ.

A time-dependent solution will help understand how the balance is achieved, which is essential to the high-speed operation and "gain switch pulse." The differential equations (2.1.16) and (2.1.17) can be numerously solved by Runge-Kutta-Fehlberg method. Figure 2-3 shows a typical dynamic behavior for a DH laser with a current injection of 1.2 times the threshold.

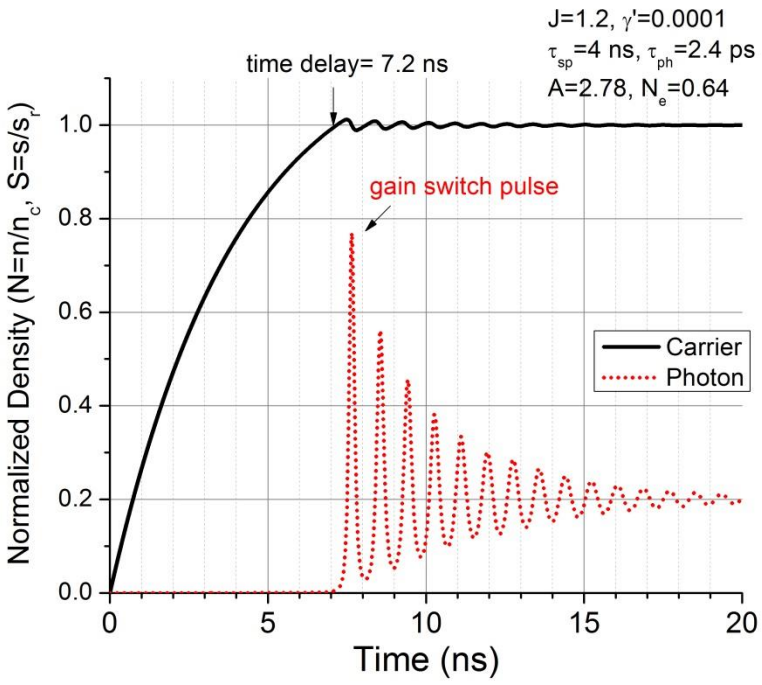


Figure 2-3. Dynamic process in a semiconductor laser before the steady state.

There's a buildup time for the carrier to accumulate from zero to threshold. This will delay the onset of laser operation by

$$t_d = \tau_{sp} \ln\left(\frac{j}{j - j_{th}/(1-\gamma')}\right) = \tau_{sp} \ln\left(\frac{J(1-\gamma')}{J(1-\gamma')-1}\right) \qquad (2.1.21)$$

At the delayed time, the photon density is still below the steady-state value, which causes the carrier overshooting the threshold value. Consequently, the corresponding gain is larger than total loss and results in "gain switch pulse". Then both carrier and photon density have damped oscillations around the steady-state value. For minimization of the delayed time and oscillation, a common practice in a signal modulation is to superimpose a signal on a dc current which is around threshold value.

2.2 Gain coefficient in semiconductor materials

In this section, more quantitative discussions are given on the transparency condition, the threshold condition, and the gain coefficient.

2.2.1 Probability inversion

In interband semiconductor lasers, photons are generated by electron-hole recombinations. Electrons are Fermions, which obey the Fermi-Dirac distribution function in thermal equilibrium:

$$f(E) = \frac{1}{\exp(\frac{E - E_F}{k_B T}) + 1} \tag{2.2.1}$$

where E_F is the Fermi energy.

A laser operates in a non-equilibrium state. However, because the interband recombination lifetime is much longer than the intraband relaxation lifetime, the carriers in the conduction band and the valence band can each arrive at **quasi-equilibrium**, with separate quasi-Fermi levels for each band. The occupation probabilities of an electron in the conduction band and the valence band are:

$$f_c(E_e) = \frac{1}{\exp(\frac{E_c - F_c}{k_B T}) + 1} \tag{2.2.2}$$

$$f_v(E_h) = \frac{1}{\exp(\frac{E_v - F_v}{k_B T}) + 1} \tag{2.2.3}$$

where F_c and F_v are quasi-Fermi energies in the conduction band and the valence band, respectively.

To amplify the photons or have a net material gain, the probability for the electron-hole recombination must be larger than its inverse process:

$$f_c(1-f_v) > f_v(1-f_c) \tag{2.2.4}$$

Then we get

$$f_c > f_v \Leftrightarrow F_c - F_v > E_c - E_v > E_g \tag{2.2.5}$$

This **inversion of the probability** of an electron between the two states is called **Bernard-Duraffourg inversion** condition [76]. This is equivalent to the population inversion mentioned in section 2.1. Note that because the valence band is almost full of electrons while the conduction band is nearly empty, it's impossible to achieve the population inversion between the conduction band and the valence band.

In equation (2.2.5), the probability inversion is equivalent to the condition that the quasi-Fermi energy difference is larger than the photon energy. In analog to the population inversion, the critical point of probability inversion, *i.e.*, $f_c = f_v$ or $F_c - F_v = E_{photon}$, can be taken as the **transparency condition**. Numerical results are presented in next sub-section.

2.2.2 Quasi-Fermi levels in bulk and QW material

The density of states ρ_{DOS} [cm^{-3} eV^{-1}] in the conduction band of a bulk material is:

$$\rho_{DOS} = \frac{\sqrt{2}}{\pi^2}\left(\frac{m_n^*}{\hbar^2}\right)^{3/2}\sqrt{E - E_c} \tag{2.2.6}$$

The carrier concentration in the conduction band is

$$n = \int_{Ec}^{\infty} \rho_{DOS} f(E) dE = N_c \cdot F_{\frac{1}{2}}(\zeta) \tag{2.2.7}$$

where N_c [cm^{-3}] is the effective density of states in the conduction band, $F_{\frac{1}{2}}(\zeta)$ is the Fermi integral:

$$N_c = 2(\frac{m_n^* k_B T}{2\pi\hbar^2})^{3/2} \tag{2.2.8}$$

$$F_{\frac{1}{2}}(\zeta) = \frac{2}{\sqrt{\pi}}\int_0^\infty \frac{\sqrt{x}}{1+e^{x-\zeta}}dx, \zeta = \frac{F_c - E_c}{k_B T} \tag{2.2.9}$$

The expression for the valence band is similar. However, the effective mass of the heavy hole is about an order larger than that of the electron. Consequently, the effective density of states in the valence band is much greater than that in the conduction band. As a result, the same carrier concentration will cause more energy shift of F_c than F_v.

Figure 2-4 shows the quasi-Fermi levels and the band edge occupation probability as a function of carrier density in InAs/GaSb heterostructure. Only Γ valley and c-hh transition are considered to highlight the essential physics. The carrier densities of electron and hole are the same to keep charge neutrality. We can see the quasi-Fermi level and the occupation probability of electron increase much more rapidly than the hole. As a result, the transparency condition requires an injected carrier density of 3.4*N_c, as indicated in the figure. The penalties associated with the effective mass asymmetry are a higher transparency carrier density, a larger free carrier absorption loss, and a larger Auger recombination loss. All together result in a higher threshold current density. One solution to alleviate this effect is lowering the valence band effective mass by changing the strain and quantum confinement [17]. Another possible solution is to have modulation doping in the active region [77].

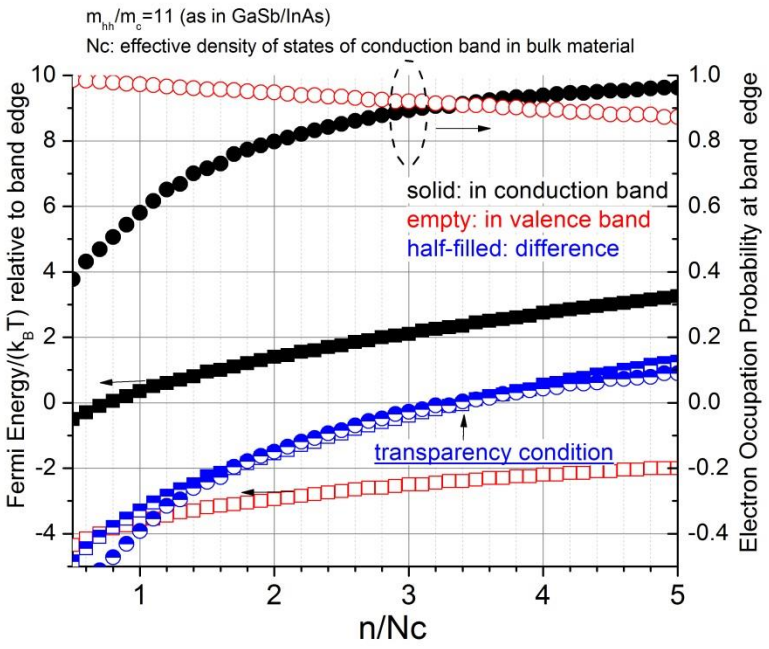


Figure 2-4. Quasi-Fermi levels as a function of carrier density in the bulk material.

For QW materials, the density of states ρ_{DOS}^{2D} [cm^{-2} eV^{-1}], the carrier concentration n^{2D} (cm^{-2}) and the effective density of states N_c^{2D} (cm^{-2}) are:

$$\rho_{DOS}^{2D} = \frac{m_n^*}{\pi\hbar^2} \tag{2.2.10}$$

$$n^{2D} = \int_{Ec}^{\infty} \rho_{DOS}^{2D} f(E)dE = N_c^{2D} \ln(1+e^{\zeta}) \tag{2.2.11}$$

$$N_c^{2D} = \frac{m_n^* k_B T}{\pi\hbar^2} \tag{2.2.12}$$

where $\zeta = \frac{F_c - E_{c1}}{k_B T}$, E_{c1} is the ground state energy in the QW.

To compare how the quasi-Fermi levels are affected in bulk and QW material, the carrier concentrations are normalized by their respective effective density of states. As shown in figure 2-5, the two curves overlap each other when the quasi-Fermi levels

are below the band edge. After that, the quasi-Fermi level for QW increases more rapidly, which means higher differential material gain.

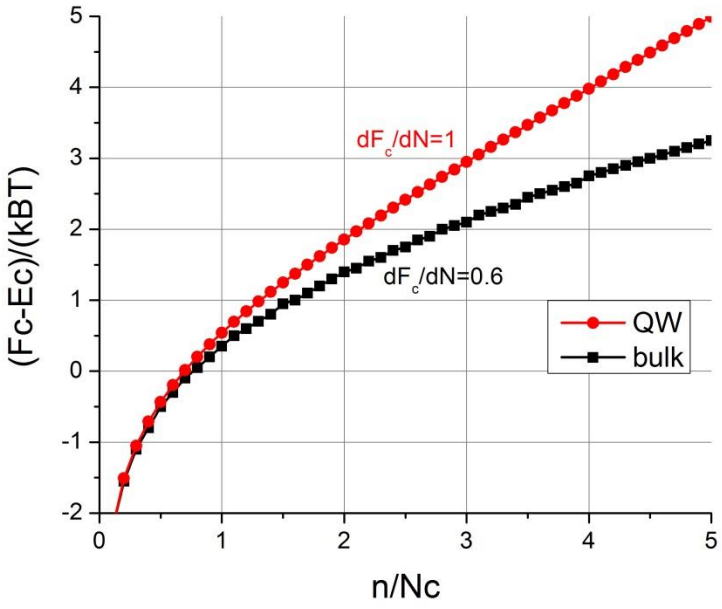


Figure 2-5. Comparison of quasi-Fermi levels between bulk and QW materials

Assume the same total carrier recombination lifetime and the same amount of normalized injected carrier density (n/N_c), the ratio of bulk current to QW current is:

$$\frac{J_{bulk}}{J_{QW}} = d_w \sqrt{\frac{m_c^* k_B T}{2\pi\hbar^2}} \xrightarrow{InAs,300K} \frac{d_w}{30nm} \tag{2.2.13}$$

This means, under "equivalent" carrier injection, the current density in bulk material is always larger than QW material unless the active core thickness is smaller than 30 nm. This makes sense because the bulk material will become QW if the thickness is sufficiently thin.

2.2.3 Gain coefficient

Once the Bernard-Duraffourg inversion condition is achieved, the injected semiconductor will have the net material gain. Within the framework of Fermi's Golden Rule, the net material gain is written as [78]:

$$g(\hbar\omega) = \frac{\pi q_e^2 \hbar}{\bar{n}_b m_0^2 c_0 \varepsilon_0 \hbar\omega} \left| M_{avg} \right|^2 \rho_t(E_t) \left[f_c(E_t) - f_v(E_t) \right] \tag{2.2.14}$$

$$\begin{aligned} bulk: \rho_t[cm^{-3}eV^{-1}] &= \frac{1}{2\pi^2} \left(\frac{2m_r}{\hbar^2} \right)^{3/2} \sqrt{E_t} \\ QW: \rho_t[cm^{-2}eV^{-1}] &= \frac{m_r}{\pi\hbar^2} \end{aligned} \tag{2.2.15}$$

where ρ_t is the reduced density of states, $\frac{1}{m_r} = \frac{1}{m_c} + \frac{1}{m_h}, E_t = \hbar\omega - E_g$. With k conservation, the Fermi distribution factor is expressed as:

$$\begin{aligned} f_c(E_t) &= \frac{1}{1 + \exp(\frac{E_t m_r / m_c - F_c}{k_B T})} \\ f_v(E_t) &= \frac{1}{1 + \exp(\frac{F_v - E_t m_r / m_v}{k_B T})} \end{aligned} \tag{2.2.16}$$

And M_{avg} [eV·s·cm^{-1}] is the momentum transition matrix element averaged over a solid angle for bulk material, or averaged over an azimuthal angle for QW material [79]:

$$\left| M_{avg} \right|^2 = \frac{1}{4\pi} \int \left| \underline{e} \cdot \underline{M}_{c-hh} \right|^2_{b,\uparrow,} d\Omega = \frac{1}{3} p_x{}^2 \equiv M_b^2 \tag{2.2.17}$$

The value of momentum transition matrix element p_x can be extracted by the experimental E_p [eV] parameter compiled in [80], which has included the effects of adjacent bands and remote bands. The relations between these symbols are:

$$P \equiv \frac{-i\hbar}{m_0}\left\langle S\left|\hat{p}_x\right|X\right\rangle = \frac{-i\hbar}{m_0}p_x$$
$$E_p \equiv \frac{2m_0}{\hbar^2}P^2 = \frac{2{p_x}^2}{m_0} \tag{2.2.18}$$

For spin degeneracy, the above expression for the matrix element assumes that a factor of two will be included in the expression for the density of states [79]. For transition of an electron to heavy-hole in a quantum well, $\left|M_{QW}^{TE}\right|^2 = 1.5\left|M_b\right|^2$ [79].

According to the relation of Einstein coefficients, the spontaneous emission rate r_{sp} [$eV^{-1}\cdot cm^{-3}\cdot s^{-1}$] is given by:

$$r_{sp}(E_{21}) = g(E_{21})\frac{\overline{n}_b^2 E_{21}^2}{\pi^2\hbar^3 c_0^2}\frac{f_c\left[1-f_v\right]}{f_c - f_v} \tag{2.2.19}$$

Then the corresponding "spontaneous emission" recombination rate R_{sp} [s^{-1}], recombination lifetime τ_{sp} [s], recombination coefficient B_{sp} [cm^6s^{-1}], and recombination current density J_{sp} [A/cm^2] can be calculated by:

$$R_{sp}(N) = \int_0^\infty r_{sp}(E,N)dE = \frac{N}{\tau_{sp}(N)} = B_{sp}(N)N^2 \tag{2.2.20}$$

$$J_{sp}(N) = q_e d_a R_{sp}(N) \tag{2.2.21}$$

A numerical example of the gain spectra and the spontaneous emission spectra is shown in figure 2-6. The QW structure has a higher gain due to its larger momentum matrix element and smaller transparency carrier density. The practical gain spectra will have some band tailing near the bandgap,which is partially due to the lineshape broadening caused by the finite lifetime of an electron at the energy levels.

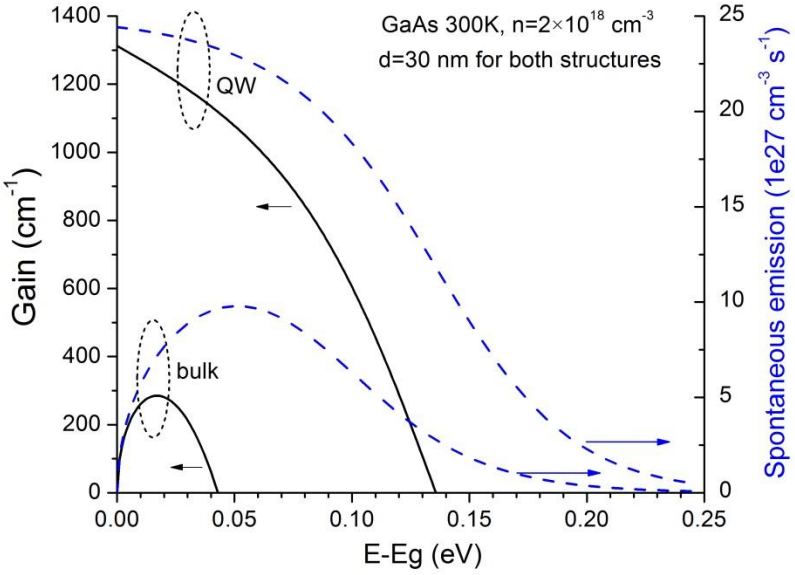


Figure 2-6. Gain spectra and spontaneous emission spectra for GaAs bulk and QW

The gain spectra for InAs has a similar shape. However, the electron effective mass in InAs is only one-third of that in GaAs, which has twofold effects. On the positive side, the smaller effective density of states in InAs results in a lower transparency carrier density and possibly lower threshold current density. On the other hand, the smaller density of states will lead to smaller differential gain, which makes its threshold more sensitive to the temperature.

2.3 Active region material and band structure

This section introduces the type-II QW materials that are the backbone of the cascade structure. The Kane model, including both 8-band model and 2-band model, is presented to calculate the band structure.

2.3.1 Type-II QW materials

Almost all commercial semiconductor lasers since the 1990s have used QWs as the light-emitting material because of their better performance. Depending on their band alignments, QWs can be classified into several types.

If the bandgap of one material completely falls within the bandgap of the other material, as shown in figure 2-7(a), this combination is called type-I QW. Because the electrons and holes are confined in the same material, a type-I QW has the advantage of large wavefunction overlap and hence, a higher gain coefficient.

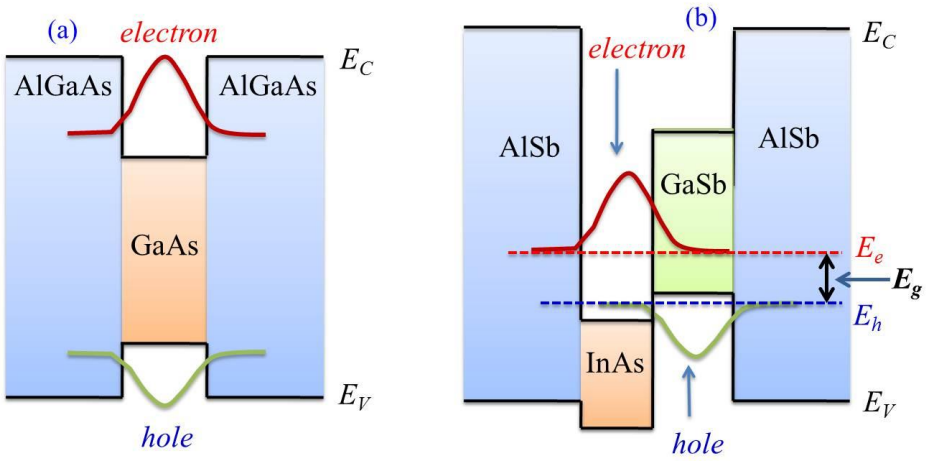


Figure 2-7. Band alignment for (a) type-I QW; (b) type-II QW [81]

If the bandgap of one material is not fully encapsulated by another material, as shown in figure 2-7 (b), this combination is called type-II QW. In this configuration, electrons and holes are confined in different materials, forming an effective "bandgap" determined by their respective quantized ground states. Although there is a smaller wavefunction overlap, the type-II QW has two distinctive advantages. First, it can have a larger tuning range of the lasing wavelength by changing the thicknesses of both materials, while the type-I QW has a long wavelength limit determined by the

material's own bandgap. Second, the spatially indirect transition in a type-II QW is not affected by the possible defect states within the material bandgap, which is very common in narrow bandgap materials.

A typical example of a type-II QW is an InAs/GaSb QW, with AlSb as their common barrier material. The conduction band edge of InAs is ~0.15 eV below the valence band edge of GaSb, forming a broken-gap alignment. This unique structure has stirred a lot of research interests. Negative differential resistance was observed as an evidence of (resonant) interband tunneling [82-84]. Resonant interband tunneling was considered as a mechanism for fast extraction from the lower lasing subband [44,85-87], a key point for high-performance quantum cascade lasers [88]. Nevertheless, an intersubband laser based on resonant interband tunneling has not been demonstrated yet. Instead, an IC laser based on the type-II QW interband transition has been proven to be more efficient, due to a much longer lifetime of interband transition. In addition, a 20~25Å thick AlSb layer sandwiched between the InAs/GaSb QW to connect the electron injector and the hole injector, can be used as a semi-metallic source [89] to facilitate the carrier transport [90], or "internally generate carriers" [51], which makes the cascade process possible.

2.3.2 Band structure modeling

As shown in figure 1-4, a cascade stage of an IC laser is composed of a "W" shaped type-II active region and two injectors of multiple type-I QWs. The structure optimization relies on the numerical calculation of the wavefunction. The popular envelope-function approximation, as an adaptation of the k·p theory [91], is an efficient

method to calculate the semiconductor band structure with decent accuracy. This approximation is convenient to implement, *e.g.*, when an external electrical field is applied.

The electron wave function satisfies the time-independent Schrodinger equation:

$$\hat{H}\Psi(r) = [\frac{\hat{p}^2}{2m_0} + V(r)]\Psi(r) = E\Psi(r) \qquad (2.3.1)$$

where $\hat{H}$ is the Hamiltonian operator and $\hat{p} = -i\hbar\nabla$ is the momentum operator.

Because the potential V(r) is periodic in a semiconductor crystal, the general solution is given by

$$\Psi_{nk}(r) = e^{ik\cdot r}u_{nk}(r) \qquad (2.3.2)$$

where n is the band index and u_{nk} has the same period as the crystal lattice. This result is the Bloch theorem, and the wavefunction $\Psi_{nk}(r)$ is called the Bloch function.

If the wavefunction $\Psi_{nk}(r)$ is substituted into the Schrodinger equation (2.3.1) , we have :

$$\left[\frac{\hbar^2k^2}{2m_0} + \frac{\hbar}{m_0}\vec{k}\cdot\hat{p} + \frac{\hat{p}^2}{2m_0} + V(r)\right]u_{nk}(r) = E_n(k)u_{nk}(r) \qquad (2.3.3)$$

At k=0, equation (2.3.3) reduces to

$$\left[\frac{\hat{p}^2}{2m_0} + V(r)\right]u_{n0}(r) = E_n(0)u_{n0}(r) \qquad (2.3.4)$$

The solutions of the equation of (2.3.4) form a complete and orthonormal set of basis functions. The terms $\hat{H}_k = \frac{\hbar^2k^2}{2m_0} + \frac{\hbar}{m_0}\vec{k}\cdot\hat{p}$ can be treated as perturbations term using perturbation theory. This method for calculating the band dispersion is known as

the k·p method. For accounting for the spin-orbit interaction, it's necessary to include the corresponding Hamiltonian:

$$\hat{H}_{so} = \frac{\hbar}{4c^2 m_0^2} \hat{\sigma} \cdot \nabla V \times \hat{p} \tag{2.3.5}$$

The zinc blende semiconductors are well described by the eight-band Kane model [91], in which the basis functions are taken from the s-state wave function and the p-state wave functions of a hydrogen atom model:

$$\begin{aligned} Y_{00} &= \frac{1}{\sqrt{4\pi}} \\ Y_{10} &= \left| Z \right\rangle \\ Y_{1\pm1} &= \mp \frac{1}{\sqrt{2}} \left| X \pm iY \right\rangle \end{aligned} \tag{2.3.6}$$

These four wave functions plus their opposite spin counterparts constitute the eight basis functions in Kane model. The 4×4 Hamiltonian matrix is

$$\underline{\underline{H}}_{4\times4} = \begin{bmatrix} E_g & 0 & kP & 0 \\ 0 & -\frac{2\Delta}{3} & \frac{\sqrt{2}\Delta}{3} & 0 \\ kP & \frac{\sqrt{2}\Delta}{3} & -\frac{\Delta}{3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.3.7}$$

For non-trivial solutions, the discriminant must be zero:

$$E'\left\{ E'\left(E' - E_g\right)\left(E' + \Delta\right) - k^2 P^2 \left(E' + \frac{2\Delta}{3} \right) \right\} = 0 \tag{2.3.8}$$

where $E'(k) = E\left(k\right) - \frac{\hbar^2 k^2}{2m_0}$. Around the band edge, $k^2 P^2 << E_g$, we can get the analytical solutions by the first-order approximation:

$$
\begin{aligned}
E_c(k) &= E_g + \frac{\hbar^2 k^2}{2m_0} + \frac{k^2 P^2 \left(E_g + 2\Delta/3\right)}{E_g\left(E_g + \Delta\right)} \\
E_{hh}(k) &= \frac{\hbar^2 k^2}{2m_0} \\
E_{lh}(k) &= \frac{\hbar^2 k^2}{2m_0} - \frac{2k^2 P^2}{3E_g} \\
E_{so}(k) &= -\Delta + \frac{\hbar^2 k^2}{2m_0} - \frac{k^2 P^2}{3\left(E_g + \Delta\right)}
\end{aligned}
\tag{2.3.9}
$$

In Kane's four-band model, the conduction band, the light-hole band, and the spin-orbit split-off band are coupled together, with their eigenenergies connected by the moment matrix element P. However, the heavy-hole band is decoupled from the other three bands. This heavy hole has a positive free-electron mass, which is apparently wrong. The main reason is the lack of the effect of remote bands. And the choice of basis functions may also be responsible. This problem can be solved by the Luttinger-Kohn model, in which the unperturbed term $\hat{H}_0^{LK}$ and perturbed term $\hat{H}_k^{LK}$ are different from the Kane model:

$$
\begin{aligned}
\hat{H}_0^{LK} &= \frac{\hat{p}^2}{2m_0} + V(r) + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{4m_0^2 c^2}\nabla V \times \hat{p} \cdot \hat{\sigma} \\
\hat{H}_k^{LK} &= \frac{\hbar}{m_0}\vec{k} \cdot \hat{\Pi} = \frac{\hbar}{m_0}\vec{k} \cdot (\hat{p} + \frac{\hbar}{4m_0 c^2}\hat{\sigma} \times \nabla V)
\end{aligned}
\tag{2.3.10}
$$

However, there is no analytical expression for the E(k) expression in the Luttinger-Kohn model, which makes it less popular than the Kane model. A self-consistent algorithm based on Luttinger-Kohn's 4-band model was developed by Chuang's group to calculate the temperature dependency of material gain [92].

2.3.3 Two-band model and formalism

One representation of the two-band model is to assume the spin-orbit interaction Δ=0. Consequently, the Kane 4×4 Hamiltonian matrix is decoupled into a single band for the heavy hole and a 2×2 Hamiltonian matrix for the conduction and the light hole valence band [93-95]:

$$\underline{\underline{H}}_{2\times 2} = \begin{bmatrix} E_g & kP \\ kP & 0 \end{bmatrix} \tag{2.3.11}$$

For non-trivial solutions, the discriminant must be zero:

$$E'\left(E' - E_g\right) - k^2P^2 = 0 \tag{2.3.12}$$

where $E'(k) = E\left(k\right) - \dfrac{\hbar^2k^2}{2m_0}$. Around the band edge, $k^2P^2 << E_g$, we can get the parabolic dispersions by the first-order approximation:

$$\begin{aligned} E_c(k) &= \frac{E_g + \sqrt{E_g^2 + 4k^2P^2}}{2} + \frac{\hbar^2k^2}{2m_0} \approx E_g + \frac{k^2P^2}{E_g} + \frac{\hbar^2k^2}{2m_0} \\ E_{lh}(k) &= \frac{E_g - \sqrt{E_g^2 + 4k^2P^2}}{2} + \frac{\hbar^2k^2}{2m_0} \approx \frac{-k^2P^2}{E_g} + \frac{\hbar^2k^2}{2m_0} \end{aligned} \tag{2.3.13}$$

So the electron effective mass around the conduction band edge is

$$\frac{1}{m_c^*} = \frac{d^2E_c}{\hbar^2dk^2} = \frac{1}{m_0} + \frac{2P^2}{\hbar^2E_g} = \frac{1}{m_0} + \frac{E_P}{m_0E_g} \tag{2.3.14}$$

Note that some published papers have a different notation for the moment matrix element. For example, $P'[cm/s] \equiv \dfrac{P}{\hbar} = \sqrt{\dfrac{E_p}{2m_0}}$ is used in reference [96,97]. For consistency, the "two-band" P parameter is evaluated at the conduction band edge by equation (2.3.14), instead of taking the E_p parameter compiled in [80]. Consequently,

the equivalent "two-band" E_p value averaged over InAs-GaSb-AlSb is 15.7 eV, smaller than the "complete band" E_p of about 22.4 eV. This discrepancy is because the two-band model has neglected effects from other bands.

In the following, the formalism for the transfer matrix method is presented.

The total wavefunction in a two-band model is:

$$\Psi_{nk}(r) = a_c(k)u_{c0}(r) + a_{lh}(k)u_{lh0}(r) \tag{2.3.15}$$

where $a_c(k)$ and $a_{lh}(k)$ are envelope functions which modulate the basis functions,

The eigenfunctions for conduction band and light hole valence band are

$$\begin{bmatrix} a_c \\ a_{lh} \end{bmatrix} = \begin{bmatrix} Ae^{ikz} + Be^{-ikz} \\ \beta_k(Ae^{ikz} - Be^{-ikz}) \end{bmatrix} = \begin{bmatrix} e^{ikz} & e^{-ikz} \\ \beta_k e^{ikz} & -\beta_k e^{-ikz} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} \tag{2.3.16}$$

$$k = \frac{\sqrt{(E - E_c)(E - E_{lh})}}{\hbar P'} \tag{2.3.17}$$

$$\beta_k = \frac{\hbar P' k}{E - E_{lh}} \tag{2.3.18}$$

The wave function probability is the superposition of both components:

$$|a|^2 = |a_c|^2 + |a_{lh}|^2 \tag{2.3.19}$$

The heavy hole valence band, which can be decoupled with the electron and light hole valence band, can be described by the single-band model. The wavefunction has a similar format:

$$\begin{bmatrix} \psi \\ \frac{1}{m^*}\frac{d}{dz}\psi \end{bmatrix} = \begin{bmatrix} Ae^{ikz} + Be^{-ikz} \\ \beta_k(Ae^{ikz} - Be^{-ikz}) \end{bmatrix} = \begin{bmatrix} e^{ikz} & e^{-ikz} \\ \beta_k e^{ikz} & -\beta_k e^{-ikz} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}, \beta_k = \frac{ik}{m} \tag{2.3.20}$$

Note that the wavevectors for electron and hole are different due to their opposite polarity:

$$CB: k = \frac{1}{\hbar}\sqrt{2m_c^*(E-V)}$$
$$VB: k = \frac{i}{\hbar}\sqrt{2m_h^*(E-V)} \tag{2.3.21}$$

Once the normalized wavefunctions for both electron and hole are obtained, the oscillator strength can be calculated by[98]

$$f = \frac{2}{m_0}\frac{\left|\left\langle \psi^f \middle| \varepsilon \cdot p \middle| \psi^i \right\rangle\right|^2}{E^f - E^i} \sim \left|\left\langle F_e(z) \middle| F_h(z) \right\rangle\right|^2 = \left|\int F_e^*(z) F_h(z) dz\right|^2 \tag{2.3.22}$$

The band parameters such as effective mass and band offset can be found in several literature reviews [80,99,100]. However, there is still uncertainty regarding the band offset, which is sensitive to the growth condition.

2.3.4 Kane's two-band model

Actually, in Kane's original paper [91], his two-band model is to use an approximation $\Delta >> E_g$, kP. Consequently, equation (2.3.8) becomes

$$E'(E'^2 - E_g E' - \frac{2}{3}k^2P^2) = 0 \tag{2.3.23}$$

One of the three solutions is still the decoupled single-band for the heavy hole. And the other two solutions are the conduction band and the light hole valence band:

$$E_c(k) = \frac{E_g + \sqrt{E_g^{\,2} + \frac{8}{3}k^2P^2}}{2} + \frac{\hbar^2k^2}{2m_0} \xrightarrow{kP<<Eg} E_g + \frac{2k^2P^2}{3E_g} + \frac{\hbar^2k^2}{2m_0}$$
$$E_{lh}(k) = \frac{E_g - \sqrt{E_g^{\,2} + \frac{8}{3}k^2P^2}}{2} + \frac{\hbar^2k^2}{2m_0} \xrightarrow{kP<<Eg} -\frac{2k^2P^2}{3E_g} + \frac{\hbar^2k^2}{2m_0} \tag{2.3.24}$$

So the electron effective mass around the conduction band edge is

$$\frac{1}{m_c^*} = \frac{d^2E_c}{\hbar^2 dk^2} = \frac{1}{m_0} + \frac{4P^2}{3\hbar^2 E_g} = \frac{1}{m_0} + \frac{2E_P}{3m_0 E_g} \tag{2.3.25}$$

It seems that the Kane's "two-band" E_p value averaged over InAs-GaSb-AlSb, 23.58 eV, is much closer to "complete band" E_p of about 22.4 eV. But actually, these two representation of the two-band model are equivalent. In another word, Kane's moment matrix element $P'' = \sqrt{\frac{3}{2}}P'$, so only a constant factor is packed into the term.

2.4 Optical wave modeling

The active QWs, usually much thinner than the emission wavelengths, have a tiny filling factor or confinement factor. Therefore, an optimized optical waveguide is required to modify the optical wave distribution to achieve sufficient modal gain. In this section, the optical wave equations are derived in detail.

2.4.1 Optical wave equations

The foundation of classical optics is Maxwell's equations:

$$\nabla \cdot \vec{D} = \rho \tag{2.4.1a}$$

$$\nabla \cdot \vec{B} = 0 \tag{2.4.1b}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{2.4.1c}$$

$$\nabla \times \vec{H} = -\frac{\partial \vec{D}}{\partial t} + \vec{J} \tag{2.4.1d}$$

where D is the displacement field, B is the magnetic field, E is the electric field, H is the magnetizing field, ρ is the free charge, and J is the electric current.

From Maxwell's equations, along with the constitutive relations, *i.e.,* $\vec{D} = \varepsilon \vec{E}$, $\vec{B} = \mu \vec{H}$,we can get the wave equations in a sourceless, homogeneous and isotropic medium:

$$\nabla^2 \vec{E}(x, y, z, t) - \mu\varepsilon \frac{\partial^2 \vec{E}(x, y, z, t)}{\partial t^2} = 0 \qquad (2.4.2a)$$

$$\nabla^2 \vec{H}(x, y, z, t) - \mu\varepsilon \frac{\partial^2 \vec{H}(x, y, z, t)}{\partial t^2} = 0 \qquad (2.4.2b)$$

where $\nabla^2 = \nabla \cdot \nabla = \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplace operator.

To separate the variables, $\vec{E}(x, y, z, t) = \vec{E}(x, y, z) \cdot e^{i\omega t}$,we get the time-independent forms of the wave equation, the Helmholtz equation:

$$\nabla^2 \vec{E}(x, y, z) + \omega^2 \mu\varepsilon \vec{E}(x, y, z) = 0 \qquad (2.4.3)$$

If z direction is taken as the propagating direction, we further separate the variable, $\vec{E}(x, y, z) = \vec{E}(x, y) e^{-i\beta z}$. Then we have:

$$\nabla^2 \vec{E}(x, y) + [\omega^2 \mu\varepsilon - \beta^2] \vec{E}(x, y) = 0 \qquad (2.4.4)$$

where $\beta = k_z = \frac{2\pi n}{\lambda} \sin\theta_i$ [cm^{-1}] is the propagating constant and θ_i is the incident angle to the interface. The real part of the propagating constant is related to the change of the phase, while its imaginary part is related to the change of the amplitude.

The vectors are fully written as $\vec{E} = E_x\hat{x} + E_y\hat{y} + E_z\hat{z}, \vec{H} = H_x\hat{x} + H_y\hat{y} + H_z\hat{z}$.

According to whether the longitudinal components are zero, all possible modes can be divided into four groups:

i. transverse electromagnetic (TEM) modes: $E_z(x,y) = 0, H_z(x,y) = 0$
ii. transverse electric (TE) modes: $E_z(x,y) = 0, H_z(x,y) \neq 0$
iii. transverse magnetic (TM) modes: $E_z(x,y) \neq 0, H_z(x,y) = 0$
iv. hybrid modes: $E_z(x,y) \neq 0, H_z(x,y) \neq 0$

If the electromagnetic wave is an ideal plane wave, *i.e.*, $\frac{\partial E(x,y,z)}{\partial x} = 0$, $\frac{\partial E(x,y,z)}{\partial y} = 0$, the solution to equation (2.4.2a) is

$$E(z,t) \sim e^{i(\omega t - \beta z)} \tag{2.4.5}$$

If the propagating constant β is a real number, this wave will propagate infinitely without any decay. This is against the law of conservation of energy. So this plane wave can only exist in a finite space.

If the electromagnetic wave is restricted in a slab waveguide, *i.e.*, $\frac{\partial E(x,y,z)}{\partial y} = 0$, the solution for TE modes is

$$E = E_y(x,z,t) \sim e^{\pm ik_x x} e^{i(\omega t - \beta z)} \tag{2.4.6}$$

where $k_x = \sqrt{\omega^2\mu\varepsilon - \beta^2} = \sqrt{k_0^2 n^2 - \beta^2}$.

2.4.2 Light propagating in a multilayer structure

It's well known that the incident angle and refraction angle satisfy the Snell's law:

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \tag{2.4.7}$$

Total reflection happens when the incident angle is large enough and $n_1 > n_2$. The critical angle $\theta_c = \arcsin(\frac{n_2}{n_1})$. Depending on whether the incident angle is greater than the critical angle, the light propagating process can generally be grouped into two cases:

1) If the incident angle is small, the light will propagate along the normal of a plane interface. The reflection and transmission coefficients at the interface are given by the Fresnel formulas.
2) If the incident angle is larger than the critical angle, the light will propagate along the parallel of a plane interface. However, not all optical modes with $\theta_i > \theta_c$ can form stable waveguide modes. From the viewpoint of geometric optics, the waveguide modes are those with incident angles satisfying the constructive interference condition. The mode order is indicated by a **modal index** $N_{\text{mod}} = \beta / k_0$, which is also widely called as **effective index N_{eff}**.

The transfer matrix formulas for these two cases are similar. The only difference in case (2) is that a scan of effective index is performed to find the eigenvalue corresponding to the waveguide mode. For a TE mode wave propagating in a multilayer structure as shown in Figure 2-8, the expression is similar to (2.4.6), and can be rewritten as:

$$E_y(x) = A\mathrm{e}^{-ik_x \cdot x} + B\mathrm{e}^{ik_x \cdot x} = A\mathrm{e}^{-\gamma \cdot x} + B\mathrm{e}^{\gamma \cdot x} \tag{2.4.8}$$

where the damping constant $\gamma = \sqrt{\beta^2 - \omega^2 \mu\varepsilon} = k_0 \sqrt{N_{eff} - n^2} = ik_x$ describes the exponential decay of the optical wave. The effective index is a constant value in all

layers. For unification of case (1) and case (2), the incident angle can actually be converted to the effective index by

$$N_{eff} = \frac{\sqrt{k^2 - k_x^2}}{k_0} = n_1 \sin\theta_1 \tag{2.4.9}$$

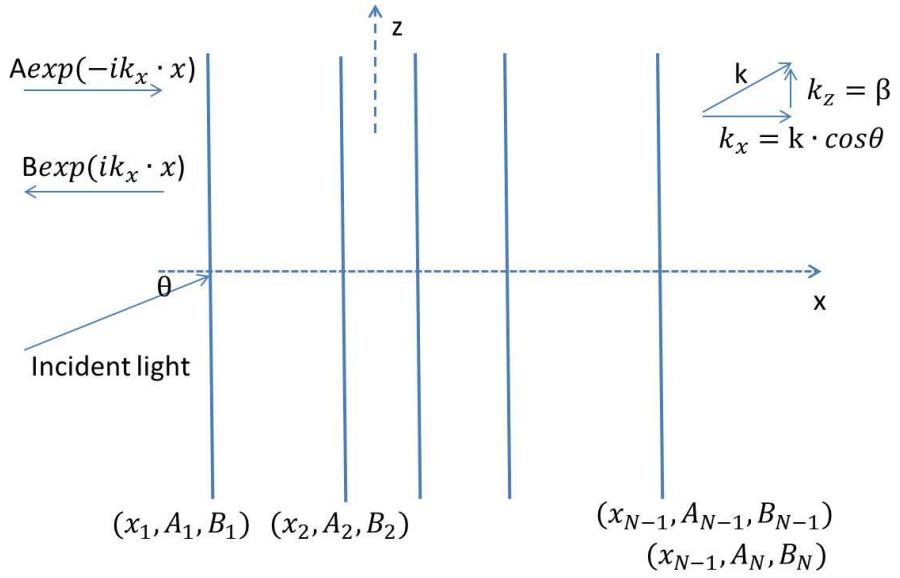


Figure 2-8. Propagation of light through a multilayer structure

According to Maxwell's equations (2.4.1a,b), if there is no net charges and no currents, the tangential components of E and H are all continuous at the interface. Therefore, E_y and H_z are both continuous at the interface. For TE modes, the existing components are E_y, H_x, and H_z. Because $H_z = \frac{i}{\omega\mu}\frac{\partial E_y}{\partial x} \propto \frac{\partial E_y}{\partial x}$, the matrix format of E_y and H_z can be written as:

$$\begin{pmatrix} E_y(x) \\ \frac{\partial E_y}{\partial x}(x) \end{pmatrix} = \begin{pmatrix} Ae^{-\gamma x} + Be^{+\gamma x} \\ -\gamma Ae^{-\gamma x} + \gamma Be^{+\gamma x} \end{pmatrix} = \begin{pmatrix} e^{-\gamma x} & e^{+\gamma x} \\ -\gamma e^{-\gamma x} & \gamma e^{+\gamma x} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} \tag{2.4.10}$$

Specifically, in each layer:

$$
\begin{aligned}
&E_1(x) = \begin{pmatrix} e^{-\gamma_1(x-x_1)} & e^{+\gamma_1(x-x_1)} \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, 0 \le x \le x_1 \\
&\cdots \\
&E_{N-1}(x) = \begin{pmatrix} e^{-\gamma_{N-1}(x-x_{N-1})} & e^{+\gamma_{N-1}(x-x_{N-1})} \end{pmatrix} \begin{pmatrix} A_{N-1} \\ B_{N-1} \end{pmatrix}, x_{N-2} \le x \le x_{N-1} \\
&E_N(x) = \begin{pmatrix} e^{-\gamma_{N-1}(x-x_{N-1})} & e^{+\gamma_{N-1}(x-x_{N-1})} \end{pmatrix} \begin{pmatrix} A_N \\ B_N \end{pmatrix}, x_{N-1} \le x \le x_N
\end{aligned}
\tag{2.4.11}
$$

By applying the continuity condition, the amplitudes (A, B) at each interface are connected by

$$
D_1(x_1) \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = D_2(x_1) \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{2.4.12}
$$

By continuous substituting the (A,B) pair in all the intermediate layers, we get

$$
\begin{pmatrix} A_N \\ B_N \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} \tag{2.4.13}
$$

For the case (1), an incoming wave from the left, as in Figure 2-8, will require B_N=0. Thus, the reflection and transmission coefficients are:

$$
\begin{aligned}
r_s &= \frac{B_1}{A_1} = -\frac{M_{21}}{M_{22}} \\
t_s &= \frac{A_N}{A_1} = M_{11} - M_{12} \cdot \frac{M_{21}}{M_{22}}
\end{aligned}
\tag{2.4.14}
$$

The energy flux is:

$$
\vec{S} = \frac{1}{2} \mathrm{Re}(\vec{E} \times \vec{H}^*) = \frac{\mathrm{Re}(\vec{k})}{2\omega\mu} |\vec{E}|^2 \tag{2.4.15}
$$

Therefore, the reflectance and transmittance are:

$$R = \left|\frac{\vec{x}\cdot\vec{S}_R}{\vec{x}\cdot\vec{S}_I}\right| = \frac{\mathrm{Re}(k_{1x})}{\mathrm{Re}(k_{1x})}|r_s|^2 = |r_s|^2$$
$$T = \left|\frac{\vec{x}\cdot\vec{S}_T}{\vec{x}\cdot\vec{S}_I}\right| = \frac{\mathrm{Re}(k_{Nx})}{\mathrm{Re}(k_{1x})}|t_s|^2 \qquad (2.4.16)$$

For TM modes, $E_z = \frac{-i}{\omega\varepsilon}\frac{\partial H_y}{\partial x} \propto \frac{\partial H_y}{\partial x}$, the matrix format is similar:

$$\begin{pmatrix} H_y(x) \\ \frac{1}{\varepsilon}\frac{\partial H_y}{\partial x}(x) \end{pmatrix} = \begin{pmatrix} e^{-\gamma x} & e^{+\gamma x} \\ -\frac{\gamma}{\varepsilon}e^{-\gamma x} & \frac{\gamma}{\varepsilon}e^{+\gamma x} \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} \qquad (2.4.17)$$

Note that the electric field of a TM mode has two components: E_x and E_z. From Maxwell equation (2.4.1d), the total electric field is:

$$|\vec{E}| = \frac{1}{\omega\varepsilon}|\nabla\times\vec{H}| = \frac{1}{\omega\varepsilon}|\hat{i}\beta H_y + \hat{k}k_x H_y| = \frac{k}{\omega\varepsilon}H_y \qquad (2.4.18)$$

Therefore, the reflection and transmission coefficients for TM modes are:

$$r_p = \frac{k_1/\varepsilon_1}{k_1/\varepsilon_1}\frac{B_1}{A_1} = \frac{B_1}{A_1}$$
$$t_p = \frac{k_N/\varepsilon_N}{k_1/\varepsilon_1}\frac{A_N}{A_1} = \frac{n_1}{n_N}\frac{A_N}{A_1} \qquad (2.4.19)$$

If there are only two layers (*i.e.*, one interface), then the transfer matrix is

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}^{TE} = \frac{1}{2\gamma_2}\begin{pmatrix} \gamma_2+\gamma_1 & \gamma_2-\gamma_1 \\ \gamma_2-\gamma_1 & \gamma_2+\gamma_1 \end{pmatrix}$$
$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}^{TM} = \frac{1}{2\gamma_2/\varepsilon_2}\begin{pmatrix} \gamma_2/\varepsilon_2+\gamma_1/\varepsilon_1 & \gamma_2/\varepsilon_2-\gamma_1/\varepsilon_1 \\ \gamma_2/\varepsilon_2-\gamma_1/\varepsilon_1 & \gamma_2/\varepsilon_2+\gamma_1/\varepsilon_1 \end{pmatrix} \qquad (2.4.20)$$

If we substitute the matrix elements in (2.4.14) and (2.4.19), the Fresnel equations are obtained.

2.4.3 Waveguide mode

For case (2) in section 2.4.2, the optical wave is confined in the core layer of a higher refractive index and decays exponentially towards the outer layer of a lower refractive index. The requirement for a waveguide mode is that no optical wave comes from both outer layers. Equation (2.4.13) becomes:

$$\begin{pmatrix} A_N \\ 0 \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} 0 \\ B_1 \end{pmatrix} \rightarrow M_{22} = 0 \tag{2.4.21}$$

For a three-layer slab waveguide (the refractive indices are n_1, n_2, and n_3, respectively), we obtain a transcendental equation:

$$\tan(k_{2x}d) = \frac{\frac{\gamma_1}{k_{2x}} + \frac{\gamma_3}{k_{2x}}}{1 - \frac{\gamma_1}{k_{2x}}\frac{\gamma_3}{k_{2x}}} \tag{2.4.22}$$

To reveal its physical insight, we rewrite it as:

$$k_{2x}d = m\pi + \arctan\frac{\gamma_1}{k_{2x}} + \arctan\frac{\gamma_3}{k_{2x}} \tag{2.4.23}$$

where m is the mode order, $\arctan\frac{\gamma_1}{k_{2x}}$ and $\arctan\frac{\gamma_3}{k_{2x}}$ are the phase shift during the total reflection. From the perspective of the ray optics, equation (2.4.23) is a condition of constructive interference. This discovery was first made by F. Goos and H. Hanchen in 1943 [101]. Because the incident angle $\theta_2 = arc\sin(N_{eff} / n_2) < 90°$, the optical wave is actually zigzagging along the cavity, which is different from our intuitive picture of a straight line propagation.

The cut-off waveguide mode is achieved when $N_{eff} \rightarrow n_1 \; or \; n_3$. Assuming $n_1=n_3$, equation (2.4.23) in the cut-off limitation becomes:

$$\frac{2d}{\lambda}\sqrt{n_2^2 - n_1^2} = m \tag{2.4.24}$$

The number of waveguide modes that a waveguide can support is

$$m+1 = floor(\frac{2d}{\lambda}\sqrt{n_2^2 - n_1^2}) + 1 = floor(\frac{2d}{\lambda} n_2 \sqrt{1 - \frac{n_1^2}{n_2^2}}) + 1 \tag{2.4.25}$$

Assuming a semiconductor waveguide with $d = \lambda, n_2 = 3.5, \frac{n_1 - n_2}{n_2} \sim 5\%$, the allowed number of waveguide modes is 3. In QW lasers, the active region (*e.g.*, d<10 nm) is much smaller than the vacuum wavelength (*e.g.*, λ>1 μm), so it is easy to obtain a single spatial mode with "diffraction limited" beam quality along the growth direction. In contrast, the lateral width (*e.g.*, w>10 μm) is usually much larger than the vacuum wavelength. For single spatial mode operation, the laser ridge is fabricated as narrow as possible.

In the practical calculation for a multiple-layer structure, the minimum values of M_{22} are found by scanning a range of possible values of N_{eff}. For semiconductors with a complex refractive index, the free-carrier absorption loss is given by

$$\alpha_{fc} = 2\,\mathrm{Im}(\beta) = \frac{4\pi}{\lambda_0}\mathrm{Im}(N_{eff}) \tag{2.4.26}$$

Because the optical wave has a penetration depth γ^{-1} into the cladding layer, not all the photons participate in the amplification of the stimulated emission. The portion confined within the active region is described by the confinement factor Γ_{active}:

$$\Gamma_{active} = \frac{\int_{active} |E_y(x)|^2 \, xdx}{\int_{all} |E_y(x)|^2 \, xdx} \tag{2.4.27}$$

2.4.4 Far-field patterns

As the propagating wave reaches a cavity facet, what happens is more than the reflectance or transmission of a plane wave, as described in the case (2) of section 2.4.2. This is because the vertical waveguide has a size comparable to the wavelength, which will lead to significant diffraction. The vertical direction is also called the fast axis due to its fast divergence. The diffraction is actually a superposition of all the emitting plane waves:

$$E_y^{out}(x,z)=\int_{-1}^{1}A_u e^{-i\left(uk_0x+k_0\sqrt{1-u^2}z\right)}du=\int_{-1}^{1}A_u e^{-i\left(u\sin\theta+k_0\sqrt{1-u^2}\cos\theta\right)k_0r}du=E_y^{out}(r,\theta) \quad (2.4.28)$$

where A_u is the amplitude of uth wave with $k_x=uk_0=k_0\sin\varphi$. It can be calculated by performing a Fourier transform of the emitting electric field:

$$A_u=\frac{1}{2\pi}\int E_y^{out}(x,0)e^{iuk_0x}d(k_0x) \quad (2.4.29)$$

With the far-field approximation $k_0r>>1$ and the stationary phase method, the angular component can be expanded around $u_0=\sin\theta$ to the first order:

$$E_y^{out}(r,\theta)\approx\int_{-\infty}^{\infty}A_{u_0}\exp[-i(1-\frac{v^2}{\cos\theta})k_0r]dv=A_{u_0}\sqrt{\frac{\lambda}{r}}\cos\theta e^{-ik_0r+i\frac{\pi}{4}} \quad (2.4.30)$$

With continuous boundary condition $E_y^{out}(x,0)=E_y^{in}(x,0)$, the far-field distribution is

$$E_y^{out}(r,\theta)=\frac{\cos\theta}{\sqrt{\lambda r}}e^{-ik_0r+i\frac{\pi}{4}}\int_{all}E_y^{in}(x,0)e^{-ik_0x\sin\theta}dx \quad (2.4.31)$$

$$E_{far}(\theta)\equiv\int_{all}E_y^{in}(x,0)e^{-ik_0x\sin\theta}dx, I_{far}(\theta)=\frac{\cos^2\theta}{\lambda r}\left|E_{far}(\theta)\right|^2 \quad (2.4.32)$$

Another approach, considering Huygens's obliquity factor for both forward and backward radiation from the facet aperture, leads to a shrinkage of the beam divergence:

$$I_{far} = O_{TE}(\theta)\left|E_{far}(\theta)\right|^2 = \cos^2\theta\left|\frac{1+\sqrt{N_{eff}{}^2-\sin^2\theta}}{\cos\theta+\sqrt{N_{eff}{}^2-\sin^2\theta}}\right|^2\left|E_{far}(\theta)\right|^2 \qquad (2.4.33)$$

where $O_{TE}(\theta)$ is the obliquity factor for TE mode. This approach gives a better agreement with experimental measurement [102] than the one without the obliquity factor [103].

Chapter 3 Material growth, fabrication and laser characterizations

3.1 Material growth

As the first step of the whole production process, the material growth is vital to the laser performance. The building blocks of the interband cascade (IC) laser are InAs, GaSb, and AlSb. The lattice constants of these three materials are 6.0584, 6.0959 and 6.1355 Å, respectively. These materials are commonly known as "6.1 Å family." They are nearly lattice-matched to each other, so the laser structure can be grown on either GaSb or InAs substrates.

The quantum structure of an active core within a laser requires the precision of 0.1 Å, which is only technically accessible by molecular beam epitaxy (MBE) . The typical growth rates are 0.3 monolayer (ML)/s or 0.38 μm/h for quantum wells (QWs), and 1 ML/s or 1.26 μm/h for bulk layers [104,105]. For a typical laser structure, the growth time is about 10 hours. The technical details of the growth are beyond the scope of this chapter.

For actual structure, any material with a lattice constant different from the substrate will cause strain. If the strain is accumulated to a critical point, dislocations or defects will form, which leads to degradation of the material quality. The overall strain of all layers can be balanced out by the sum of tensile strain and compressive strain. Each single layer has its own critical thickness to avoid the formation of the defect. Using the Energy Balancing or Force Balancing approach, the critical thickness for single strained layer is given by Matthews-Blakeslee formula [106]:

$$H_c = \frac{b(1-v\cos^2\theta)}{8\pi f(1+v)\cos\varphi}\left[\ln\left(\frac{H_c}{b}\right)+1\right] \tag{3.1.1}$$

where $b = a_s / \sqrt{2}$ is the magnitude of the Burger's vector of the dislocation, f is the absolute value of lattice mismatch, θ is the angle between the dislocation line and its Burger's vector, φ is the angle between the slip direction and that direction in the film plane which is orthogonal to the line of intersection of the slip plane and the interface. Poisson's ratio $v = \frac{C_{12}}{C_{11} + C_{12}}$. For a slip plane, $\theta = \varphi = 60^o$.

Applying the above approach, the maximum thickness of a single epitaxial layer grown on InAs can be obtained. Within the 6.1 Å family, GaSb and AlSb have a compressive strain of -0.617% and -1.26%, resulting in a critical thickness of 185 Å and 71 Å, respectively. In the active region, the ternary material GaInSb is used as the host layer for holes. If a linear interpolation is used to obtain the lattice constant and elastic stiffness coefficients, $Ga_{0.7}In_{0.3}Sb$ has a compressive strain of -2.46% and a critical thickness of 27.6 Å, and $Ga_{0.6}In_{0.4}Sb$ has a compressive strain of -3.06% and a critical thickness of 19.4 Å. All these materials contribute to the compressive strain, which requires tensile strain to achieve sufficient strain balance. For this purpose, GaAs and AlAs are inserted into the injection region. They have an average tensile strain of 7.3% and an average critical thickness of 3.4 Å. This thickness is only about one monolayer.

Bennett pointed out that the experimental critical thickness could be an order higher than the theoretical value [107]. Actually, most of the InAs-based IC lasers have a 30-Å-thick $Ga_{0.7}In_{0.3}Sb$ or $Ga_{0.65}In_{0.35}Sb$ layer in the active region [66,71]. Currently, no study has been performed on this discrepancy.

3.2 Device fabrication

A typical fabrication process has several cycles of applying photoresist, photolithography, and etching. These cycles use different masks for different purposes such as forming the ridge, depositing an insulating layer, opening contact windows and depositing metal contacts. The technical details and specific steps can be found in the appendix of reference [81].

There are two sets of fabrication flow, according to the width of the ridge. A laser with a ridge with larger than 100 μm is called a broad-area (BA) device; while a laser with a ridge width less than 30 μm is referred to as a narrow ridge (NR) device.

The broad-area device is relatively easy to fabricate due to the absence of an insulator layer on the ridge sidewall. In addition, larger size helps to achieve a more uniform fabrication result and less uncertainty caused by the wet etching. So the broad-area devices can be used in a benchmark test (*i.e.*, pulsed threshold current density at 300 K), allowing quick feedback to the material grower and the structure designer.

In contrast, the fabrication of narrow-ridge devices requires a more precise control of the etching process, which results in a much lower yield. Compared with the broad-area device, the narrow ridge device requires a less current injection to reach the threshold and therefore generates less internal heat. Consequently, the narrow ridge laser is able to work at a much higher temperature in continuous wave mode.

Reducing device the size is one strategy to decrease the generated heat. Another strategy is to increase the thermal dissipation. For this purpose, a thick gold layer of several micrometers can be electroplated on top of the metal contact. A more efficient

way to conduct heat is epi-down mounting [63]. In this way, the generated heat can directly flow into the heatsink without going through the thick substrate.

3.3 Laser characterizations

There are many aspects of performance for a laser, such as threshold current density, threshold voltage, maximum operating temperature in pulsed/CW mode, characteristic temperature T_0, maximum output power at specific temperatures, slope efficiency, lasing wavelength, tuning range by current/temperature, thermal resistance, differential resistance, waveguide loss, differential gain and group refractive index. With the restriction of time, it's impossible to characterize every single device fully. In this section, we introduce some core characterization techniques which are essential to our goals in the research projects.

3.3.1 Light-current-voltage characteristic

Emission spectrum and light-current-voltage characteristic are two basic characterizations for a laser. While emission spectrum is an accurate way to determine the lasing wavelength and threshold current, light-current-voltage characteristic can be used directly to obtain information such as output power, slope efficiency, and differential resistance. The light-current-voltage characteristic contains rich physical information of a working laser above its threshold, which is intensively explored [108].

The electrical measurement is an efficient way to analyze its transport property. The current-voltage characteristic of an ideal diode is known to follow the Shockley diode equation [109]:

$$I = I_s\left[\exp(\frac{eV}{nk_BT}) - 1\right] \tag{3.3.1}$$

where $I_s = eA\left(\sqrt{\frac{D_p}{\tau_p}}\frac{n_i^2}{N_D} + \sqrt{\frac{D_n}{\tau_n}}\frac{n_i^2}{N_A}\right)$ is the reverse bias saturation current. The ideality factor n equals to 1 for the ideal diffusion current.

This equation was originally proposed for p-n homojunction. However, the exponential increases in the current have also been found in heterojunctions, light emitting diodes (LEDs), and quantum well lasers, even unipolar quantum cascade lasers. This is a strong indication that a **space charge region** exists in every semiconductor laser. The space charge region produces a potential that is called diffusion voltage $V_D = \frac{kT}{e}\ln\frac{N_A N_D}{n_i^2}$. Hence, the Shockley equation can be rewritten as [110]:

$$I = eA\left(\sqrt{\frac{D_p}{\tau_p}}N_A + \sqrt{\frac{D_n}{\tau_n}}N_D\right)\left[\exp(\frac{e(V - V_D)}{nk_BT}) - 1\right] \tag{3.3.2}$$

Because eV_D is roughly equal to the bandgap, equation (3.3.2) explains why most LEDs have a turn-on voltage that is close to the bandgap. The electrical characteristics of DH lasers had been analyzed by early researchers [111-113]. Similar research should also be carried on cascade lasers to study the carrier transport mechanism.

After a laser reaches the threshold, the stimulation recombination becomes the dominant mechanism, consuming almost all the excess carriers besides the threshold conditions. Consequently, the carrier concentration is pinned, as discussed in section 2.1.1, and the voltage ceases to increase. However, any actual V-I curve, excluding the

voltage drop on circuit resistance, has never been an ideal horizontal line. It is because the stimulation recombination, no matter how fast, still takes the time to achieve the dynamic balance between gain and loss. The loss, either electrical or optical, increases modestly as soon as more current is injected. As a result, the threshold is "inflated" and the voltage increases. Therefore, the actual differential resistance is always larger than zero.

The above insight also explains the light-current rollover phenomenon. The slope efficiency per facet per stage η_{slope} can be expressed as:

$$\eta_{slope}[W/A] = \frac{\eta_d E_{ph}}{E_e} = \frac{\eta_d 2\hbar\pi c}{\lambda[\mu m]} = 1.24\frac{\eta_d}{\lambda[\mu m]} \tag{3.3.3}$$

The differential quantum efficiency per facet η_d is related to the internal quantum efficiency η_i, internal loss α_i, and mirror loss α_m by:

$$\eta_d = \frac{1}{2}\eta_i\left(\frac{\alpha_m}{\alpha_i + \alpha_m}\right) \tag{3.3.4}$$

For natural cleaved facet without any treatment, the reflections (R) for both facets are the same. The mirror loss is:

$$\alpha_m = -\frac{\ln(R)}{L} \tag{3.3.5}$$

Hence, the differential quantum efficiency becomes a linear function of the cavity length L:

$$\frac{2}{\eta_d} = \frac{1}{\eta_i} - \frac{\alpha_i}{\eta_i \ln(R)}L \tag{3.3.6}$$

If uniform-quality lasers with different cavity lengths are available, this equation can be used to extract the relevant parameters such as internal efficiency η_i and internal loss α_i. There are two critical details:

(1) internal efficiency and internal loss are current density dependent. So the value of differential quantum efficiency should be taken at the same current density for different lengths.

(2) researchers often use the simple Fresnel formulas to get the reflection at the emitting facet. For example, if the effective index of the waveguide mode is 3.40, then the reflection $R=(3.4-1)^2/(3.4+1)^2=0.30$. However, this approach ignores the effect of diffraction, which is significant when the active core has a similar size to the vacuum wavelength. A more comprehensive calculation gives a reflection of ~0.35. The exact value depends on the specific structure. The value of 0.41 was used in [25]. A more practical and quick solution is to use the group refractive index, which can be inferred from the amplified spontaneous emission (ASE) spectrum:

$$n_g = \frac{2L}{\Delta\tilde{k}} \tag{3.3.7}$$

where $\Delta\tilde{k} = \frac{1}{\lambda_1} - \frac{1}{\lambda_2}$ is the difference of wavenumber between two adjacent peaks.

3.3.2 Double-modulation technique in pulse spectra

In CW operation, the biggest problem is that the laser generates more heat than it can dissipate until the temperature gradient is large enough to achieve the thermal equilibrium. Consequently, the actual temperature in the active region may be significantly greater than the temperature at the heatsink. On the one hand, making

ridge narrow (*e.g.*, a few micrometers), or using facet coating, can reduce the threshold current and therefore reduce the heat generation; on the other hand, using epi-down mounting or thick electroplated Au can improve the heat dissipation. However, all these strategies will increase the complexity of fabrication.

In contrast, the pulse operation can almost eliminate the heating problem and reveal the fundamental property of a laser. The pulse operation usually employs small duty cycles of 1%~5% with pulse widths of 0.2~1 μs. The less generated heat comes at the cost of the decreased laser light by the same ratio. As a result, the weak light is submerged by the ambient blackbody radiation, causing difficulty for the detection. For the extraction of the small signal in pulse spectra, a technique called "double modulation" is developed since the 1980s [114], and has been increasingly used in photoluminescence and electroluminescence measurement [115,116].

The experimental setup is shown in Figure 3-1. The essential part of a Fourier transform infrared (FTIR) spectrometer is a Michelson interferometer, featuring one fixed reference mirror and a second movable mirror, which passes repetitively through the zero-delay position corresponding to the balanced interferometer configuration. The detector, either inside or outside the FTIR, receives the coherent spatial superposition of the signals from the two mirrors as a function of their time-modulated relative delay. The resulting interferograms are then passed through a lock-in amplifier which filters out the random noise from the background. The "purer" signal is fed back to the electronic board of the FTIR for the final processing.

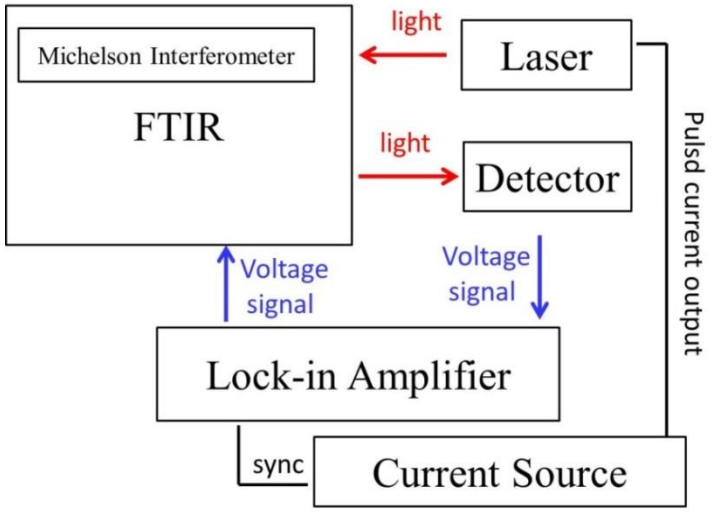


Figure 3-1. Double modulation setup

The technique is called "double modulation" because: (1) the low-frequency time-delay modulation provided by the scanning mirror; (2) the high-frequency modulation of the pulsed laser beam intensity given by the current source and extracted by the lock-in amplifier.

After the light beams pass through the Michelson Interferometer, the amplitude of the signal received by the detector is a superposition of two beams:

$$A(\lambda) \propto (1 + e^{-i\phi(\lambda)}) \tag{3.3.8}$$

where λ is the wavenumber of the investigated beam, and the phase difference $\phi(\lambda)$ is

$$\phi(\lambda) = 2\pi \frac{d}{\lambda} = 2\pi \frac{v_m t_m}{2\lambda} \tag{3.3.9}$$

where d is the optical path difference between the two interfering beams, v_m is the speed of the movable mirror; t_m is the time of moving for the mirror. The mirror moves back and forth between the zero path difference (ZPD) and the first destructive interference, *i.e.*, $\phi(\lambda) = \pi$. In these periodic cycles, the frequency is:

$$f_m = \frac{v_m}{\lambda} \tag{3.3.10}$$

For avoiding aliasing, the sampling rate should be at least larger than twice the maximum frequency of the signal. This minimum sampling rate is known as Nyquist rate. In FTIR, it is called Fourier frequency $f_{Fourier}$:

$$f_{Fourier} = 2 f_m = \frac{2v_m}{\lambda} \tag{3.3.11}$$

For a better restoration of the spectrum information, the external modulation frequency is recommended to be 5 times [116] or 10 times [117] of Fourier frequency in the rapid-scan mode. In contrast, in step-scan mode, the internal Fourier frequency is 0 Hz because the movable mirror stands still during the data acquisition. However, the step-scan mode is unsuitable for obtaining fine spectral features, because it typically requires several hours, during which the chirp of the sample and random effect are not controllable [116].

The actual setting depends on the target wavelengths and equipment capabilities. For example, the current source **AVTECH AV-107C-B** has a maximum modulation frequency of 5 kHz and the target wavenumber is 2500 cm^{-1}, so the mirror speed should be less than 1 cm/s to avoid possible fake signal. The lock-in amplifier is set at a time constant of 300 μs, at a slope of 6 dB/octave, and a sensitivity of 5 mV. Note that the most popular lock-in amplifier, SR 830, only has a fast output of X, which is phase dependent. The solution is, either to monitor the phase or to use another lock-in amplifier (*e.g.*, EG&G 7265 DSP) that has a fast output of R (amplitude).

3.3.3 Pulse power calibration

As discussed in 3.3.1, the power-current relation is very useful to extract the internal efficiency and internal loss. And such extraction can only be done in pulse operation to eliminate the heating effect. The pulse measurements employ a liquid-nitrogen cooled MCT detector and a lock-in amplifier. It's necessary to calibrate the actual power before each practical use [108].

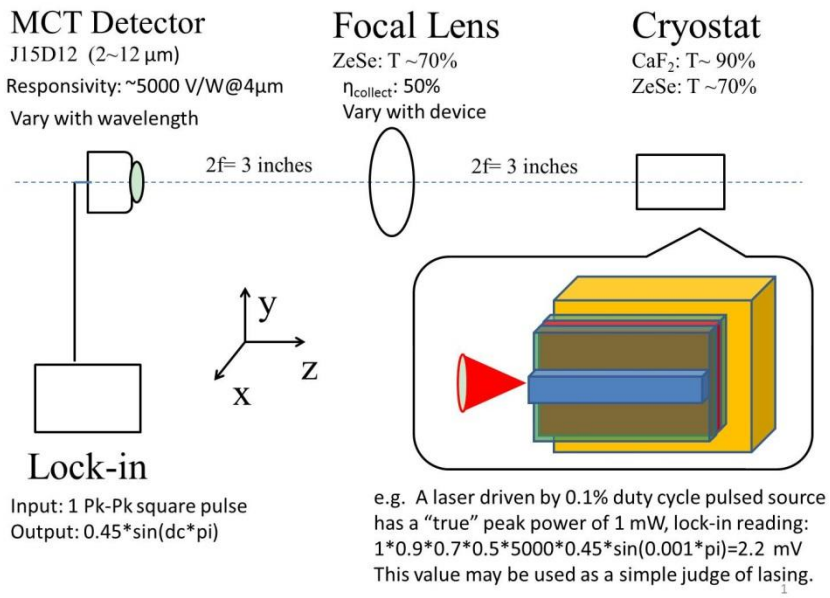


Figure 3-2. Experimental setup for pulse power measurement

The schematic experimental setup is shown in Figure 3-2. The reading displayed in lock-in amplifier is actually much smaller than the peak value of input signal. As proof in Appendix A, the reading for an input square pulsed wave with a peak-to-peak amplitude of 1 V is:

$$V_{lock-in} = 0.45\sin(\Lambda \cdot \pi) \tag{3.3.12}$$

where Λ is the duty cycle of the pulse.

Ideally, if there's no propagation loss, a peak power of P is 100% received by the detector with a responsivity R, the lock-in reading is:

$$V_{lock-in} = \mathrm{P}_{laser}\mathrm{R}_{MCT}\,0.45\sin(\Lambda\cdot\pi) = \mathrm{P}_{laser}\mathrm{g}_{ideal}\sin(\Lambda\cdot\pi) \tag{3.3.13}$$

where g_{ideal}=R*0.45 [V/W] is the ideal conversion factor.

However, the actual power received by the detector is reduced by non-ideal factors such as the transmission loss of the cryostat window and focal lens, and uncollected light by the significant divergence. All these factors can be packed into a term, collecting efficiency for lock-in amplifier $\eta_{lock-in}$:

$$V_{lock-in} = \mathrm{P}_{laser}\eta_{lock-in}\mathrm{g}_{ideal}\sin(\Lambda\cdot\pi) \tag{3.3.14}$$

For examples, at 4 μm, the transmission ratios of CaF_2 window and ZnSe lens are 90% and 70%, respectively. The collecting efficiency for the lens is about 50%, which will be discussed in more detail in Chapter 4. The responsivity of the MCT detector is 5100 V/W at 4 μm, according to the data sheet provided by the manufacturer. So the conversion factor $\eta_{lock-in}\mathrm{g}_{ideal} \approx 1600\ V/W$.

However, this is only a rough estimation. A more realistic approach is to use the thermopile power meter to calibrate. We also pack all possible loss into a term, η_{PM}, so the relation between peak power and power meter reading is

$$P_{PM} = P_{laser}\eta_{PM}\Lambda_2 \tag{3.3.15}$$

where Λ_2 is the duty cycle of the current source.

Due to the lack of a suiable current source (*e.g.*, short pulse, high frequency [108]), the calibration is done at 80 K to alleviate the heating effect. After getting two I-P curves at 80K by power meter and lock-in amplifier, respectively, the experimental g factor is:

$$g_{\exp} = \frac{V_{lock-in}/\sin(\Lambda_1\pi)}{P_{PowerMeter}/\Lambda_2} \xrightarrow[\Lambda_2=1]{\Lambda_1<<1} \frac{\eta_{slope,lock-in}}{\Lambda_1\pi\eta_{slope,PM}} \tag{3.3.16}$$

For bridging the gap between these two g factors, equation (3.3.14) and (3.3.15) are substituted into equation (3.3.16):

$$g_{\exp} = \frac{\eta_{lock-in}g_{ideal}}{\eta_{PM}} > \eta_{lock-in}g_{ideal} \tag{3.3.17}$$

This partially explains the discrepancy between the experimental g factor and the predicted g factor. The latter actually requires further calibration of the collecting efficiency and detector response, which adds more complexity and possible errors. So equation (3.3.16) should be used in the practical calibration , and equation (3.3.14) is used to understand the physical origin of the g factor.

3.3.4 Gain spectrum

For extraction of the internal efficiency and internal loss by equation (3.3.16), another challenge is the requirements of highly uniform material and device fabrication. To avoid these issues, Hakki and Paoli developed a method in 1973 [118] to extract the gain coefficient by using a single laser. A detailed derivation of this approach and clarification of the relevant concepts are presented as follows.

As mentioned in equation (2.4.6), the optical wave propagating in a laser waveguide is described as $E = E_0 e^{i(\omega t-\beta z)}$. For a passive waveguide, *i.e.*, no injection in laser, the imaginary part of the complex propagation constant represents the waveguide loss $\alpha_{waveguide}$, and we set it as negative.

$$\alpha_{waveguide} = -2imag(\beta) = -2k_0 imag(N_{eff}) \tag{3.3.18}$$

For simplicity, we only consider free carrier absorption loss as waveguide loss. Instead of solving equation (2.4.21), here we estimate the waveguide loss by the confinement factor Γ:

$$\alpha_{waveguide} = \Gamma \alpha_{active\ region} + (1-\Gamma)\alpha_{non-active\ region} \tag{3.3.19}$$

For a laser under injection, the loss can be compensated by the increased gain until it reaches the threshold condition:

$$\Gamma g = \alpha_{total} = \alpha_{mirror} + \alpha_{\text{int}} \tag{3.3.20}$$

where g is material gain and internal loss α_{int} is assumed to be equal to waveguide loss to simplify our discussion. To avoid confusion, Γg is termed "modal gain", and $g_{net} = \Gamma g - \alpha_{\text{int}}$ is termed "net modal gain". Thus, the complex propagation constant is $\beta = k + j\frac{1}{2} g_{net}$.

After a round trip, the amplitude of a photon $E_0(\lambda)$ generated by the spontaneous emission becomes

$$E_0 e^{g_{net} L} e^{i2kL} \sqrt{R_1 R_2} \tag{3.3.21}$$

Below threshold condition, $e^{g_{net} L} \sqrt{R_1 R_2} < 1$, the amplitude is a damped oscillation. However, it can still be amplified by constructive interference when $2k^+ L = 2\pi \cdot N$. The total electric field is:

$$E^+_{total}(\lambda^+) = E_0(\lambda^+) \sum_{n=0}^{\infty} [\sqrt{R_1 R_2}\, \exp(g^+_{net} L)]^n = \frac{E_0(\lambda^+)}{1 - \sqrt{R_1 R_2}\, \exp(g^+_{net} L)} \tag{3.3.22}$$

Alternatively, a photon with an adjacent wavelength λ^- may encounter destructive interference when $2k^-L = 2\pi \cdot N - \pi$. The total electric field for this wavelength is:

$$E_{total}^-(\lambda^-) = E_0(\lambda^-)\sum_{n=0}^{\infty}[-\sqrt{R_1R_2}\exp(g_{net}^-L)]^n = \frac{E_0(\lambda^-)}{1+\sqrt{R_1R_2}\exp(g_{net}^-L)} \tag{3.3.23}$$

For these two adjacent wavelengths, the spontaneous emission intensity and gain coefficient should be nearly equal: $\left|E_0(\lambda^+)\right|^2 = \left|E_0(\lambda^-)\right|^2$, $g_{net}^+ = g_{net}^-$. So we have

$$g_{net} = \frac{1}{L}\left[\ln\frac{E_{total}^+ - E_{total}^-}{E_{total}^+ + E_{total}^-} + \ln\frac{1}{\sqrt{R_1R_2}}\right] = \frac{1}{L}\ln\frac{r(\lambda)-1}{r(\lambda)+1} + \alpha_{mirror} \tag{3.3.24}$$

where $r(\lambda)$ is the depth of modulation [119,120]:

$$r(\lambda) = \frac{E_{total}^+}{E_{total}^-} = \sqrt{\frac{P_{\max}}{P_{\min}}} \approx \sqrt{\frac{P_{\max} + P_{\max}'}{2P_{\min}}} \tag{3.3.25}$$

where P_{max} and P_{min} are the consecutive maximum and minimum of the Fabry-Perot spectra intensity. P_{max} and P_{max}' are the maximum intensities of two adjacent longitudinal modes with the spacing

$$\Delta(\frac{1}{\lambda}) = \frac{1}{2n_gL} \tag{3.3.26}$$

where n_g is the group refractive index due to the dispersion of the optical medium.

After the gain spectra are obtained at various current levels, the peak gain can be fit to a linear equation:

$$g_{peak} = \Gamma g - \alpha_{\text{int}} = g_{diff}(I - I_{tr}) - \alpha_{\text{int}} \tag{3.3.27}$$

where $g_{diff} = \Gamma \frac{dg}{dI}$ is the differential modal gain per unit current and I_{tr} is the transparency current. At the transparency point $I=I_{tr}$, material gain g=0, and internal loss can be evaluated from the low-energy part of the gain spectra.

The lower and upper bounds can be used to understand the gain spectra. The lower limit of the net modal gain is the internal loss with a negative sign, and the upper bound is mirror loss with a positive sign. This is consistent with [121].

An example of gain spectra for an IC laser is shown in Fig.3-3. A polarizer is used to filter out the TM light. This laser has a different modal gain of 2.23 cm^{-1}/mA (per current) or 3 cm/A (per current density), which is comparable to 2.2 cm/A from a 12-stage IC laser with an early injector design [120] or 5.9 cm/A from a 10-stage IC laser [122]. However, a large leakage current exists for this device, leading to an overestimated transparency current density of 88 A/cm^2. As a result, the threshold current density is 96 A/cm^2, much higher than that (22 A/cm^2) of the best device from the same processed wafer.

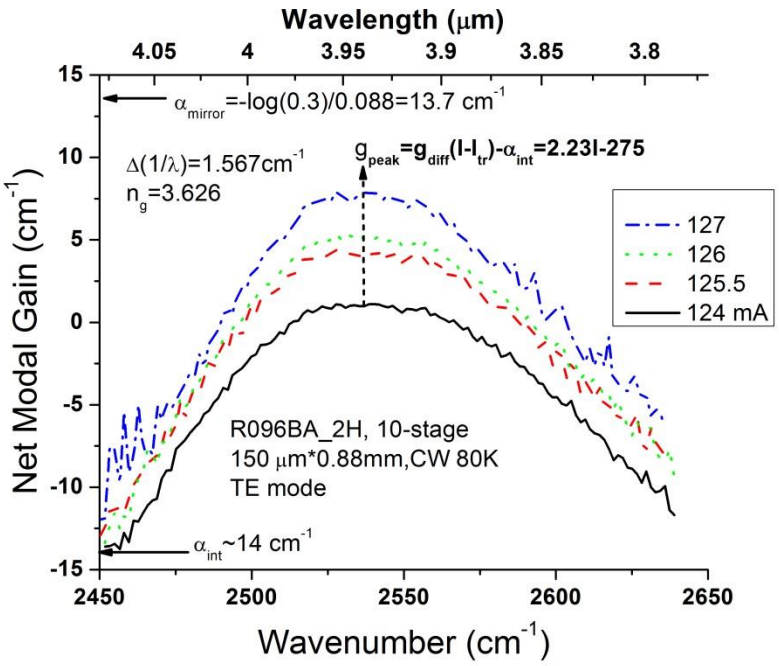


Figure 3-3. Gain spectra for an IC laser

However, the device is not well chosen in this example. In Hakki-Paoli's original paper [119], they picked a device with a cavity length of 0.38 mm and a ridge width of 10 μm, to reduce higher-order frequency modes and spatial modes. Besides, selecting a uniform and stable device is essential to obtain the "clean" spectra.

3.3.5 Thermal resistance

In a continuous-wave laser, the most severe problem is the large temperature difference between the heatsink and the laser's active core. Because the laser threshold is very sensitive to the temperature, the significant temperature increase can cause the laser cease to work. This problem can be described by thermal resistance:

$$R_{thermal}[K/W] = \frac{\Delta T}{\Delta P} = \frac{T_{active} - T_{\text{heatsink}}}{P_{\text{inject}} - P_{\text{radiate}}} \tag{3.3.28}$$

where $P_{\text{inject}} = UI$ is the power injected into the laser and $P_{radiate}$ is the power converted into radiation, including blackbody radiation, spontaneous emission and stimulated emission. Below threshold, more than 90% of the power doesn't convert into radiation at room temperature [123]. So the thermal resistance of a laser can be empirically determined by the following equation:

$$R_{thermal} = \frac{\Delta T}{U_{th} I_{th}} \tag{3.3.29}$$

For further insight, the thermal resistance should be related to a more fundamental material property, thermal conductivity k [W/(m·K)], which is defined by the Fourier's Law of thermal conduction:

$$J_P = -k \frac{dT}{dx} = -k \frac{\Delta T}{t} \tag{3.3.30}$$

where $J_P = \frac{P}{A} = \frac{P}{wl}$ is the heat (power) flux density. And t, w, l is the thickness, width, and length of a particular layer, respectively. Substitute (3.3.30) into (3.3.29), we have:

$$R_{thermal} = \frac{t}{kwl} \tag{3.3.31}$$

We immediately recognize the analogy between thermal resistance and electrical resistance. The physics behind them is that Ohm's law has the exactly same format:

$$J_e = -\sigma_e \frac{dU}{dx} \tag{3.3.32}$$

If the laser structure is considered as a thermal circuit, the contribution of each layer can be analyzed with the guidance of in-series connection. Table 3.1 shows an example of a narrow ridge IC laser with 10-stage. The value of conductivity for active region and superlattice is taken from[124] based on a 3ω method [125]. The total

thermal resistance for this laser (R145) is 40.6 K/W. Obviously, the cascade region and the InAs/AlSb superlattice layer are the major contributors to the total thermal resistance, due to their extremely low thermal conductivity. Hence, these two regions are better to be made thinner when the thermal issue has eaten up other benefits such as the modal gain.

Note that the specific thermal resistance R_{sth} is also widely used and can be derived by:

$$R_{sth} = \frac{t}{k} \tag{3.3.33}$$

While this physical quantity is convenient in comparison to the same amount of input power density J_P, it can't be used in a thermal circuit if the circuit components have different sizes.

As shown in Table 3-1, the estimated total thermal resistance for a typical device is about 40 K/W, larger than the experimental value of about 30 K/W. The overestimation is mainly because this one-dimensional model is oversimplified. First, the cladding layers actually have a larger width than the ridge, so the effective R_{th} for cladding layer should be much smaller. Second, in a narrow ridge laser, the ridge is covered by SiN and thick electroplated gold, which are equivalent to an additional thermal resistance connected to the thermal circuit in parallel. A better estimation is to use a two-dimensional model based on the finite element method. Published results show that the best thermal dissipation can be achieved by a buried structure mounted epilayer down on a diamond heat sink [126,127].

Table 3-1. Thermal resistance of each layer from cascade region to bonding area

region	material	k	thickness	width	length	R_{th}	R_{sth}
		W/mK	μm	μm	mm	K/W	Kcm^2/kW

cascade core	InAs/GaSb/AlSb	**1.55**	0.42	10	2	**13.5**	2.7
lower spacer	InAs	27	0.43	10	2	0.8	0.2
SCL	InAs/AlSb SL	**2.65**	1.03	10	2	**19.4**	3.9
lower cladding	InAs	27	1.62	10	2	3	0.6
substrate	InAs	27	100	500	2	3.7	
bonding	In	87	10	500	2	0.11	
Total						40.6	7.4

Possible confusion by equation (3.3.31) is that the larger ridge width will reduce the thermal resistance R_{th}. Then, why the narrow ridge laser has better thermal dissipation than the broad-area laser? The key is to distinguish power and power density. Ideally for any size of lasers (assume the same loss), they share the same threshold current density and thus the same input power density. But the larger ridge generates a larger amount of wasted energy, which has to go through the layers under the ridge, including cladding layer, substrate, heat sink. For example, these layers are packed into a term "$R_{th,sink}$". For a 10-μm-wide device with a threshold power of P, the temperature difference between the heatsink and ridge bottom is $P \cdot R_{th,sink}$; while for a 100-μm-wide device, this temperature difference will be 10 $P \cdot R_{th,sink}$. And the large temperature rise will increase the optical loss in that region, which again increase the threshold and add a positive feedback to the thermal power. So the broad-area laser has a much lower operating temperature in CW mode.

A drawback of Fourier's Law of thermal conduction (3.3.30) is that it doesn't include the speed of heat transport. In other words, it assumes an infinite speed of propagation of heat signal, which is against the theory of relativity. The theory of relativistic heat conduction [128] is proposed to avoid this problem,

In experiments, a low duty cycle (*e.g.*, 0.1%) and a short pulse (*e.g.*,200 ns) are used to avoid the thermal accumulation. However, researchers found that the

temperature of the active region increases immediately as soon as the pulsed current is injected [126,129,130]. To roughly quantify this self-heating in pulsed laser, I use a laser R144BA_1G, which is able to work up to 377 K in pulsed mode [73]. At 377 K, the threshold current is about 6A and threshold voltage on the device is about 4V. Assuming 200-ns-pulse power is entirely absorbed by the cascade region with a size of 0.4 μm×100 μm× 2 mm and the specific heat of 0.25 J/K/g, the temperature increase is 42 K (compare with 3.6 K at room temperature). An early theoretical analysis predicted an increase of 65 K at a current density of 4 kA/cm^2[131]. The actual temperature increase will be smaller than this value because the temperature gradient will drive the heat to the heatsink until a steady-state is achieved. This vast temperature increase under a high injection partially explains the pulsed power rollover or efficiency droop [108].

Chapter 4 Far-field patterns and beam qualities of plasmon-enhanced waveguide interband cascade lasers

Far-field patterns and beam qualities of interband cascade lasers are essential for The integration with optical systems. This chapter presents a detailed study of far-field patterns of plasmon-enhanced waveguide IC lasers with different ridge dimensions and metal contact configurations. Distinctive far-field patterns, corresponding to the fundamental and the first excited optical modes, are observed with different metal contact arrangements. The experimental results are well explained by the simulation. Furthermore, the concepts of brightness and quality factor are clarified, with examples of IC lasers given in the end.

4.1 Waveguide modeling

4.1.1 Plasmon-enhanced waveguide

The criterion of a good waveguide for a semiconductor laser are: (a) a large difference in refractive index between the active core layer and the cladding layer; (b) a small free carrier absorption loss; (c) a large thermal conductivity.

The cascade core of an IC laser, consisting of a “W” shape InAs/GaInSb active region, an InAs/AlSb electron injection region and a GaSb/AlSb hole injection region, has an equivalent refractive index of about 3.45. However, the substrate material, either GaSb (n=3.73) or InAs (n=3.49), has a refractive index larger than the cascade core. So they are not suitable to be used as the cladding layer to prevent the optical leakage .

Instead, they are used as a spacer layer to accommodate the wave expansion from the active QWs.

One solution is to use InAs/AlSb superlattice (SL) as the cladding layer on the GaSb substrate, as shown in Figure 4-1(a). In this structure, the refractive index contrast is Δn/n~3%, which can be further increased by inserting a GaSb spacer layer. The advantage of InAs/AlSb SL is that the doping level in SL is relatively small, *e.g.*, 1×10^{17} cm^{-3} [61], so the free carrier absorption loss is kept minimal.

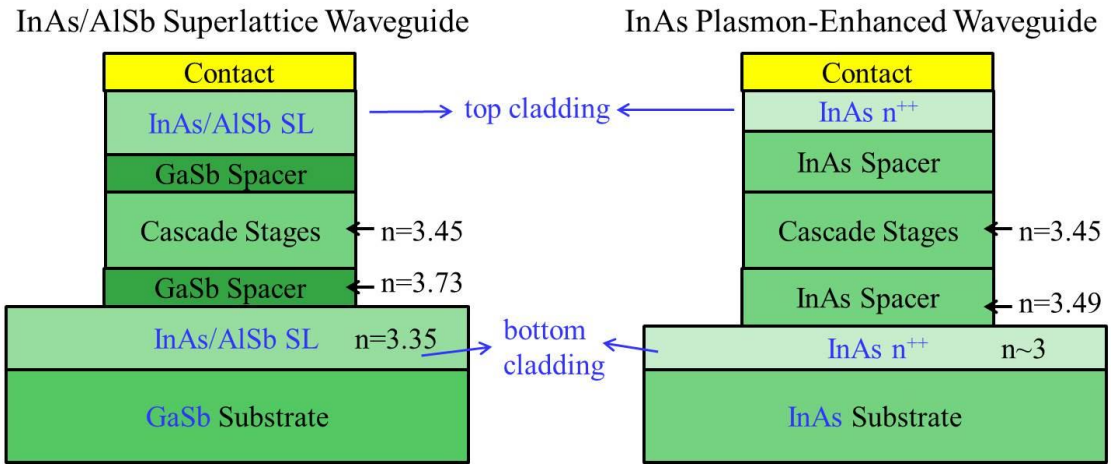


Figure 4-1. Two types of waveguides for IC lasers

Another solution is to use highly doped InAs (*e.g.*, $n>5\times10^{18}$ cm^{-3}) as the cladding layer on the InAs substrate, as shown in Figure 4-1(b). Such waveguide was originally called "**plasmon-enhanced**" waveguide [132] because the bulk plasmon effect in the highly doped InAs enhances the optical confinement by reducing the refractive index significantly, *e.g.*, to ~3. In this scenario, the refractive index contrast Δn/n is ~13%, which can efficiently prevent the leakage of laser light. Compared to the SL, the highly doped InAs has a larger thermal conductivity and is much easier to grow

by MBE. The latter becomes a principal advantage when an IC laser is extended to a longer wavelength which requires a thicker cladding layer.

Figure 4-2 shows a comparison of optical intensity and index profiles for two waveguides, both of which are close to their respective optimums. We can see that the SL waveguide has a smaller waveguide loss. In contrast, the spacer layer in the plasmon-enhanced waveguide is kept thick enough to reduce the penetration of optical waves into the high-loss InAs n^{++} layer. As a result, the plasmon-enhanced waveguide has a smaller confinement factor of the cascade core and a larger waveguide loss compared with the SL waveguide, which was not explicitly pointed out by earlier researchers [33,133].

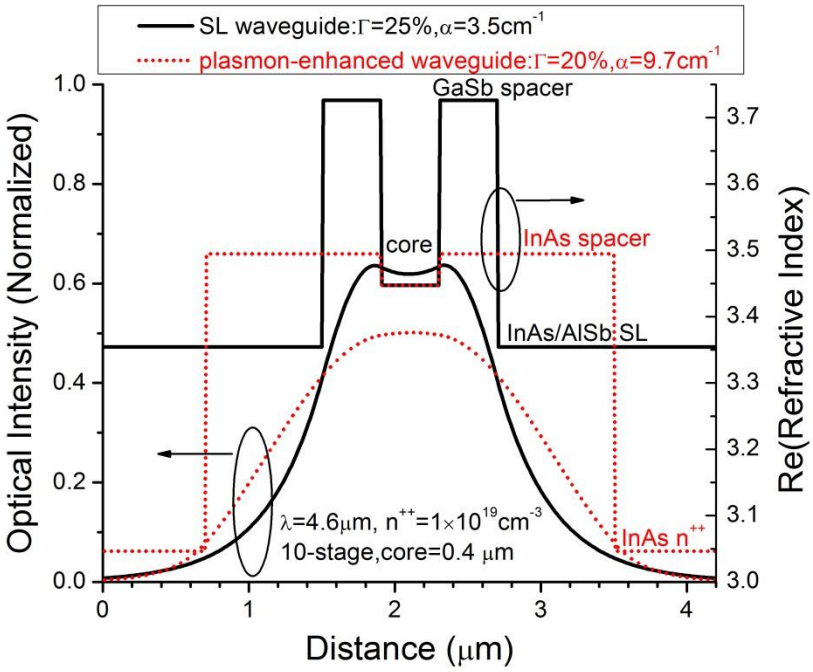


Figure 4-2. Comparison of optical intensity and index profiles for two waveguides

Historically, the first "plasmon-enhanced" waveguide was reported in 1995 for QC lasers at 8.4 μm [132]. The top InGaAs layer was n-type doped to 7×10^{18} cm^{-3} to suppress the coupling between the fundamental waveguide mode and the high-loss surface plasmon mode propagating along the interface between the metal contact and semiconductor. At longer wavelengths, metal had been used as the cladding layer [134] [135].

Note that the plasmon-enhanced waveguide is sometimes called "plasmon waveguide" for short, which, however, may cause confusion. Because "plasmon waveguide" widely refers to a totally different waveguide, the **plasmonic** waveguide. In principle, the "plasmon-enhanced" waveguide is still an optical waveguide [136] which confines light in a high index region. On the other hand, the plasmonic waveguide confines light exactly at the metal surface. Further discussion on the plasmonic waveguide is given in Appendix C.

4.1.2 Calculation of refractive index

The refractive index of an intrinsic material at the specific wavelength is a well-established constant [137]. However, the refractive index of a doped semiconductor, which changes with the doping concentration and wavelength, is not readily available in the literature. The Lorentz-Drude model is a useful strategy to solve this problem.

According to the Lorentz model, which is within the framework of classical mechanics, an electron interacting with electromagnetic field can be approximated by a damped oscillator:

$$m\frac{d^2}{dt^2}\vec{r}(t)+m\gamma\frac{d}{dt}\vec{r}(t)+m\omega_0{}^2\vec{r}(t)=-e\vec{\mathbf{E}}(t) \tag{4.1.1}$$

where γ is the damping constant, ω_0 is resonant frequency of electron motion. The restoring electric field by the positively charge background of atomic core $\vec{E}=\frac{Ne\vec{x}}{\varepsilon_0}$.

After solving equation (4.1.1), we can obtain the relative permittivity ε_r or dielectric constant, given by:

$$\varepsilon_r=\frac{\varepsilon}{\varepsilon_0}=1-\frac{Ne^2}{\mathrm{m}\varepsilon_0}\frac{1}{\omega^2-\omega_0^2-i\omega/\tau} \tag{4.1.2}$$

In the Drude model, the electrons are regarded as free electrons without restoration force, so ω_0=0, equation (4.1.2) becomes:

$$\varepsilon_r=1-\frac{Ne^2}{\mathrm{m}\varepsilon_0}\frac{1}{\omega^2-i\omega/\tau} \tag{4.1.3}$$

With the consideration of the potential field in a semiconductor, the effective mass of the electron is used. The complex refractive index is:

$$\left(n+ik\right)^2=\varepsilon_r=\varepsilon_\infty\left[1-\frac{\omega_p^2}{\omega^2-i\omega/\tau}\right] \tag{4.1.4}$$

where n is the real part of the refractive index, *k* is the imaginary part, also known as extinction coefficient, ε_∞ is the high-frequency limit of dielectric constant (which equals one for metal), $\tau=1/\gamma$ is the electron scattering time, ω_p is the **plasma frequency**:

$$\omega_p^2=\frac{Ne^2}{\mathrm{m}^*\varepsilon_\infty\varepsilon_0} \tag{4.1.5}$$

where N is the electron density, m^* is the effective mass of the electron. According to this equation, any semiconductor can exhibit certain plasmon effect if a high level of doping concentration is possible, which is also true for the large bandgap material like GaAs [121].

In (4.1.4), ε_∞ is taken to be 12.25 for InAs [138]. However, the values for effective mass and time constant are not readily available due to their dependence on doping concentration. In a narrow bandgap material, the band nonparabolicity has a strong effect on the effective mass m^*, which is given by [139]:

$$\frac{\hbar^2}{m^*(k)} = \frac{1}{k}\frac{dE(k)}{dk}\bigg|_{k=k_F} \tag{4.1.6}$$

where the energy-wavevector dispersion curve E(k) is calculated by a rigorous eight band Kane model [91]. Note that equation (4.1.6) is different from the conventional definition, $\frac{\hbar^2}{m^*(k)} = \frac{d^2E(k)}{dk^2}$. The term $m^*(k)$ is evaluated at the Fermi wavevector k_F given by [139]:

$$k_F^3 = 3\pi^2 N \tag{4.1.7}$$

For example, the electron densities of 5×10^{18}, 1×10^{19} and 2×10^{19} cm^{-3} correspond to $\varepsilon_F = \frac{k_F^2\hbar^2}{2m_0}$ of 0.011, 0.017 and 0.027 eV, respectively.

Figure 4-3 shows the energy-wavevector dispersion curve and energy-dependent effective mass in the conduction band of InAs. The 2-band model is essentially the approximation to the 8-band model with the assumption $\Delta >> E_g$ [91]. The parameter E_p

is evaluated at the conduction band edge to keep self-consistence. We can see the two sets of curves are very similar. For a typical doping of 1×10^{19} cm^{-3}, the effective mass of InAs is about $0.066*m_0$. Other researchers using a set of nonparabolicity coefficients obtained a slightly smaller effective mass [140].

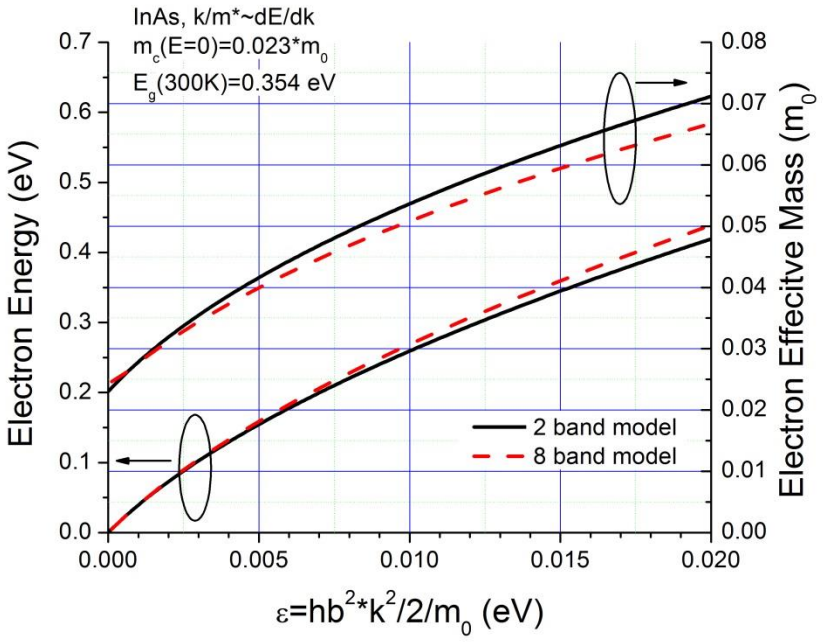


Figure 4-3. Energy E(k) and effective mass m*(k) as a function of wavevector

The other remaining unknown parameter is the damping time constant, also known as scattering lifetime, which can be inferred from its relation to mobility:

$$\tau = \frac{\mu_e \mathrm{m}_e^*}{\mathrm{e}} \tag{4.1.8}$$

The mobility is often characterized by an empirical model, such as the Caughey-Thomas formula [141]:

$$\mu = \frac{\mu_{\max} - \mu_{\min}}{1+(N/N_{ref})^{\alpha}} + \mu_{\min} \tag{4.1.9}$$

where all the fitting parameters are derived by trial and error from extensive experimental data. The actual mobility may be affected by the material quality and compensation factor [142]. One group has compiled the published data of InAs in the Caughey-Thomas formula [143]:

$$\mu = \frac{3.3\times10^{4}}{1+(N/1.1\times10^{18})^{0.32}} + 1000 \qquad (4.1.10)$$

However, this equation predicts that the time constant is increasing when the electron density is larger than 10^{17} cm^{-3}, which is not a consistent with the expectation. Instead, Baranov [144] provided an equation that is able to predict a decreasing trend:

$$\mu_e[cm^2/Vs] = \frac{3\times10^{4}}{1+(\frac{N}{8\times10^{17}})^{0.75}} \qquad (4.1.11)$$

With all the parameters at hand, the complex refractive index of InAs is calculated at a wavelength from 3 to 13 μm, as shown in Figure 4-4. The arrows indicate three plasmon wavelengths of 12 μm, 9.5 μm, and 7.4 μm, corresponding to three typical doping concentration of 0.5×10^{19}, 1×10^{19} and 2×10^{19} cm^{-3}. As the optical wavelength approaches the plasmon wavelength, the refractive index decreases sharply, and the free carrier absorption loss is even more dramatically increased. This implies that the benefit of a better wave confinement always comes at a greater cost of more loss, which may be a fundamental reason impeding the development of plasmon-waveguide based IC lasers. The long-wavelength IC lasers have lasing wavelengths near the plasmon wavelength [68], and their unexpectedly high threshold may be related to this issue.

Both 2-band and 8-band models are compared in Figure 4-4. The 8-band model gives a slightly smaller value of the complex dielectric constant due to a slightly smaller effective mass as shown in Figure 4-3, but overall they are similar.

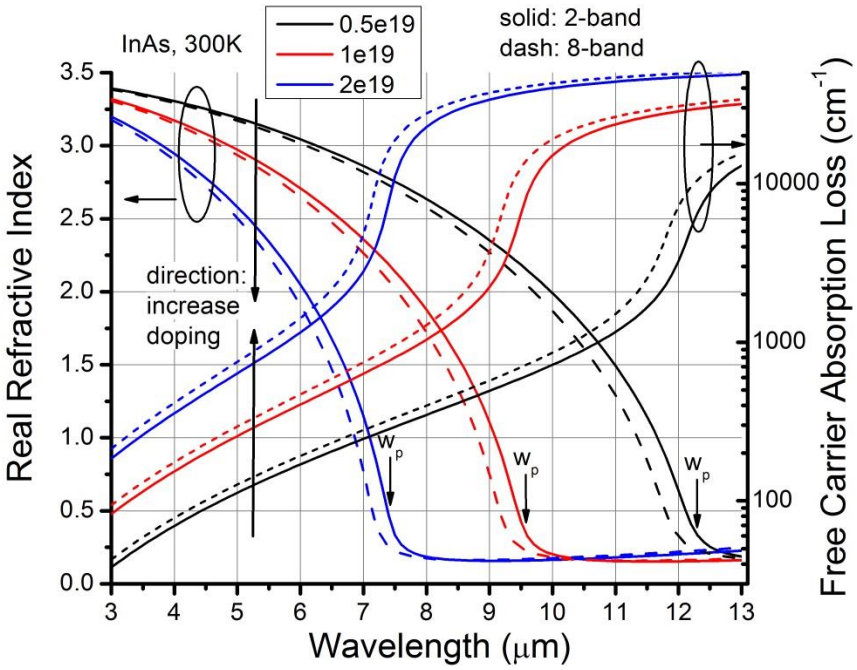


Figure 4-4. Complex refractive index of InAs

Note that the parameters used in the calculation should be calibrated with experimental data. The Hall measurement is routinely used to calibrate the carrier concentration in the growth. Reflectance measurement on the plasmon edge was used to extract the plasmon frequency as well as the damping constant [139]. A more precise and efficient way is based on the Berreman effect, in which the reflectance minimum is polarization dependent [145].

For further insight into the free carrier absorption loss, equation (4.1.4) is decomposed into the real part and imaginary part.

$$\mathrm{Re}(\varepsilon_r) = \varepsilon_\infty(1 - \frac{\omega_p^2}{\omega^2 + 1/\tau^2}) \approx \varepsilon_\infty(1 - \frac{\omega_p^2}{\omega^2}) \tag{4.1.12}$$

$$\mathrm{Im}(\varepsilon_r) = \varepsilon_\infty \frac{\omega_p^2}{\omega^3\tau + \omega/\tau} \approx \varepsilon_\infty \frac{\omega_p^2}{\omega^3\tau} \tag{4.1.13}$$

The above approximation is based on the fact that the damping time constant for heavily-doped InAs is in the range of 0.1~0.3 ps. They are equivalent to the wavelengths of 90~30 μm, far beyond the mid-wave infrared region that we are concerned with. In addition, the imaginary part of refractive index is much smaller than the real part:

$$\varepsilon_r = (n + ik)^2 = n^2 - k^2 + i2nk \approx n^2 + i2nk \tag{4.1.14}$$

Then the free carrier absorption loss is:

$$\alpha_{fc} = \frac{4\pi k}{\lambda} = \frac{4\pi}{\lambda}\frac{\mathrm{Im}(\varepsilon_r)}{2n} \approx \frac{2\pi}{\lambda n}\varepsilon_\infty \frac{\omega_p^2}{\omega^3\tau} \propto \lambda^2 \tag{4.1.15}$$

This is the well-known empirical relation that the free-carrier absorption loss increases proportional to the square of wavelength.

4.1.3 Device structure and waveguide simulation

The InAs plasmon-enhanced waveguide was first used in QC lasers at 10 μm in 2002 [146]. Their cascade regions are typically 30~40 stages (2~3 μm in thickness) and the doping concentration in the InAs cladding layer was 3×10^{18} cm^{-3} for the lasing wavelength of 10 μm [147,148] or 12 μm [149], and 6×10^{18} cm^{-3} for the lasing wavelength of 6 μm [150]. For QC lasers at 20 μm, the stage number is doubled to 72

and the doping in the cladding layer is reduced to 3×10^{18} cm^{-3} to avoid excessive loss [136].

The InAs-based IC lasers at 5.9 μm were first reported in 2009 [66]. They have 10 cascade stages, and the doping in InAs cladding layer is 6×10^{18} cm^{-3}, the same with [150]. These preliminary lasers were able to work up to 184 K in CW mode [133] and were used for our far-field investigation. From top to bottom, the whole waveguide structure consists of a 35-nm-thick n^+-type InAs top contact layer, a 2.15-μm-thick n-type InAs spacer layer, 0.5-μm-thick cascade stages, a 1.65-μm-thick n-type InAs spacer layer, and a 1.5-μm-thick n^+-type InAs bottom cladding layer. The thin top InAs cladding layer was used to simplify the MBE growth and device fabrication.

The InAs spacer layer, with a background doping concentration of about 1×10^{16} cm^{-3}, is not intentionally doped. The function of the spacer layer is to reduce the penetration of the optical wave into the high-loss InAs n^{++} cladding layer. Note that the terms "spacer layer" and "separate confinement layer (SCL)" have been used interchangeably in the IC laser community, although SCL is usually referred to a layer with a refractive index intermediate between the active core and the cladding layer [151]. This will be further elaborated in section 5.3.1.

Pieces of the wafer were processed into ridge-waveguide lasers with different top Ti/Au contact configurations on the ridges. As illustrated in Figure 4-5, the NR device A, B, and C has a ridge width of 15, 20, and 40 μm, respectively. A 200-nm-thick SiO_2 layer is deposited for electrical insulation from the wire bonding pad, and a Ti/Au (~ 30/170 nm) layer is placed on the top of each device for contacts and the wire

bonding pad. Device A has only one 2-μm-wide contact window on the top of the ridge stripe, which is the commonly adopted configuration for Fabry-Perot cavity diode lasers. Devices B and C have two contact windows near the two edges of the ridge with a spacing of 8 and 28 μm, respectively. This double-contact window configuration was used in the past for DFB IC lasers to minimize losses from the top metal contact [58]. Device D (not shown in Figure 4-5), is a broad-area laser, with a mesa width of 150 μm and a 100 μm-wide Ti/Au contact on the top of the center part of the mesa. The cavity lengths of the devices are around 1.8 mm.

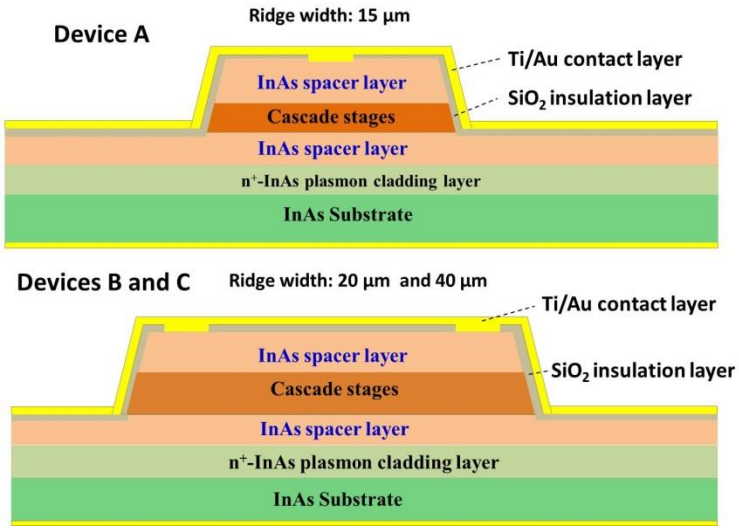


Figure 4-5. Device structure for narrow-ridge IC laser. Device A has a one metal contact trench in the center; Devices B and C have two metal contact trenches near the edges.

The far-field measurements were taken for λ~5.6 μm lasers at 80 K. For the simulation of the optical wave profile in the waveguide at low temperature, the refractive indices are modified by the empirical temperature coefficients [152]. The complex refractive index of Au, Ti and SiO_2, taken directly from the Palik's handbook [137], are 4.44-30i, 2-9.92i and 1.212-2.27e-3i, respectively. For expediting the

calculation, an equivalent refractive index n_{eq} is used to represent the cascade core, by taking a weighted average of the dielectric constant of the constituent materials—InAs, AlSb, GaSb, and InGaSb. It's found that changes in the weighting of the materials have little appreciable effect on the value of n_{eq}.

Figure 4-6 shows one-dimensional optical wave and refractive index profile for the waveguide in the vertical direction. The insulation layer and top contact layer are also included to simulate the actual situation. The equivalent index for the cascade core at 80 K is 3.411, slightly smaller than the InAs spacer layer of 3.413. The free carrier absorption loss is calculated to be 1.5 cm^{-1}, which only counts from the doped InAs layer. The confinement factor of the cascade region is 16.7 %. As we see, the optical wave is pushed towards to the substrate direction due to the small dielectric constant of SiO_2 and the negative dielectric constant of metal.

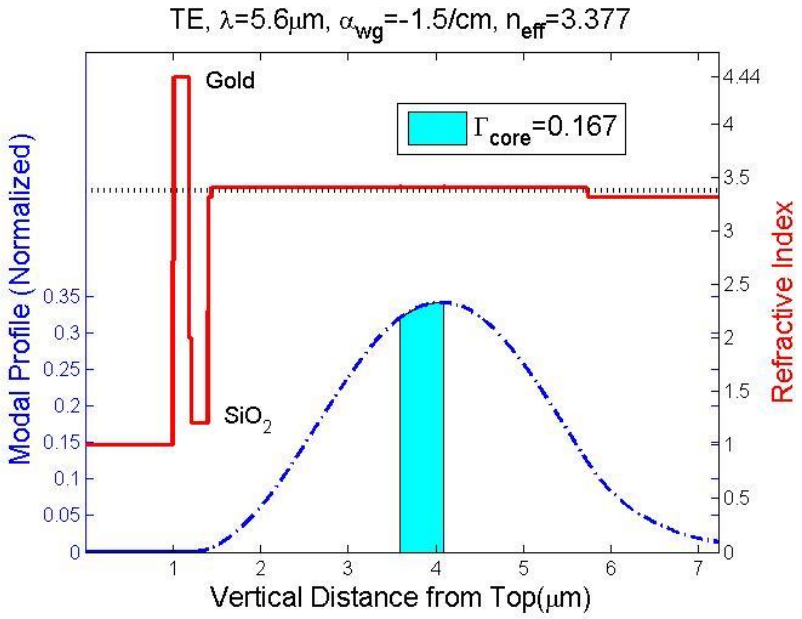


Figure 4-6. One-dimensional optical wave and refractive index profile

4.2 Far-field pattern measurement

4.2.1 Experimental setup

The lasers were operated in pulsed mode, driven by a pulsed current source with a pulse width of 1 μs and a repetition rate of 500 Hz. The emitting light from one facet of the laser was collected by mercury cadmium telluride (MCT) detector, which has an active area of 1 mm × 1 mm. A lock-in amplifier triggered by the pulsed current source was used to display the pulsed signal received in the MCT detector. Under high injection conditions, an IR neutral density filter was placed right in front of the MCT detector to prevent the detector from saturation.

Figure 4-7 shows two experimental setups for the far-field measurement, which can be called translation method and rotation method, respectively. In both cases, the laser and its cooling system *(e.g.*, a cryostat) were set on a three-dimensional (3D) translational stage.

For the translation method as shown in Figure 4-7(a), the MCT detector was placed on a three-dimensional (3D) translational stage 19 mm away from the laser facet. The far-field profile P(x, y, z_0) was obtained by moving either the detector or the laser in the z = z_0 plane. Thus, a two-dimensional far-field pattern is readily available, To facilitate the further analysis, it's necessary to convert the Cartesian coordinate profile into a spherical coordinate profile $P(r, \varphi, \theta)$. The spherical coordinate profile is more convenient and can be obtained by:

$$P(r,\varphi,\theta)\big|_{r=r_0,\varphi=\varphi_0} = P(x,y,z_0)\frac{r^2}{z_0^2\cos\theta} = P(x,y,z_0)\frac{r^2}{z_0^3} \qquad (4.2.1)$$

where $r = \sqrt{x^2 + y^2 + z_0^2}$ .The ratio r^2/z_0^2 is based on the assumption that the spherical surface $4\pi r^2$ and $4\pi z_0^2$ should receive the same amount of light. The factor $\cos\theta$ is due to the difference of effective light flux between oblique incidence and normal incidence.

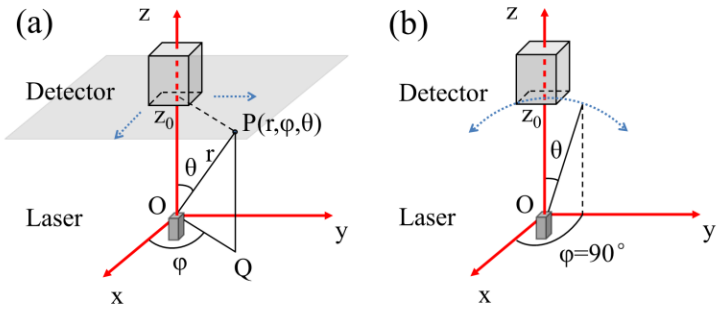


Figure 4-7. Two experimental setups for far-field measurements: (a) translation method. (b) rotation method.

For the rotation method as shown in Figure 4-7(b), the MCT detector was placed on the arm of a rotation stage right below the laser. In this approach, the angle-dependent profile $P(\theta)$ is readily available. In addition, the laser light is always normal incident on the active area of the detector, circumventing the problem caused by the finite field of view of the detector. This is a significant advantage over the translation method. In the latter method, the highly divergent laser beam along the vertical (growth) direction (about ±60°) will not be fully "seen" by the detector with a field of view of ±33°.

4.2.2 Experimental and simulated results

Figure 4-8 (a) shows the 2D far-field pattern of device A operating at $I=1.1*I_{th}$, which is collected by the translation method. A single-lobe profile was observed along the vertical direction while the double peaks appeared along the lateral direction, which

indicates the dominance of the first-order lateral mode over the fundamental (zero-order) mode. The far-field profile was then converted into the angle-dependence along both directions as shown in Figure 4-8 (b). Along the growth direction, the full width at half maximum (FWHM) is 49.9º, which is not accurate due to the limited field of view for the detector. Along the lateral direction, the FWHM of the far-field is 23.1º for each lobe. This observation suggests that the commonly used center metal contact configuration might introduce more loss to the fundamental mode than the first-order mode, leading to a more divergent beam with two peaks in the far-field pattern, which is undesirable.

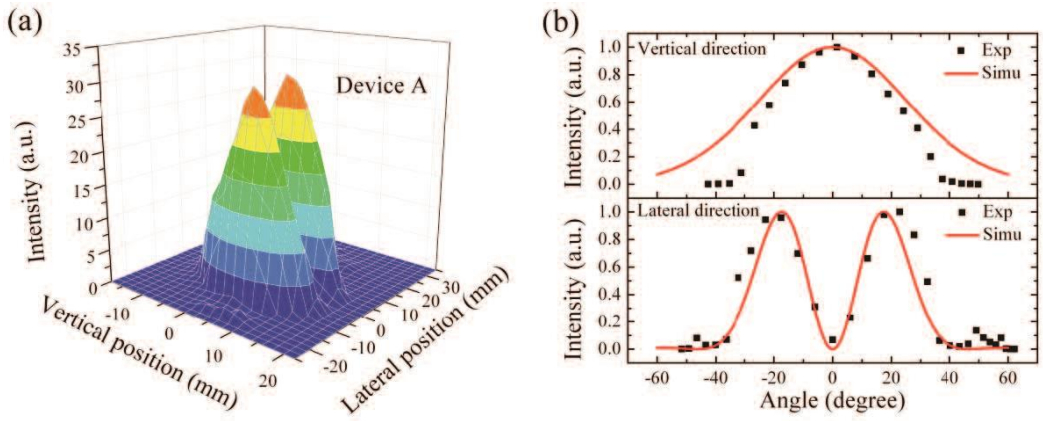


Figure 4-8. (a) Measured 2D far-field profile for device A at 18 mA. (b) Far-field profiles along vertical direction (top panel) and lateral direction (bottom panel).

The vertical far-field profile can be calculated by the conventional 1D slab waveguide model as described in section 2.4. For the lateral far-field profile, a cost-efficient model is the effective index method [153]. This method divides the 2D structure vertically into several sections and uses the modal index of each section as an effective index to form an equivalent slab waveguide along the lateral direction, as shown in Figure 4-9. The simulation indicates that the waveguide loss for the fundamental TE mode is about 13% larger than that for the first-order mode. In

addition, the modal reflectance of the facet for the fundamental TE mode is smaller than the first-order mode. As a result, the first-order lateral mode has a lower threshold gain required for lasing.

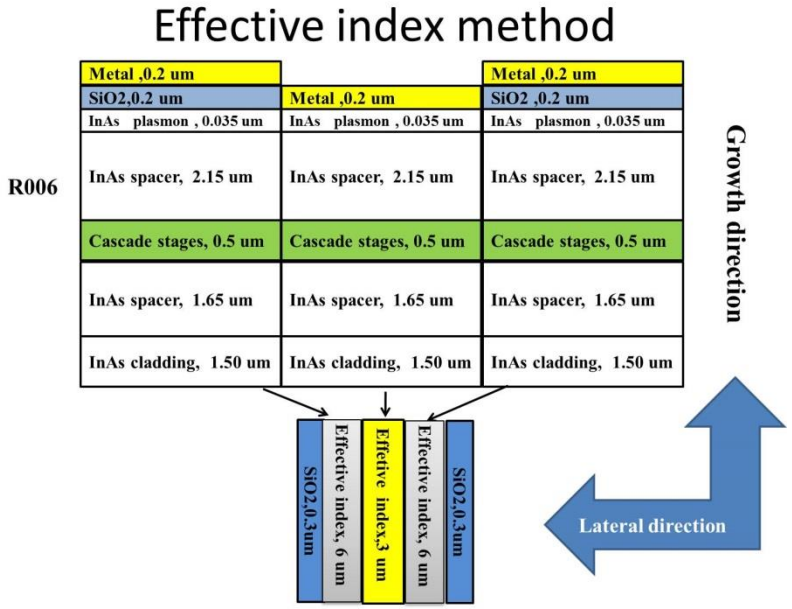


Figure 4-9. Example of the effective index method applied to device A.

The calculated far-field for the first-order lateral mode, as shown in solid lines in Figure 4-8(b), is in good agreement with our experimental result. The calculated FWHM for each lobe is 19.4^{o} , slightly smaller than the experimental result. This may be due to the over-simplification of the effective index method and the possible experimental uncertainties.

The calculated vertical far-field intensity, as shown in the top panel of Figure 4-8 (b), is higher than the measured value when the angle is larger than 30°. This can be explained by the limited field of view of the detector as mentioned earlier. The detector is located inside a dewar about 10 mm away from the dewar window. Based on the geometry of the dewar and the size of the active area, we estimate that a reduced

amount of laser light reaches the detector when the incident angle is larger than 33°, and no light reaches the detector when the angle is greater than 38°. Therefore later measurements on the vertical far-field are done by using a rotation stage.

Figure 4-10 is the lateral far-field profiles of device A under various injection conditions (1.1 to 10 times I_{th}), which show that the valley floor between two lobes rises with the increasing current. This is because the fundamental TE mode also reaches the threshold gain as the injection current increases. The inset to Figure 4-10 gives an illustrative far-field profile that was formed by equal contributions from the fundamental mode and the first-order mode. Again, the sharp drops near 33° of the measured intensity were due to the limited field of view of the detector.

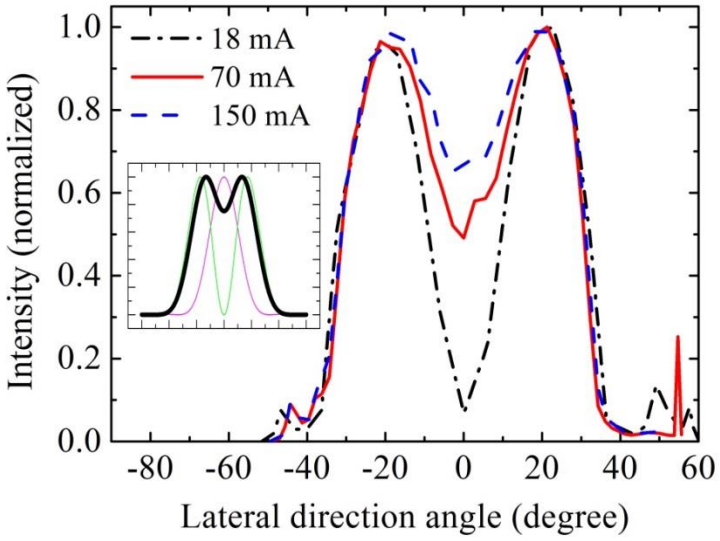


Figure 4-10. Lateral far-field profiles of device A at different currents. Inset: far-field (dark black line) consists of 50% fundamental mode and 50% first mode

The measured far-field profiles of devices A through D along both the vertical and the lateral directions are plotted in Fig. 4-11. All the devices were driven at current levels of around $1.1\times I_{th}$ at 80 K. All the vertical far-field profiles obtained by the

rotation stage have an FWHM of about 55°, which is in good agreement with the calculated value of ~59° for the fundamental vertical mode. These vertical far-field patterns were examined under higher currents, but no appreciable change was found.

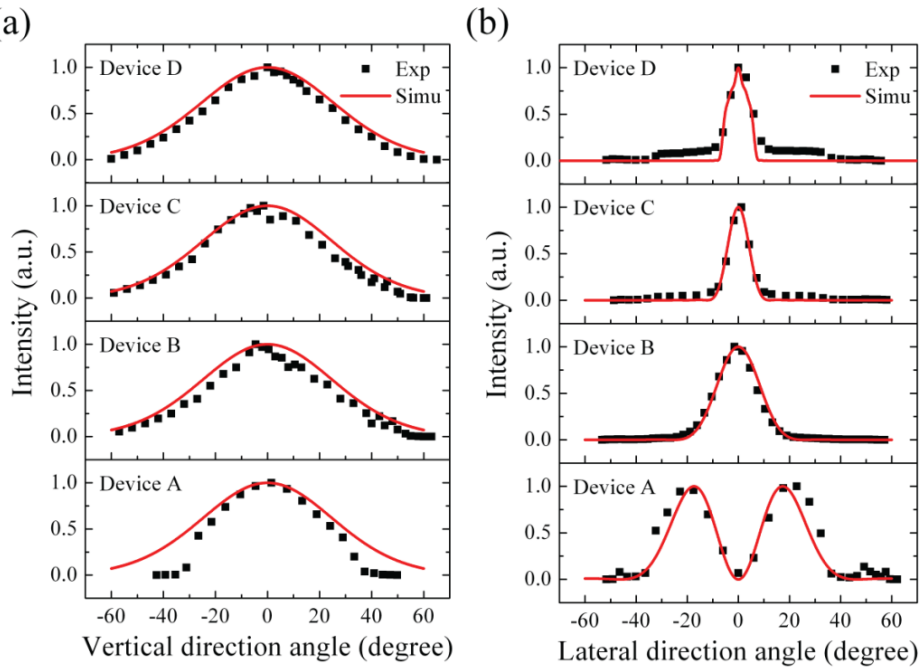


Figure 4-11. Vertical and lateral far-field profiles of four devices

Unlike device A, the lateral far-field profiles of devices B and C exhibit a single-lobe profile. The two edge metal contacts on the ridge lead to a single-lobe lateral far-field profile, corresponding to the fundamental lateral mode. In addition, the larger spacing between the two metal contacts made a less divergent far-field pattern as indicated by the smaller FWHM (~9.2^{o}) for the 40-μm-wide ridge device C compared to the FWHM value (~17.3^{o}) for the 20-μm-wide ridge device B. This is in an excellent agreement with the simulation results based on the fundamental lateral mode. However, for the 150-μm-wide device D, the observed lateral far-field is single-peak with an FWHM of 10.1^{o}, which is a combination of many higher-order lateral modes with

significant divergence, due to their similar thresholds. A superposition of multiple modes (0~4_{th} order modes TE_{00}:TE_{01}:TE_{02}:TE_{03}:TE_{04}=4:4:3:3:2) was used to achieve a reasonable agreement with the experimental curve as shown in Figure 4-11(b).

In conclusion, far-field patterns of plasmon-enhanced waveguide IC lasers have been investigated along the vertical and the lateral directions. It has been observed that the commonly adopted center metal contact configuration leads to an undesirable double-lobe far-field pattern associated with the first-order lateral mode in plasmon-enhanced waveguide IC lasers. With two metal contacts placed near the edges of the ridge, a more desirable, less divergent single-lobed beam was exhibited in the far-field. The observed far-field profiles are in good agreement with simulation results. This study suggests that appropriate metal contact configurations need to be considered in the design of plasmon-enhanced IC lasers, especially when a thick top semiconductor cladding layer is absent.

4.3 Beam quality

In fact, the measured far-field profile is by itself not the most informative figure of merit [154]. Because lasers with different ridge widths can in principle have the same angular divergence, due to the combinations of higher lateral modes. In this section, the Gaussian beam is first introduced as a starting point. Then the concept of brightness and its importance is discussed in detail. At last, the beam quality factors of IC lasers are calculated.

4.3.1 Gaussian beam

Gaussian modes are the mathematical solutions to the Helmholtz equation in the paraxial approximation. The lowest order mode TEM_{00} is the fundamental Gaussian beam, or simply called the **Gaussian beam**. If the Gaussian beam is propagating along the +z direction, as shown in Figure 4-12, the amplitude of its electric field is given by:

$$E(x,y,z) = E_0 \frac{w_0}{w(z)} \exp^{-[\frac{r}{w(z)}]^2} \tag{4.3.1}$$

where w(z) is the beam width, w_0=w(0) is the waist size at the beam's focus, z is the axial distance from the beam's focus, $r = \sqrt{x^2 + y^2}$ is the radial distance from the center axis of the beam. From equation (4.3.1), we can see the beam width is where the amplitude decays by 1/e, or the intensity drops to $1/e^2$ (≈ 13.5%).

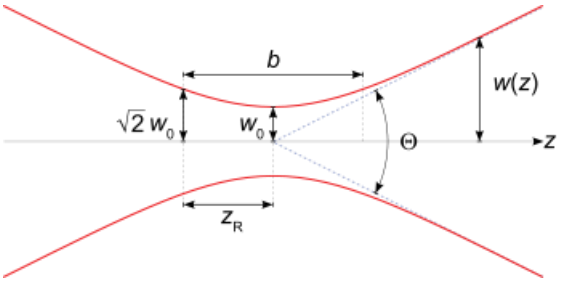


Figure 4-12. Spread of the Gaussian beam [155]

The beam width w(z) is z-dependent, given by

$$w^2 = w_0^2[1 + (z / z_R)^2] \tag{4.3.2}$$

$$z_R = \frac{\pi w_0^2}{\lambda_0} \tag{4.3.3}$$

where z_R is called the Rayleigh range. At the wave plane of z=0, w(0)=w_0 is the smallest beam width, called beam waist. At the wave plane of z=z_R, the beam width w(z_R)=$w_0/\sqrt{2}$, and the amplitude E(x, y, z_R)= E(x, y, 0) /$\sqrt{2}$.

At the far field condition or Fraunhofer diffraction ($z>>z_R$), equation (4.3.2) becomes:

$$w = w_0 z / z_R \tag{4.3.4}$$

The beam width increases linearly with z. This means that far from the waist, the beam width is cone-shaped, with a divergence half angle by

$$\frac{\Theta}{2} = \text{atan}(\frac{w_0}{z_R}) = \text{atan}(\frac{\lambda_0}{\pi w_0}) \tag{4.3.5}$$

This inverse relationship between the beam waist and the far-field divergence is a fundamental characteristic of diffraction. For the paraxial approximation ($w_0>2\lambda/\pi$), the divergence half angle is simply $\lambda/(\pi\, w_0)$.

The laser beam quality is quantified by the **beam parameter product (BPP)**, a product of the beam's divergence and the beam waist. This value is λ/π for an ideal Gaussian beam.

The **beam quality factor**, also known as $\mathbf{M^2}$ (M-squared), is calculated by the ratio of the BPP of a real beam to that of a Gaussian beam at the same wavelength. The M^2 for a Gaussian beam is one; while all actual laser beams have M^2 values greater than one. A laser beam with M^2 close to one is said to have the diffraction-limited quality.

Note that the beam divergence is also measured by the full width at half maximum (FWHM). The conversion between it and the beam width is given by:

$$e^{-2(\frac{FWHM/2}{w})^2} = 0.5 \rightarrow w = 0.85 \cdot FWHM \tag{4.3.6}$$

4.3.2 Brightness

Figure 4-13(a) shows a monochromatic beam of uniform intensity and phase wavefront is incident on a screen S containing an aperture D. The beam will have a finite divergence θ_d due to diffraction:

$$\theta = \beta \frac{\lambda}{D} \tag{4.3.7}$$

where the factor β is a numerical coefficient of the order of unity, whose value depends on the shape of the amplitude distribution and on the way in which both the divergence and the beam diameter are defined. For Fraunhofer diffraction, the factor β is one. For a Gaussian beam with a circular cross-section of diameter D, the factor β is about $2/\pi$.

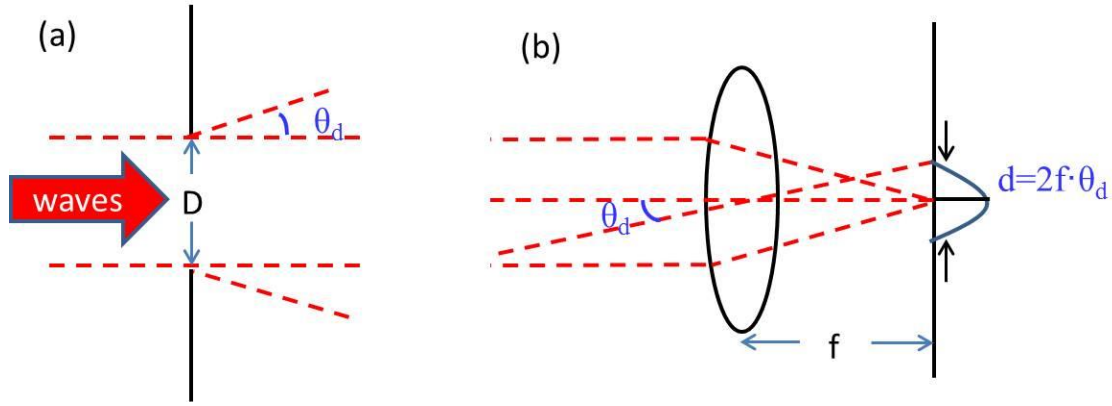


Figure 4-13. (a) Divergence of a plane wave due to diffraction ;(b) Intensity distribution in the focal plane of a lens for a divergent beam.

The **brightness** of a given source of electromagnetic waves is defined as the power emitted per unit surface area per unit solid angle [156]:

$$dP = B\cos\theta dS d\Omega \tag{4.3.8}$$

where B [W/(cm^{2}sr)] is the brightness. For a small divergence angle θ, cosθ is close to 1.

For an arbitrary beam with a circular cross-section of diameter D and with a divergence θ_d, the surface area is $\pi D^2/4$, and the emission solid angle is $\pi(r\theta)^2/r^2 = \pi\theta^2$. Equation (4.3.8) becomes:

$$P = B\pi^2 D^2 \theta_d^2 / 4 \tag{4.3.9}$$

For a Gaussian beam, $\theta_d = \frac{2}{\pi}\frac{\lambda}{D}$, so the brightness is P/λ^2.

The importance of brightness is better understood if a lens is used to focus the laser beams, as shown in Figure 4-13(b). The beam divergence θ_d of the waves will make the focal spot has a finite diameter $d=2f\theta_d$. For an ideal lens, the overall power in the focal plane equals to the power of the incoming wave. The peak intensity in the focal plane is:

$$I_p = \frac{4P}{\pi d^2} = \frac{P}{\pi f^2 \theta_d^2} = \frac{\pi B D^2}{4 f^2} \xrightarrow{D_L = D} \pi B (N.A.)^2 \tag{4.3.10}$$

where $N.A. = n\sin\theta \cong \frac{D_L}{2f}$ is the lens numerical aperture. So the peak intensity in the focal plane of a lens depends only on the beam brightness.

Therefore, a high-brightness laser requires a high output power, a diffraction-limited beam quality and a matched lens fabricated at its facet.

4.3.3 Beam quality factors for IC lasers

For a practical laser, the beam quality depends on whether the laser works at its fundamental transverse spatial mode. In the growth direction, the thickness of the active QWs is much smaller than the lasing wavelength in the median. According to equation (2.4.25), it's easy to obtain single-mode guidance and thus an essentially diffraction-

limited beam quality with an M^2 factor only slightly above 1. Anyway, the small aperture size leads to a "fast" divergence. Hence, this vertical direction is also called "fast axis."

In the lateral direction, the stripe width is much larger which allows many spatial modes coexist. As a result, the optical fields of higher-order modes superposition on the fundamental mode, leading to an overall divergence much larger than that of a diffraction-limited beam corresponding to the stripe width (close to the beam waist). Although the beam quality is worse in this direction, the larger size leads to a relatively smaller divergence. So the lateral direction is also called "slow axis."

A rigorous calculation of the beam quality factor is to get the beam variance along transverse coordinates, and spatial-frequency coordinates by integral [156]. Another equivalent but more convenient way is to compare the beam divergence of a practical laser to that of a Gaussian beam with the same beam waist. In this way, the calculated near-field for the practical laser is used to obtain the beam waist, the value of which is assigned to the Gaussian beam for evaluating the diffraction-limit divergence angle.

The calculated near-field for R006 in the growth direction, as shown in Figure 4-6, has a beam waist w_0=2.26 μm and an estimated divergence half angle of 45.2°. The experimental value for the $1/e^2$ divergence half angle, extracted from Figure 4-11(a), is 47.5±2°. This means the M^2 factor is close to one and verifies that the fast axis has a diffraction-limited beam quality.

In the lateral direction, due to the lack of effective waveguide, the near-field intensity distribution is approximately constant. Consequently, the Gaussian full waist

can be taken to be the stripe width [154]. According to equation (4.3.5), the Gaussian divergence half angles for the ridge widths of 20 μm, 40 μm, and 150 μm are 5.11°, 2.55°, and 0.68°, respectively. The corresponding experimental values for the $1/e^2$ divergence half angle, extracted from Figure 4-11(b), are 15.5±2°, 8.0±2° and 10.0±2°. So the corresponding M^2 factors are 3.0, 3.1 and 14.7, respectively.

Figure 4-14 shows the beam quality factors for devices from four other representative wafers. In Figure 4-14 (a), the lateral far-field was fit by the Gaussian function and an excellent fitting was achieved with a half beam width of 8.3°. By comparing it with the diffraction angle calculated by equation (4.3.5), an M^2 factor of 3 was obtained for this 30-μm-wide device. In Figure 4-14(b), we can see the beam quality gradually degrades as the current increases. Figure 4-14(c) shows that single frequency mode broad-area DFB lasers [157] can have multiple spatial modes. Due to the large size of the ridge, the M^2 factor is as large as 25, undesirable for practical use. Figure 4-14(d) shows several narrow-ridge lasers with the capabilities of CW room temperature operation. Even the ridge is narrow down to 10 μm; the M^2 factor doesn't improve significantly. More advanced fabrication techniques, such as corrugated sidewalls, can improve the beam quality by increasing the loss of higher-order lateral modes [158].

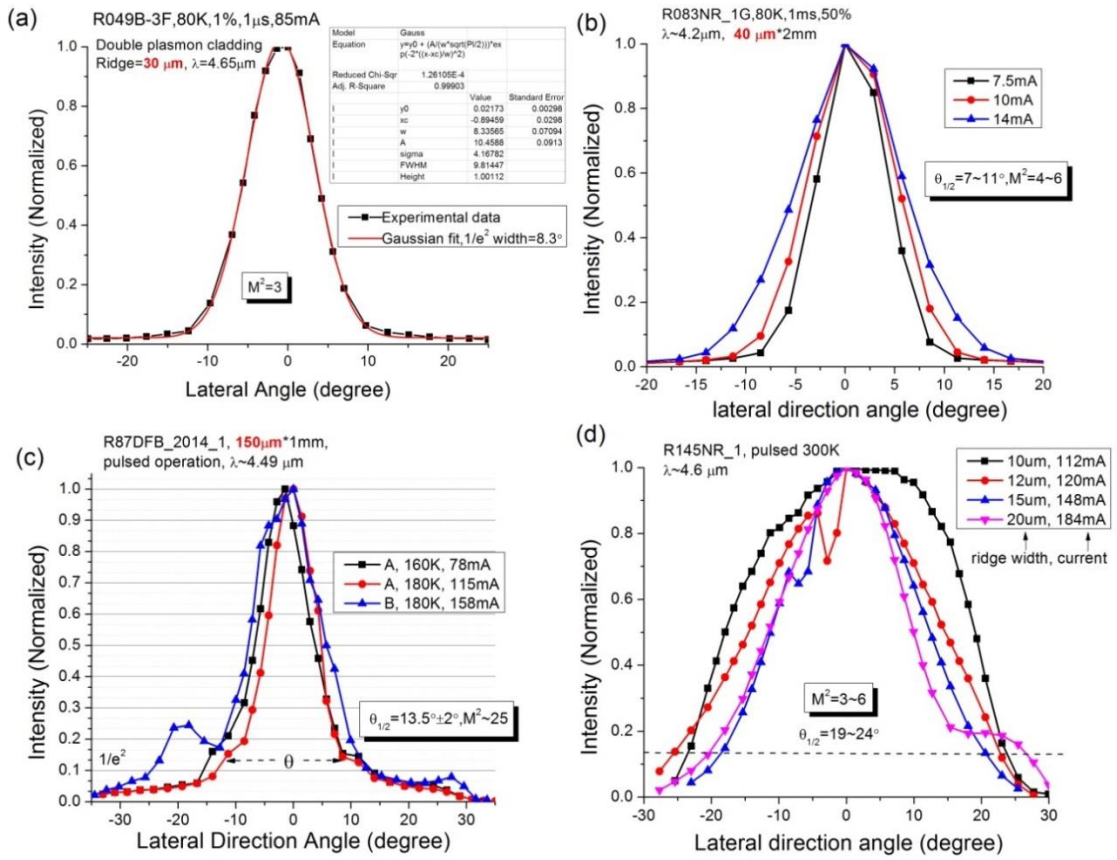


Figure 4-14. Beam quality factors for devices from four representative wafers

Actually, the beam quality is more important for high power lasers, in which the increased active gain area naturally introduces more spatial modes. Strategies such as angle-tilt DFB lasers [159] and tapered ridge [160], were developed to obtain M^2 of about 2. The "effective brightness", defined as output power divided by M^2, had been achieved as high as 175 mW [160] in CW operation at room temperature.

Chapter 5 InAs-based interband cascade lasers above room temperature

With a series of design improvements as well as the consistent material growth and device fabrication, the threshold current densities of InAs-based IC lasers are sharply reduced to about 250 A/cm^2. This chapter reviews the three key steps during this development, which increase the pulse operating temperatures to 315 K, 335 K, and 375 K, respectively. And the corresponding maximum operating temperatures in CW mode are increased to ~250 K, ~280 K and ~310 K. The last one, CW operation above room temperature, marks a significant milestone for the InAs-based IC lasers.

5.1 InAs-based IC lasers with asymmetric cladding layers

5.1.1 Device structure and waveguide modeling

The InAs-based IC lasers in the preliminary exploration, as mentioned in section 4.1.3, have asymmetric cladding layers [161]. The top cladding layer is composed of 35-nm-thick n^+-type InAs, 200-nm-thick SiO_2, and a Ti/Au contact layer. This practice simplifies the material growth and reduces the thermal resistance, but whether the device performance is improved or not depends on both material quality and waveguide design.

A pair of waveguides is intended to compare the symmetric and asymmetric cladding layer. For a fair comparison, they both have identical cascade core, each of which consists of an electron injection region (6 InAs/AlSb QWs), a "W" shape type-II active region and a hole injection region (2 GaSb/AlSb QWs). The "shortened injector"

is implemented in the electron injection region with the first InAs QW of 6.1 nm. The n-type doping concentration in the electron injection region is ~1.5×10^{17} cm^{-3}, without using the "carrier rebalancing" strategy.

Figure 5-1 shows the profiles of optical mode and refractive index along the growth direction for two waveguides. The thicknesses of the spacer layers in both waveguides are optimized individually to achieve smallest threshold gain. R49 is the reference structure, in which the upper waveguide is the mirror image of the bottom waveguide. It's dubbed as "symmetric cladding" for convenience. In contrast, R48, dubbed as "asymmetric cladding", is intentionally designed with a thin top cladding layer. To reduce the penetration of the optical wave into the high-loss metal layer, the top spacer layer of R48 is 300 nm thicker than that of R49. As shown in Figure 5-1, R48 has a slightly smaller confinement factor, but the free carrier absorption loss is ~20% smaller. This is because, at λ=5.3 μm, the complex refractive index of SiO_2 is 1.39-0.0023i, while the refractive index is 2.90-0.012i for the n^{++} InAs with a doping concentration of 1×10^{19} cm^{-3}. So the SiO_2 insulation layer serves as a better cladding material due to a larger index contrast and a smaller optical loss.

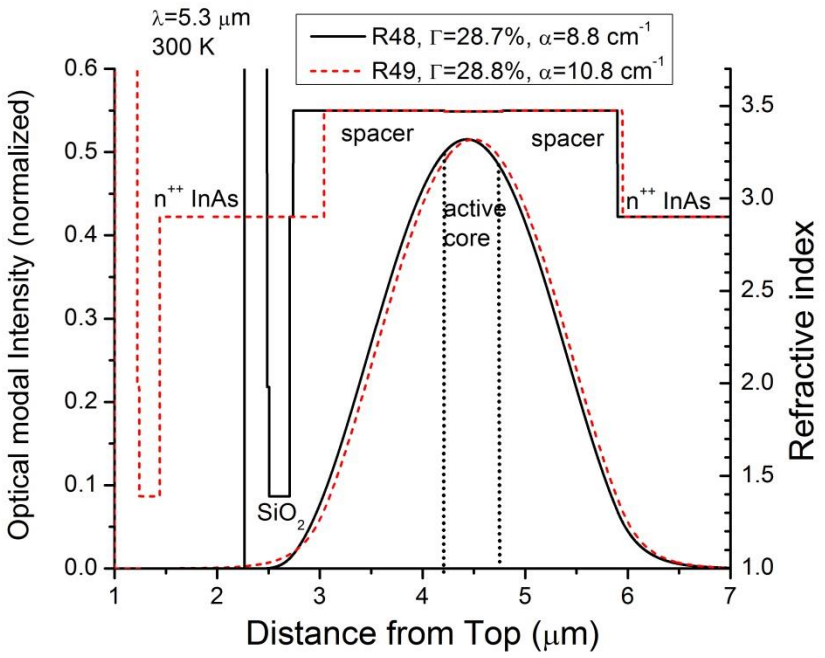


Figure 5-1. Optical mode and refractive index profiles along the growth direction

The above calculation is based on a simplified 1D model along the growth direction. The optical absorption loss due to intersubband transition and excess non-equilibrium carriers in the cascade region is not included. In the lateral direction, the top metal contact configuration, similar to Figure 4-5, has a small impact on the threshold gain of the asymmetric waveguide R48, because the metal contact directly touching the semiconductor is narrow (~3 μm). In contrast, there is no such impact for the symmetric waveguide, because the optical wave is well confined within the n^{++} InAs layer.

5.1.2 Experimental results

Two sets of IC laser structures, comprised of 10 and 12 cascade stages, were grown in a Gen II MBE system on n-type InAs (001) substrates. The first set of 10-stage wafers, R48 and R49, have the waveguide structures as shown in Figure 5-1. The

second set of 12-stage wafers, R65 and R66, have the same spacer and cladding layer thickness with R48 and R49, respectively. The cascade region is designed and grown with Al(Ga)As interfaces to achieve strain-balance.

After growth, the wafers were processed into narrow-ridge lasers, with ridge widths ranging from 15 to 40 μm. The processed wafers were cleaved into laser bars with cavity lengths of 1-3 mm with facets left uncoated. They were mounted epi-side-up on copper heat-sinks for measurements. The testing results show that these lasers with either symmetric cladding layers or asymmetric cladding layer can work in CW mode around 250 K, and in pulsed mode above room temperature near 5.3 μm. Their temperature dependent spectra and light-current curves have been reported in [69,162].

It's still not conclusive whether the "asymmetric waveguide" has comparable performance with the "symmetric waveguide", due to non-uniformities in the wafers (*e.g.*, different defect densities), processing variations, and different laser sizes. Figure 5-2 (a) shows the temperature dependent threshold current densities for lasers with the various sizes from the four wafers. For the 10-stage set, R48 lasers have larger threshold current densities than that of R49 when the same size is picked. However, lasers from both wafers have very similar maximum pulsed operating temperature and power slope efficiency, as shown in Figure 5-2 (b), so neither waveguide has a dominating advantage over the other. A better conclusion can be drawn if the broad-area lasers are fabricated.

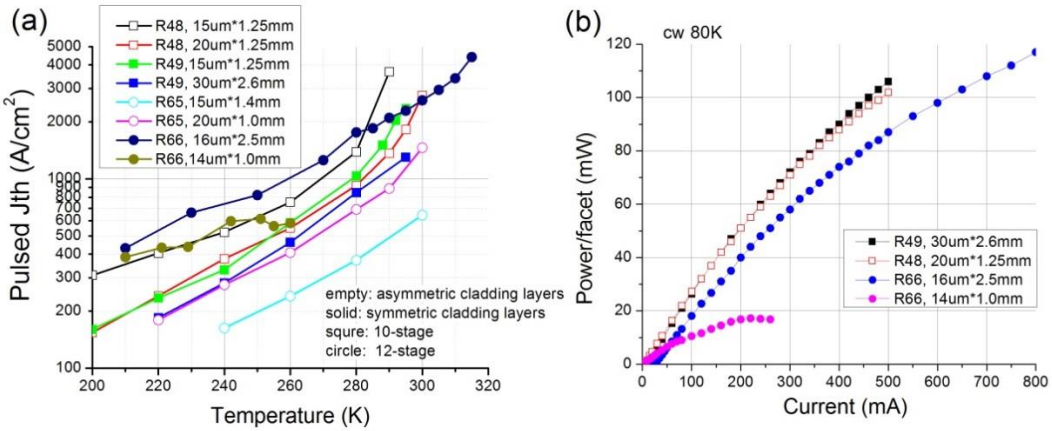


Figure 5-2. (a) Temperature dependent threshold current densities in pulsed mode for lasers from 4 wafers; (b) Power-current curves for lasers from 3 wafers.

For the 12-stage set, it seems that R65 lasers with the asymmetric waveguide have a much smaller threshold than that of R66 lasers. But one laser from R66 achieved the highest pulsed operating temperature (315 K) among them. This means that the intrinsic structure of R66 could be better. The high threshold current densities may be due to the defect-related Shockley-Read-Hall recombination or other leakage channel introduced by imperfect passivation during the processing. Besides, another laser from R66, with the size of 14 μm*1.0 mm, exhibited a non-monotonic threshold current at temperatures between 220~260K. Furthermore, it had a much smaller slope efficiency than its "normal" peers. Interestingly, a vast tunable range of 280 cm^{-1} was found from this laser. The details of this tunable laser were reported in [163].

In conclusion, although the numerical model predicts a lower loss for the asymmetric waveguide, the further experiment is required to confirm this finding. If it is true, the practical benefits include: (1) the MBE growth of the IC laser structure is simplified with a reduced growth time; and (2) with a very thin top cladding layer, a top DFB grating can be simply integrated with an IC laser for strong coupling.

5.2 InAs-based IC lasers with heavily-doped injector

5.2.1 Carrier rebalancing

In the conventional ICL design, the doping concentrations of the InAs QWs in the electron injector is doped moderately with Si to 4×10^{17} cm^{-3}. In 2011, Naval Research Lab reported a bold strategy named "carrier rebalancing" [51]. They increased the doping to 5×10^{18} cm^{-3}, which reduced the pulsed threshold current density for a 5-stage IC laser by half, to 167 A/cm^2. Because the injector states adjacent to the active region are $5*k_BT$ higher than the active region [164], the injected electrons reside mostly in the injector. On the other side, nearly all the injected holes transfer to the active GaInSb QW. They thought the active hole population substantially outnumbers the electron population, which could be compensated by n-doping the electron injector QWs much more heavily.

However, the carrier density "rebalancing" between electron and hole seems to be more a fitting result by playing with the parameters such as Auger coefficient and the normalization thickness [51], the values of which were later corrected by an order of magnitude [25]. Furthermore, the direct observation by scanning voltage microscopy (SVM) measurements [165] showed there was no net charge distribution in the electron injector, in either scenario when the InAs QWs were heavily-doped (as in wafer R84, R90), or moderately doped (as in wafer R94). The charge carriers mostly accumulated in the active region and were clamped with a density of 7×10^{11} cm^{-2} above the threshold at 78 K. In addition, a carrier leakage was found in the valence band of first AlSb layer from the active region to the electron injector, which was still not understood.

Anyway, heavily doping in the vicinity of the active region is not entirely new in the long history of laser development. In 1988, a theoretical study proposed that n-type modulation doping of InGaAs/InP QW active layers could reduce the transparency level, which is applied in other material systems having s strong conduction band/valence band effective mass asymmetry[77]. The doping concentration of 3×10^{18} cm^{-3} in the SCL layer of a GaAs/AlGaAs single QW laser was able to bring down the threshold current by 30% [166].

I tentatively propose a new explanation based on the origin of material gain. Figure 5-3 shows the occupation probability as a function of carrier injection in InAs/GaSb QWs, where the electron effective mass is from InAs and the heavy hole effective mass is from GaSb. The Bernard-Duraffourg inversion condition, as mentioned in section 2.2.1, requires that the occupation probability of an electron at the conduction band edge is larger than that at the valence band edge, i.e., $f_c>f_v$. The difference of occupation probability is plotted as the blue line in Figure 5-3. Due to the small effective mass of electron in the conduction band, its probability has a sharp increase during the initial injection. A high doping concentration, *i.e.*, $n\sim N_c$, will significantly elevate f_c and reduce the required injection to reach the transparency condition. This head start helps achieve a smaller threshold carrier density and hence the lower threshold current density and lower internal loss.

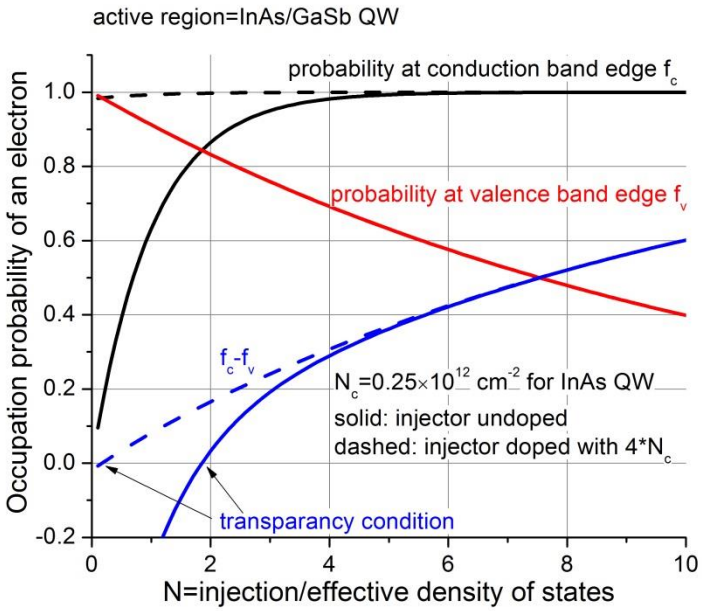


Figure 5-3. Carrier density dependent occupation probability for active QWs with injectors undoped and heavily doped

With the substitution of momentum matrix and density of states for QW, the gain coefficient, as in equation (2.2.14), can be explicitly written as:

$$g(\hbar\omega) = \frac{q_e^2}{4\bar{n}_b c_0 \varepsilon_0 \hbar d_a} \frac{E_p}{E_g} \frac{m_r}{m_0} (f_{c0} - f_{v0}) \xrightarrow[d_a=3nm]{InAs/GaSb} 2800 cm^{-1} * (f_{c0} - f_{v0}) \quad (5.2.1)$$

where d_a is the normalized thickness of active region, f_{c0} and f_{v0} are the electron and hole occupation probability at the respective band edge. In this simplified formula, the gain coefficient is proportional to the inversion probability.

Figure 5-4 shows the differential inversion probability as a function of the normalized injected carrier. A sharp decrease in the differential gain is due to the fact that the occupation probability f_{c0} quickly approaches unity and is nearly saturated. The decrease is even more severe for the high-doped QW, which can be fit into an **exponential decay** of 0.1*exp(-0.1*N). This prompts us that the inversion probability is:

$$f_{c0} - f_{v0} = 1 - e^{-N/10} \approx \frac{N}{10} - \frac{N^2}{200} \tag{5.2.2}$$

where N is the injected carrier density normalized by the effective density of states. This equation means that the commonly assumed **linear** gain-injection relation is only valid for small injection.

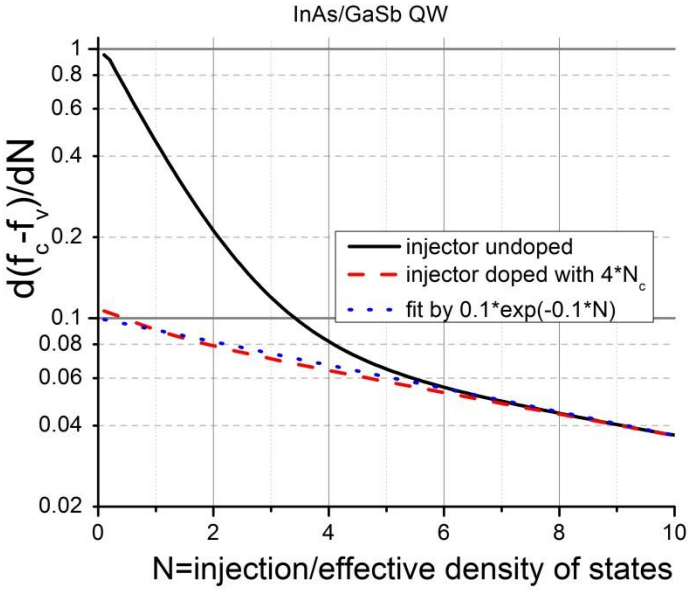


Figure 5-4. Normalized differential gain in InAs/GaSb QWs with injectors undoped and heavily doped

Figure 5-3 and Figure 5-4 offer some profound insights into the fundamental limit of any semiconductor laser based on the interband transition. First, the differential gain is rapidly decreasing at high injection. It's always the top strategy to make the device work at a low injection. For a high power laser, a low threshold carrier density means a low internal loss, which can help improve the power slope efficiency. Besides, if the required threshold gain for a laser is too high, a large injection may still not reach the requirement because the increment of gain will be marginal.

Second, at a higher temperature, the superlinear threshold current density is expected due to three aspects: (a) to maintain the same transparency condition, the transparency carrier density increases linearly with T because the 2D effective density of states $N_c^{2D} = \frac{m_n^* k_B T}{\pi \hbar^2}$. (b) the optical loss, such as free carrier absorption loss, increases because the thermal energy $k_B T$ can help ionize more carriers that previous are bound to the nucleus. (c) the carrier loss, such as Auger recombination, increases by n^3, which significantly reduces the carrier lifetime. All these factors together contribute to the exponential growth in the threshold current density with a characteristic temperature T_0 much smaller than room temperature. However, a quantitative derivation of the low T_0 is not yet reported.

Third, the differential modal gain can be directly predicted by:

$$\Gamma \frac{dg}{dJ} = \frac{\Gamma \tau}{q} \frac{dg}{dn_{2D}} = \frac{\Gamma \tau g_0}{q N_c^{2D}} \frac{d(f_{c0} - f_{v0})}{dN} \tag{5.2.3}$$

For example, a 10-stage IC laser, as shown in Figure 4-2, has a confinement factor of 25% for the whole cascade core. But strictly speaking, the exact active core composed by InAs/GaInSb/InAs QWs, has a normalized thickness of about 3 nm, which reduces the confinement factor to only 1.5%. At an injection of $4*N_c$, the carrier lifetime is taken to be 6 ns, and the differential inversion probability is 0.067. Thus, the differential modal gain is 0.42 cm/A, or 0.042 cm/A per stage. Note that the injection and internal efficiency are assumed to be 100% and the carrier lifetime will be reduced by other possible nonideal factors. So this ideal value is an upper limit. It is comparable to a recent theoretical value of 0.02 cm/A per stage[25]. For conventional IC lasers with

"carrier rebalancing," experimental values were reported to be 0.06 cm/A at 270 K [120] and 0.011 cm/A at 300 K [122].

Fourth, the optimal n-type doping concentration is about 4 times the effective density of states for conduction band. As shown in Figure 5-3, this level of doping can make the occupation probability f_c close to unity. The transparency condition is soon achieved once the external electrical field sweeps excess carriers into the active region. Because the effective density of states in QW is a 2D value, the conversion to a 3D value is to divide it by the layer thickness. At longer wavelengths, the InAs QWs in the electron injector are usually wider for a smooth carrier transport with proper band alignment. As a result, the bulk doping concentration is generally smaller. In addition, the choice of which QW layer to dope is also critical. Experiments showed that the doping in the vicinity QW near active region introduced more loss than benefit [71], which is still not quantitatively understood.

5.2.2 Band structure and experimental results

To implement the "carrier rebalancing" design, a series of wafers were grown with the injector doping concentrations varying from 1.2×10^{18} cm^{-3} to 6.3×10^{18} cm^{-3}. Generally, lasers from these wafers have a lower threshold and higher pulsed operating temperature than those without the "carrier rebalancing" design. The optimal doping concentration is in excellent agreement with the proposal in section 5.2.1.

Figure 5-5 (a) shows the diagram of one complete cascade stage for an 8-stage laser wafer R083, the waveguide of which is similar to the asymmetric waveguide R48

(shown in Figure 5-1). The middle 3 InAs QWs of the 5 QWs in the electron injection region were heavily doped to 4.3×10^{18} cm^{-3}.

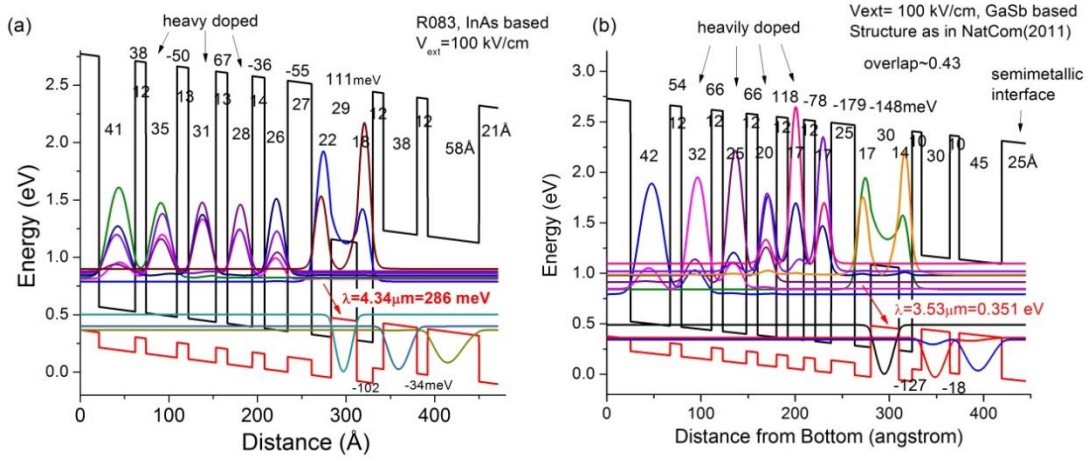


Figure 5-5. Band diagrams of one complete cascade stage and eigenenergy alignments for (a) R083 (b) reference structure in [51]

The fundamental eigenenergy in each InAs QW is calculated by the two-band model under an electric field of 100 kV/cm. The band offsets are corrected by the strain effect based on the InAs substrate. The eigenenergies for heavy holes in the Ga(In)Sb layers are calculated by the one-band model. The wavefunctions are plotted as a superposition over their respective eigenenergies. The number above the AlSb barrier is the energy change along the direction of electron or hole transport. The next two lines of numbers are the thicknesses of each layer.

As a reference, the well-established structure reported in [51] is plotted in Figure 5-5(b) for a side-by-side comparison. Both of them have roughly equal total thickness for one stage, although the reference structure has one more QW in the electron injector. A significant difference is the eigenenergy alignment in the conduction band between the two structures. In R083, the first QW in the active region has the lowest energy. In contrast, in the reference structure, the first QW in the electron injector has the lowest

energy, and the electron energy is elevated all the way until the last QW in the injector. So far, there is no systematic study on the optimal layer arrangement for an efficient carrier transport in the injector or to achieve an effective population inversion in the active region. Instead, the structure like in Figure 5-5(b) is more a collection of many empirical findings that have been patented [167-169].

After the MBE growth of the structure R083, standardized broad-area lasers were fabricated by contact lithography and wet chemical etch, in order to provide rapid feedback for wafer screening and design evaluation. The etch proceeded to the InAs spacer layer below the active region. Major characterizations for a typical device is shown in Figure 5-6. This laser is able to work up to 335 K in pulsed mode and 237 K in CW mode. It has a lasing wavelength of 4.86 μm at 300K, which is longer than the calculated value of 4.34 μm, mainly due to the uncertainty of the band offsets.

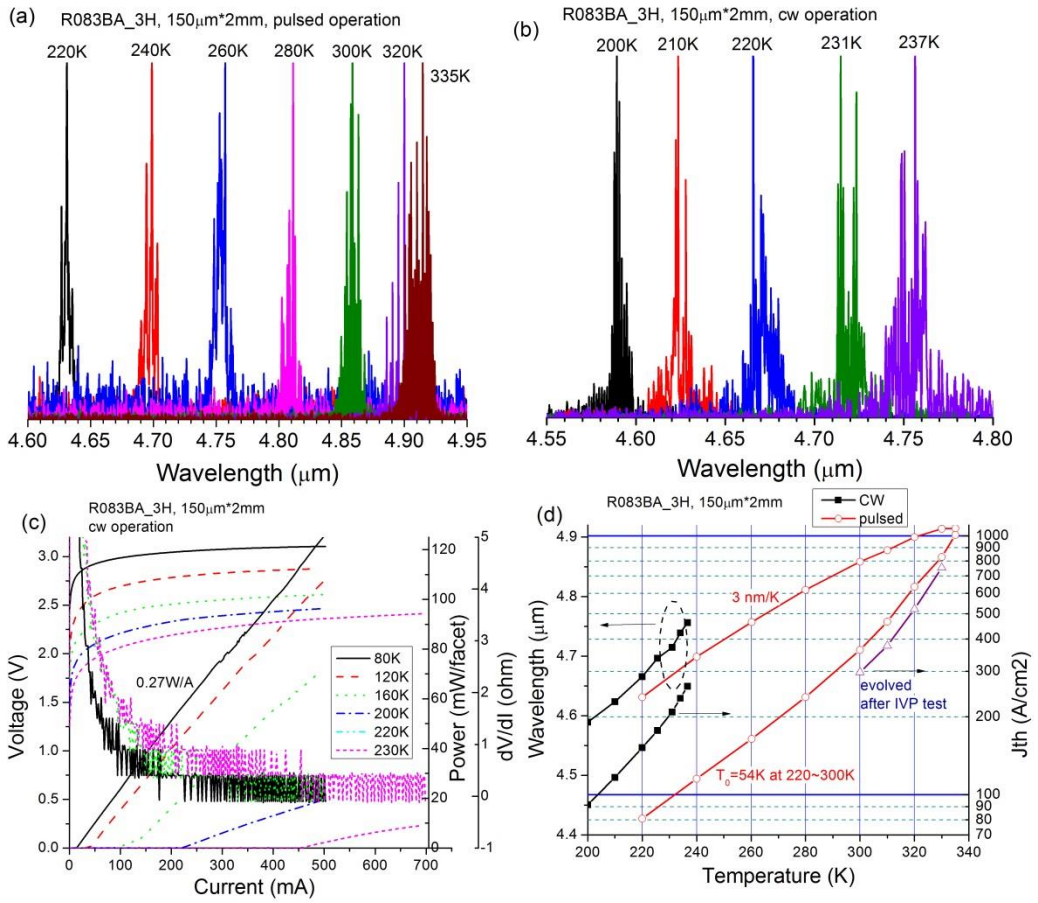


Figure 5-6. Characterizations for device R083BA_3H. (a) pulse spectra; (b)CW spectra; (c) IVP and differential resistance in CW operation; (d) temperature dependence of lasing wavelengths and threshold current densities

The pulsed threshold current density at 300K is the most important benchmark we used to judge the basic quality of a laser. This value was 360 A/cm^2 in the initial test, which however evolved to 300 A/cm^2 after the IVP test, as shown in Figure 5-6(d). This evolution may be due to the non-uniformity of the material and annealing-like effect during the high injection.

Figure 5-6(c) shows the IVP curves of the laser in CW mode. The slope efficiency at 80K is 0.27 W/A per facet, a relatively low value due to the large size of the cavity length. Another laser with a length of 1.2 mm achieved a slope efficiency of ~0.40 W/A and a maximum output power of 250 mW at 700 mA. The differential

resistance extracted by the figure is about 0.2±0.1 Ω at high injection. This translates into a differential resistivity dV/dJ of 0.6 mΩ*cm^2, which is comparable to the reported value of 1 mΩ*cm^2[170].

The low threshold current densities achieved in broad-area lasers suggest the possibility of CW operation at room temperature with narrow-ridge devices. In principle, the narrower ridge is preferred. However, due to the available masks and fabrication facility, the narrowest ridge processed from this wafer was 15 μm, compared with 7.4 μm accessible by another lab [171]. Figure 5-7 shows the CW spectra for a 15 μm-wide laser at several temperatures. The maximum CW operating temperature is 282 K, with a lasing wavelength of 4.85 μm. The corresponding threshold current density and threshold power is 1 kA/cm^2 and 0.9 W, respectively.

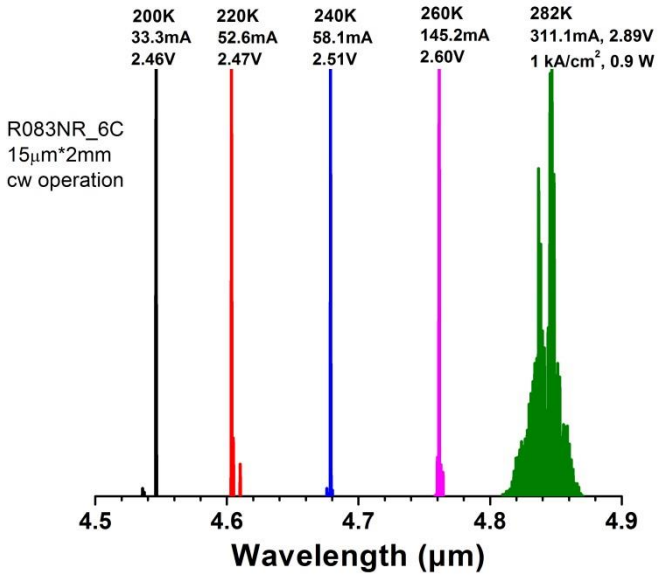


Figure 5-7. CW spectra for device R083NR_6C

5.3 InAs-based IC lasers above room temperature in CW mode

5.3.1 SL as the inner cladding layer

As mentioned in section 4.1.1, the plasmon-enhanced waveguide has a smaller confinement factor of the cascade core and a larger loss than the superlattice waveguide. The physical origin of this disadvantage is that to achieve the same real refractive index as the InAs/AlSb SL, InAs in the cladding layer has to be n-type doped at 1×10^{18} cm^{-3}, an order of magnitude higher than that of InAs in SL, as shown in Figure 5-8. As a result, the threshold current densities of InAs-based IC lasers are generally greater than those of GaSb-based IC lasers.

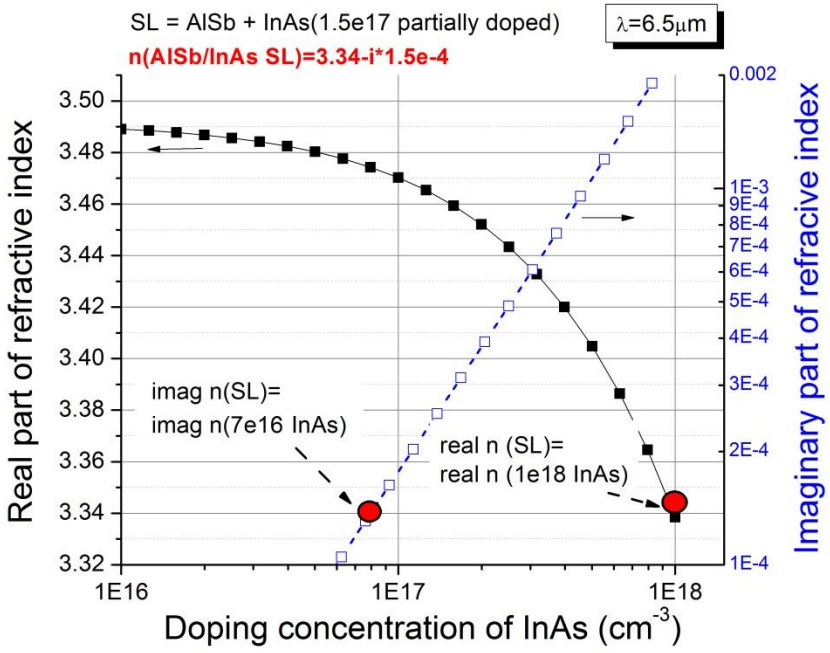


Figure 5-8. Complex refractive index of InAs and SL

Actually, the InAs/AlSb SL has been widely used as the inner cladding layer of short wavelength (*e.g.*, λ~3 μm) InAs-based QC lasers [172-174]. The bandgap of intrinsic InAs at room temperature is 0.354 eV, corresponding to a wavelength of 3.5 μm. For preventing the absorption of emission light by the intrinsic InAs, each SL

period is composed of 25 Å InAs and 15 Å AlSb to achieve an equivalent band gap of 0.7 eV. Furthermore, the cladding layer consists of InAs with an n-type doping of 3×10^{19} cm^{-3}, which shifts the interband absorption edge by more than 0.3 eV due to the band filling effect. This heavily doped InAs has a plasmon wavelength of 7 μm and a real refractive index of -0.3.

Similarly, the InAs/AlSb SL should also be applicable to InAs-based IC lasers. It can be used at the intermediate position between the n^{++} InAs cladding layer and the InAs spacer layer. In the development of semiconductor lasers, "separate confinement layer (SCL)" was used to call the intermediate layer with a refractive index between the core layer and the cladding layer. However, in the IC laser community, the terms "spacer layer" and "SCL" have been used interchangeably to refer to the low-doped InAs or GaSb layers that sandwich the cascade core region. For the short-wavelength QC lasers, the intermediate layer consisting of InAs/AlSb SL was called spacer layer [172]. For the long-wavelength QC lasers, the intermediate layer consisting of low-doped GaAs was also called SCL [121]. For the avoidance of confusion, it's worth to trace back to the origin of SCL.

The first DH lasers, capable of CW operation above room temperature [11], was primarily a three-layer structure: an n-type cladding layer, an active core layer (less than 1 μm) and a p-type cladding layer. Soon an additional layer named "large optical cavity (LOC) " [175] was used to increase the optical width and reduce the catastrophic degradation at the emitting facet. This additional layer had a larger bandgap and a lower refractive index than the recombination region. Researchers realized that further decreasing the active region thickness would result in an expansion of the optical

distribution and an increase of the threshold current density. The symmetric version of LOC was used to optimize the optical profile and the carrier confinement separately [151]. Thanks to the advanced growth technology such as MBE, QW lasers [12] quickly became the new norms, and graded-index (GRIN) separate-confinement heterostructure (SCH) lasers rapidly gained popularity due to their low thresholds and narrow Gaussian beams [13,14]. Even for ultraviolet lasers with wavelengths as short as 250 nm, GRIN-SCH was still a good option [176].

Due to the distinctive nature of type-II broken-band gap QW, the concept of SCL originally used for the type-I direct-bandgap QW is worth revisiting. In IC lasers, the carrier confinement in the active core of InAs/GaInSb QWs is actually achieved by the electron injector and the hole injector. In a larger scale, neither InAs nor GaSb spacer layer is able to prevent the carrier escape from the cascade region (which is also a possible reason for carrier leakage). From the perspective of the optical waveguide, either the InAs spacer in an InAs-based IC laser or the GaSb spacer in a GaSb-based IC laser has a refractive index larger than the cascade stage, which is not able to directly prevent the leakage of the optical wave by total internal reflection . Nevertheless, they can concentrate the optical profile to the center layer so that less optical wave will spread into the outer cladding layer.

It's also worth to note that the concept of "double cladding" was developed in single-mode optical fiber as early as 1974 [177]. In the double cladding structure, the **inner cladding** even had a smaller refractive index than the outer cladding. As a result, the core area can be made roughly doubled, and the loss can also be reduced for single-mode operation, compared to the single cladding fiber.

After this historical analysis, I think a better name for the InAs/AlSb SL in the short-wavelength QC lasers is "inner cladding layer" to avoid confusion. The reason why QC lasers lack a "spacer layer" as in IC lasers is because of their much larger stage number. As a rule of thumb, the optimal thickness of the spacer layer is that the sum of total spacer layers and the active core is equal to the wavelength in the medium. For example, a $\lambda \sim 4.6$ μm IC laser will have the optimal total thickness (active core + spacer layer) of 4.6/3.4=1.35 μm. For QC lasers, such a spacer layer is avoided by increasing the stage number and making their active core roughly equal to this thickness.

In Chapter 4, two types of waveguide structures are compared in Figure 4-2. Here we introduce a third waveguide structure (blue dotted in Figure 5-9), which consists of an InAs spacer layer, an SL inner cladding layer, and an n^{++} InAs outer cladding layer. The confinement factor of 23% and the waveguide loss of 6.0 cm^{-1} are halfway values between the conventional SL cladding waveguide (black solid) and the n^{++} InAs cladding waveguide (red dashed). This is easily understood because the new waveguide is essentially a hybrid of the other two waveguides. As a result, the new waveguide will have a threshold gain ~30% smaller than the conventional plasmon-enhanced waveguide for InAs-based IC lasers.

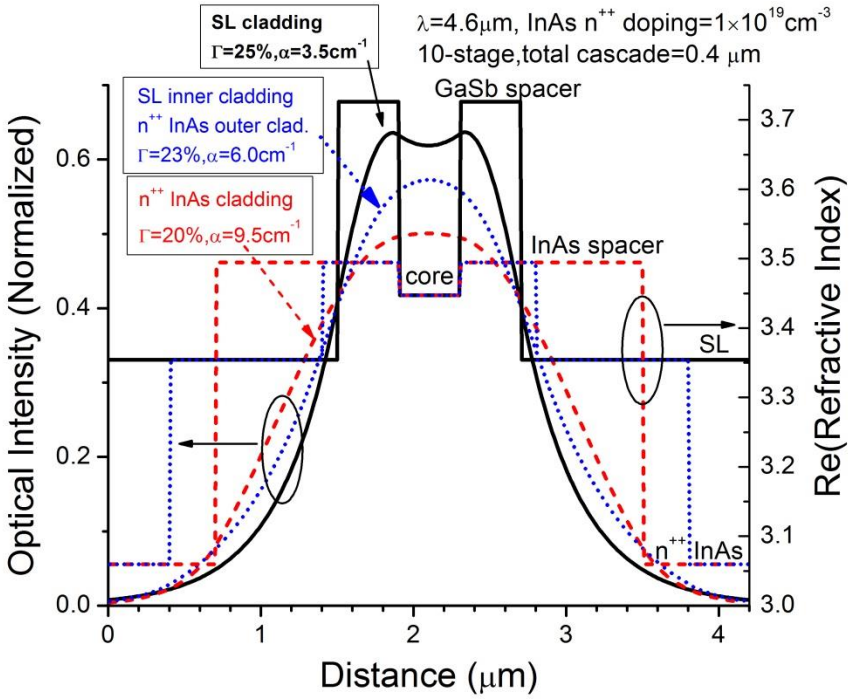


Figure 5-9. Three types of waveguides in IC lasers: GaSb spacer/SL cladding (Black solid), InAs spacer/SL inner cladding/n^{++} InAs outer cladding (blue dotted), and InAs spacer/ n^{++} InAs cladding (red dashed)

5.3.2 Experimental results

IC laser structures with the waveguide similar to the green dotted line in Figure 5-9 were grown on *n*-type InAs substrates in a Gen-II MBE system. The cascade stage was similar to Figure 5-5 (a). Three InAs QWs in the injector region were doped with Si to 3.3×10^{18} or 1.6×10^{18} cm^{-3}. The superlattice (SL) in the inner cladding was composed of 25-Å-thick InAs and 23-Å-thick AlSb, in which each AlSb layer contains a 3-Å-thick AlAs interface for strain balance. A digitally graded InAs/AlSb(As) QW region was inserted as a transition/connection bridge between the SL layers and other regions for smoothing carrier transport, similar to GaSb-based IC lasers.

The laser wafers were processed into deep-etched broad-area (150- and 100-μm-wide) mesa stripe lasers and narrow-ridge (10-, 12-, 15-, and 20-μm-wide) lasers by

contact photolithography and wet chemical etching. The laser bars were mounted epilayer side up on copper heat sinks with indium solder, and placed on the cold finger of a cryostat for measurements in CW and pulsed modes. In pulsed measurements, the applied current pulse width was 1 μs at a repetition rate of 5 kHz. When the applied current was larger than 1 A, the pulse width was reduced to 250 ns to avoid possible Joule heating.

In pulsed operation, a broad-area (BA) device from a 15-stage wafer R140 had a threshold current density J_{th} of 247 A/cm^2 near 4.6 μm at 300 K, the lowest ever reported among mid-IR semiconductor lasers at similar wavelengths. Another BA device from a 10-stage wafer R144 lased at temperatures up to 377 K near 5.1 μm, the highest operating temperature reported for electrically-pumped interband lasers at this wavelength.

Figure 5-10 (a) shows threshold current densities for several representative broad-area and narrow-ridge (NR) lasers made from 12-stage and 10-stage wafers are plotted as a function of the heat-sink temperature. The characteristic temperature, T_0 (~46-57 K in the neighborhood of 300 K), is comparable to that of state-of-the-art GaSb-based IC lasers in the 3-4 μm wavelength region [34].

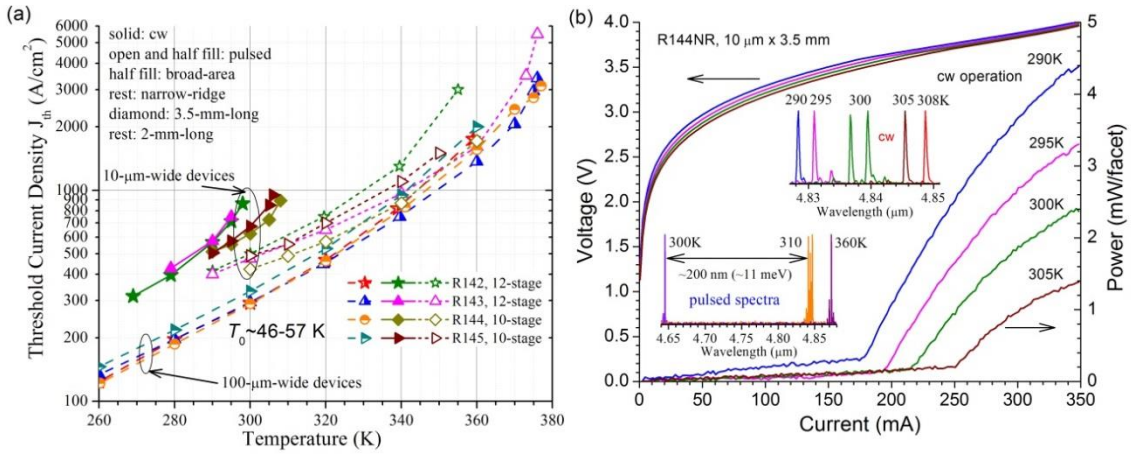


Figure 5-10. (a) Threshold current density vs. heat-sink temperature (b) Current-voltage-light characteristics for a narrow-ridge device in CW operation. Two insets show CW and pulsed lasing spectra, respectively

NR devices (from 10- and 12-stage wafers) with a top layer of ~4 μm electroplated gold were able to work in CW mode at room temperature and above at wavelengths from 4.6 to 4.9 μm, as indicated in figure 5-10 (a). The CW operating temperatures achieved by the 10-stage lasers were higher than those of the 12-stage lasers because the thermal resistance is smaller in 10-stage lasers.

Figure 5-10 (b) shows a 10-stage NR device with a long cavity (3.5 mm) that was able to work in CW mode at temperatures up to 308 K near 4.85 μm. The maximum threshold current density is limited to below1 kA/cm^2 in CW operation, as shown in Figure 5-10(b).This may imply their relatively high thermal resistance. By comparing threshold current densities in CW and pulsed modes, the specific thermal resistance for the 10-μm-wide ridge lasers was 6.5~11 Kcm^2/kW, which is still higher than the reported value (5.4 Kcm^2/kW) for a 12-stage 10-μm-wide ridge IC laser with a similar structure and thicker SL cladding layers [57]. This suggests there is still room for improving thermal dissipation of these InAs-based IC lasers even without employing epilayer-down mounting.

In pulsed mode, the NR lasers worked at temperatures up to 376 K, which is comparable to BA lasers. However, their threshold current densities were generally higher than that for BA lasers at temperatures below 320 K, as shown in Figure 5-10 (a). For example, the pulse threshold current density at 300 K in a 10-μm-wide NR laser was about 45% to 71% higher than that in a BA laser with the same cavity length. This difference is substantially greater than the ~21% difference reported for IC lasers in the 3-4 μm wavelength region [178]. This suggests a somewhat significant current leakage from the sidewalls due to imperfect passivation, and implies further room to achieve better performance by reducing this surface leakage.

Another behavior that was observed from a NR device was a substantial mode hopping (~200 nm in wavelength or ~11 meV in transition energy, which is within the width of a typical material gain peak (>20 meV)) when the temperature was changed from 300 to 310 K, as shown inset in Fig. 5-10(b). This type of mode hopping was not observed from BA devices made from the same wafer. Hence, material non-uniformity may be one of the several possible factors causing the frequency-mode hopping.

In summary, InAs-based IC lasers were demonstrated with low threshold current densities and CW operation at room temperature. This progress was made by the insertion of SL as inner cladding in the waveguide, which enhances the optical confinement and reduces internal absorption loss.

Chapter 6 Long-wavelength InAs-based interband cascade lasers

6.1 Challenges of long-wavelength interband cascade lasers

While the GaSb-based interband cascade (IC) lasers have been achieved with outstanding performance in the wavelength range of 3~4 μm, the InAs-based IC lasers are developed with the ultimate goal of long wavelength operation, where the waveguide thickness is increased proportionally for sufficient optical confinement. Compared with the superlattice cladding layer, the n^{++} InAs cladding layer has the advantages of an easier MBE growth and a significantly larger thermal conductivity (as in table 3-1), but at the cost of a smaller confinement factor and a larger waveguide loss (as in figure 4-2). The pros and cons have to be weighed for a chosen approach.

Long wavelength IC lasers have been reported to have high threshold current densities. For example, the threshold current density for a 10-stage (or 11-stage) λ~6 μm InAs-based IC laser is about 1 kA/cm^2 near room temperature [70,179]. The higher threshold than that in short wavelength is partially due to the not yet optimized design of waveguide and cascade, but the fundamental physical origin should attend to the wavelength dependency of loss and gain.

On the one hand, the optical loss due to free carrier absorption is known to increase with wavelength by λ^2, as proved in equation (4.1.15). On the other hand, the differential material gain is expected to decay exponentially, as discussed in section 5-2-1. As a result, the threshold carrier density is increasing rapidly to compensate for the increased loss. A numerical model to predict the wavelength dependence of threshold current density will be interesting and worthwhile. For a fair comparison in the calculation, each region of the waveguide, including the cladding layer, the spacer layer

and cascade core region, is thickened proportionally with wavelength. Such treatment ensures the same optical confinement for each wavelength. The wavelength dependence of internal loss and Auger coefficient is described in [25].

Figure 6-1 shows a tentative calculation of the wavelength dependent threshold current density with the above ideas. This predicts a threshold current density below 800 A/cm^2 at 7 μm. However, an accurate model requires more experimental data and more considerations to bridge the gap between the two waveguides.

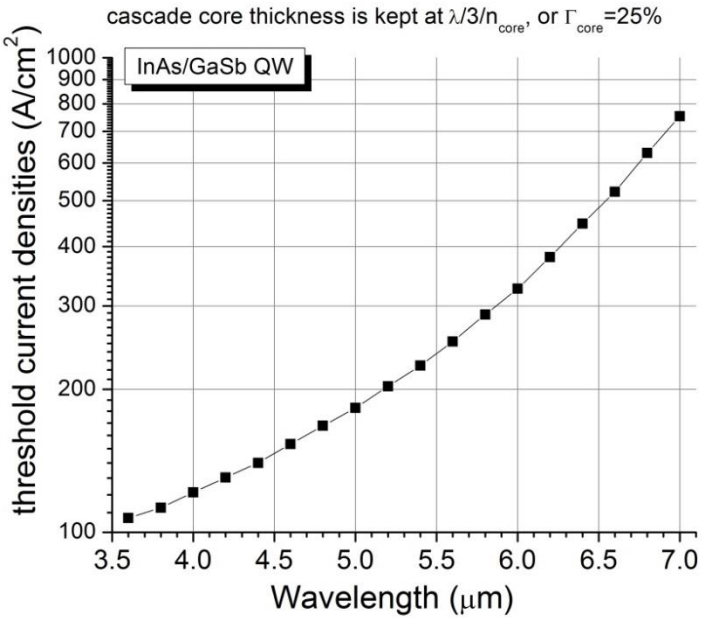


Figure 6-1. Calculated threshold current densities as a function of wavelength.

6.2 Interband cascade lasers at wavelength beyond 11 μm

In this section, attempts are made to extend the lasing wavelengths beyond 11 μm. The design considerations and simulations are presented. The preliminary results show unexpected high threshold current density, which may be related to the negative differential resistance found in the IV testing. Amplified spontaneous emissions around 12 μm are also presented.

6.2.1 Design considerations and simulations

Before this work, the longest wavelength achieved by IC lasers was 10.4 μm [180]. The laser wafer, labeled as R069, has the cascade structure as shown in Figure 6-2 (a) and a waveguide structure as shown in Figure 6-3. One laser from this wafer, with the size of 20 μm×1 mm, has a very low threshold current density of 10 A/cm^2 at 80 K. Its threshold increases sharply to 970 A/cm^2 at the maximum operating temperature of 190 K in pulsed mode.

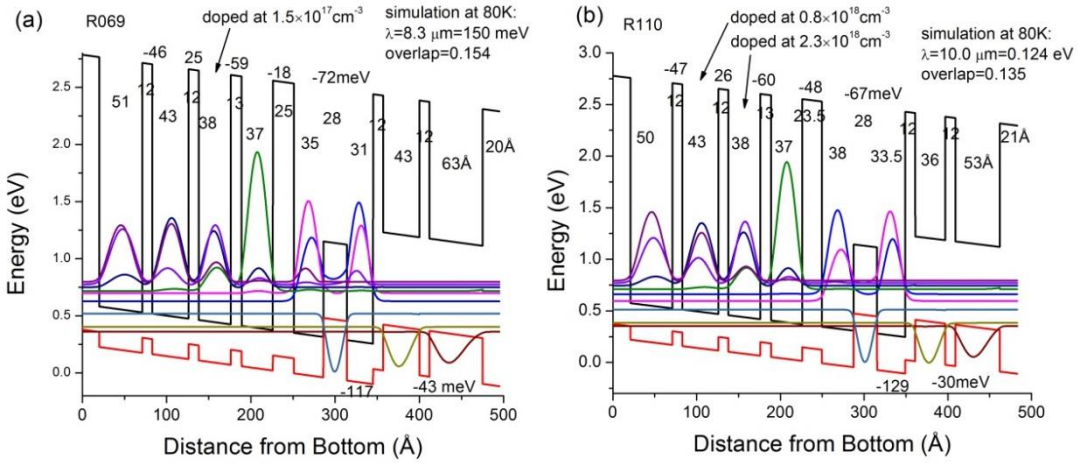


Figure 6-2. Band diagrams of one complete cascade stage and eigenenergy alignment for (a) R069 (b)R110

It is interesting to explore the long wavelength limit for IC lasers. In principle, the interband transition energy in a type-II QW active, not restricted by the bandgap of the continuant material, can be extended to THz frequency range [181] by tailoring the QW layer thickness. However, the greatest challenges are the rapidly increased loss and the even more sharply decreased differential gain, as discussed in the previous session. So the design guidelines are (a) to minimize the possible internal loss by reducing doping, and (b) to increase the modal gain by increasing the number of stages as well as optimizing electron transport in the injector and wavefunction overlap in the active region.

The cascade structure and the calculated wavefunction of a long-wavelength laser wafer R110 are shown in Figure 6-2 (b). Compared with R069, the thicknesses of two InAs QWs in the active region are increased by 3 Å for interband transition at a longer wavelength. Because the electron wavefunction spreads more in the InAs QWs, the overlap between electron and hole wavefunctions is reduced by 12%, from 15.4% to 13.5%. For compensation of this reduction, the stage number is increased from 15, as in R069, to 20. Another difference between these two structures is that two QWs of the electron injector in R110 are heavily doped in the order of 10^{18} cm^{-3}, and only one InAs QW of the injector in R069 is moderately doped in the order of 10^{17} cm^{-3}.

Figure 6-3 shows the waveguide comparison between R069 and R110. Both of them have a 35-nm-thick n^{++} InAs top cladding layer, and the SiO_2 and metal layers also serve as the top cladding layer. For the bottom InAs cladding layer, both of them use a doping concentration of 7×10^{18} cm^{-3}. The confinement factor of the cascade core for R110 is slightly larger than R069. This is because R110 have a greater number of

cascade stages and a smaller refractive index of the n^{++} InAs layer. However, the latter comes at a huge price that leads to a waveguide loss increased by 80%. The significant loss of the InAs cladding layer is due to the closeness of lasing wavelength to plasmon wavelength. For R110, the lasing wavelength of 10.8 μm is actually the same with the plasmon wavelength. Consequently, the real part of the refractive is close to zero and the imaginary part, proportional to the free carrier absorption loss, becomes dominant. This scenario should be avoided in the future design, by reducing the doping concentration to push away the plasmon wavelength.

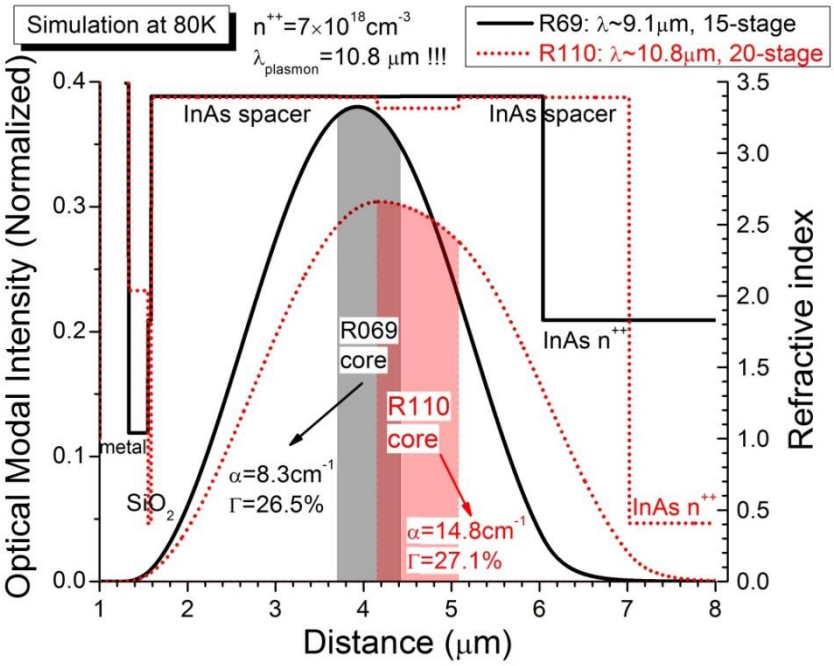


Figure 6-3. Waveguide comparison between R069 (black solid) and R110 (red dotted). Complex refractive indices at 80K are corrected by temperature coefficients.

Note that in Figure 6-2 (b), the eigenenergy difference between the InGaSb QW and the next GaSb QW is 129 meV, which is close to the transition energy. This may cause additional loss such as Auger recombination. In addition, the injection doping concentration should be reduced according to the target temperature.

6.2.2 Experimental results

The IC laser structure R109, as presented in the last section, was grown on an epi-ready S-doped (n~2.5×10^{18} cm^{-3}) InAs substrate using an Intevac GEN II MBE system equipped with a valved As cracker and an unvalved Sb cracker [68]. From the material characterization techniques such as differential interference contrast microscopy and X-ray diffraction, the structure has an excellent crystalline quality, with the average cascade stage thickness only 3% more than the design value.

After the growth, a piece of the wafer was fabricated into broad-area IC lasers. The 150-μm-wide mesas were defined using UV contact lithography and wet chemical etching. The etching proceeded to the bottom InAs spacer to prevent lateral current spreading. A 220-nm-thick, 90-μm-wide SiO_2 insulating layer was deposited and lithographically defined (through lift-off) in the center of the mesa stripe, followed by the deposition and lithographic defining (through lift-off) of a 30/300-nm-thick and 100-μm-wide Ti/Au layer that contacted on the top InAs layer for current injection. After thinning the wafer down to ~150 μm, the processing was finished by Ti/Au metallization on the substrate side of the wafer. The fabricated wafer was cleaved into bars with cavity lengths ranging from 1 to 2 mm, which was mounted epi-side up on a copper heat sink with indium solder and then wire bonded.

In pulse mode (1 μs at 5 kHz), a 1.6-mm-long laser worked near 10.8 μm at 80 K, and operated at temperatures up to 130 K near 11.2 μm, as shown in Figure 6-4(a), the longest wavelength achieved among III-V interband lasers. At 80 K, the threshold current density was 72 A/cm^2, much higher than the ~10 A/cm^2 for IC lasers R069 near

~9 μm. Even if we consider the doubled loss in the n^{++} InAs layer, this threshold is still higher than our estimation. In the light-current characteristics, the light intensity of the laser did not increase rapidly immediately after the quasi-lasing spectra were observed, which implies a tiny differential gain or a relatively high loss.

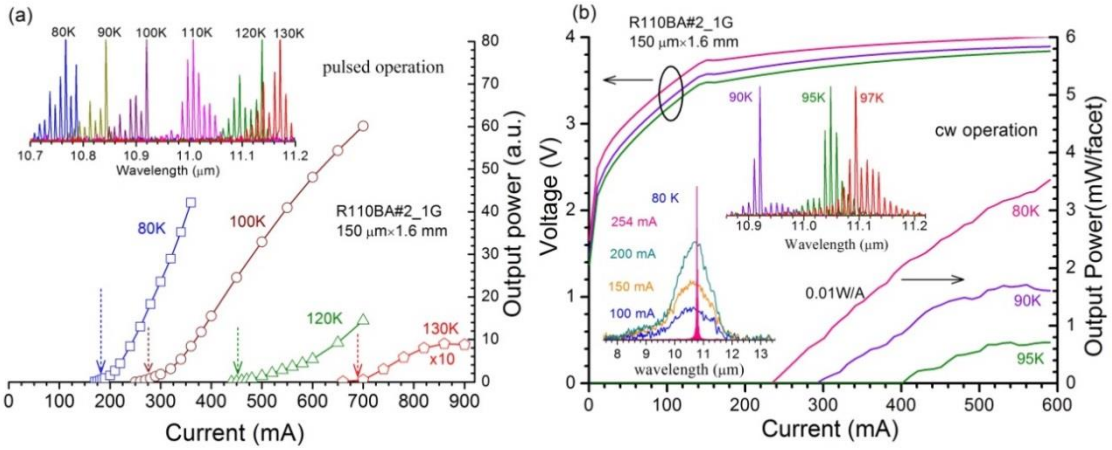


Figure 6-4. (a) Light-current curve for a broad-area IC laser in pulsed mode. The arrows point to threshold currents where spectra were taken. Inset: pulse spectra. (b) Current-voltage-light characteristics in CW operation. Insets: CW lasing spectra at 80 K to 97 K; emission spectra at 80 K with different injection currents.

In CW mode, the same device worked up to 97 K with an emission wavelength at 11.2 μm, as shown in Figure 6-4 (b). The power slope efficiency was only about 0.01 W/A at 80 K, corresponding to a differential external quantum efficiency (per stage value) of less than 1%. If this is mainly due to the internal loss, then the loss will be an order higher than the waveguide modeling suggests. So a very low internal efficiency or injection efficiency was also accompanied in this scenario.

6.2.3 Negative differential resistance

With a close scrutiny of the current-voltage characteristic in Figure 6-4(b), we identify a kink around 150 mA, which is well before the threshold current as revealed

by the light-current curves in the same figure. For further investigation of this unusual behavior, emission spectra were taken at several currents from 100 mA to 254 mA. As shown in the inset of Figure 6-4 (b), electroluminescence below the threshold (~231 mA) were broad (>120 nm), in contrast to the narrow lasing spectrum. This confirmed that the device did not lase before the current reached 231 mA.

The kink is actually where a negative differential resistance (NRD) appears if the first derivative of the current-voltage curve is taken. Furthermore, the temperature dependence of NRD is studied to investigate its possible origins. Figure 6-5 (a) shows the I-dV/dI curves for the same laser at a series of temperatures. We can see the NRD appeared earlier as the temperature increases. At temperatures above 210K, the abrupt drop in the differential resistance disappeared. For clarity, the current and voltage were extracted when the NRD was identified, as shown in Figure 6-5 (b). Both of them exhibited a linear decrease with the temperature. The temperature dependence can be related to the resonant tunneling [182], which had been intensively studied on the InAs/AlSb/GaSb type-II broken-gap structure [183]. Briefly, the temperature dependence of the resonant tunneling is because the increased thermal spread of the carrier at higher temperature enhances the non-resonant component of the tunneling current [184]. The detailed mechanism of NDR is worth further investigation.

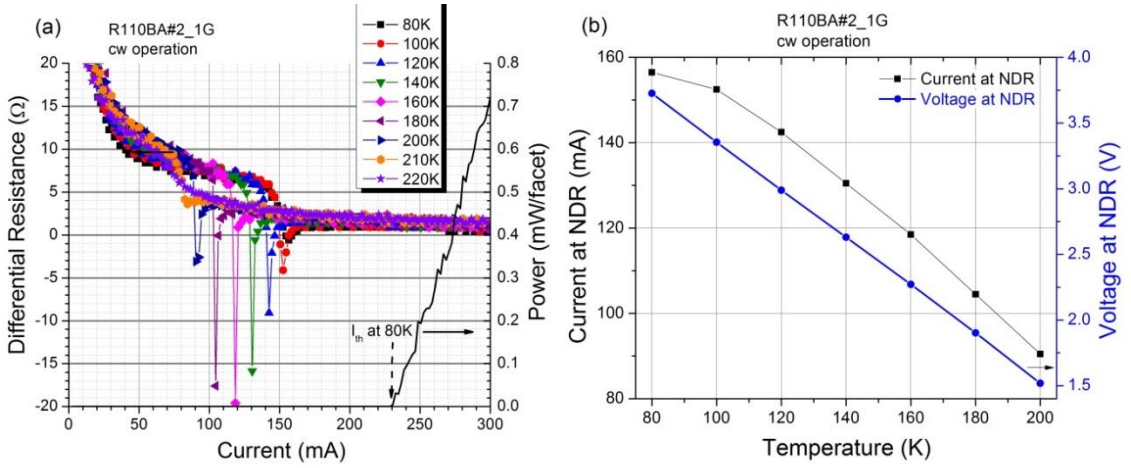


Figure 6-5. (a) Differential resistance as a function of dc current at temperatures of 80 K to 220 K.(b) Injected currents and voltages corresponding to the negative differential resistance at different temperatures.

From the same wafer, other IC lasers with larger threshold current densities around 300 mA at 80K, were observed to have the abrupt drop of differential resistance of around 50 mA. This suggests that the NDR is sensitive to the material non-uniformity. It's still not clear whether the NDR is related to the defect state, which is more dominant at low temperature. Further study of NDR should be able to reveal more details of the underlying physics as well as offer some hints on the design of long-wavelength of IC lasers.

6.2.4 Amplified spontaneous emission around 13 μm

The IC laser structures R111 and R112 were designed and grown. They both have the same cascade structure with a target wavelength of 13 μm at 80 K. Because the InAs QWs are thicker, the wavefunction overlap is reduced to 12.1%. For the compensation of this reduction, the stage number is increased to 23. The waveguide structures of them are similar to R109. The only difference between them is that the

doping concentration of the plasmon layer is 7×10^{18} cm^{-3} for R111 and 1.5×10^{19} cm^{-3} for R112.

The waveguide calculation shows that R111 has a confinement factor of 30.9% with a waveguide loss of 22.4 cm^{-1} while R112 has a confinement factor of 31.7% with a waveguide loss of 23.4 cm^{-1}. Although a higher confinement factor compared to R109, both of them have a much larger loss, which may prevent them from working even at 80 K.

The broad-area lasers fabricated from wafer R111 and R112 indeed could not work at 80 K in pulse mode. From the current-voltage characteristics, the negative differential resistances were also found in lasers from both wafers. The NDR happened at 30~40 mA for lasers from R111, and happened at 10~20 mA for lasers from R112. It's not clear whether the different currents to trigger NDR is related to the different doping in the InAs cladding layer.

Although these lasers could not generate stimulated emission, they should in principle have amplified spontaneous emission (ASE) because the transparency condition at 80 K is easy to achieve. Special attentions are made to get the ASE for the long wavelength emission: (a) choose suitable detector, lens and optical window that have cut-off wavelength longer than the target wavelength; (b) in the optical alignment, to prevent the very weak light from submerging by the noise background, the interferogram received by the detector, should be processed by a lock-in amplifier (*e.g.* Signal recovery 7265) that is able to display the overall intensity; (c) in the spectrum collection, the resolution should be chosen to be relatively coarse and average at

relatively large times to minimize the fluctuation, and background will be subtracted in the end.

With the above tricks, the obtained ASE spectra for typical devices from R111 and R112 are as shown as in Figure 6-6. Both of them have emission wavelengths around 12~13 μm, in good agreement with the calculation. For R111, the ASE spectrum could be obtained at current as low as 25 mA, indicating that the transparency current should be lower than this value. And the gain width of larger than 200 cm^{-1} suggests the injection level is very high, which can be inferred from the relation between Fermi level and gain coefficient. Compared with R111, the intensity of ASE for R112 is slightly smaller. This could be explained by the larger loss due to the higher doping of the cladding layer.

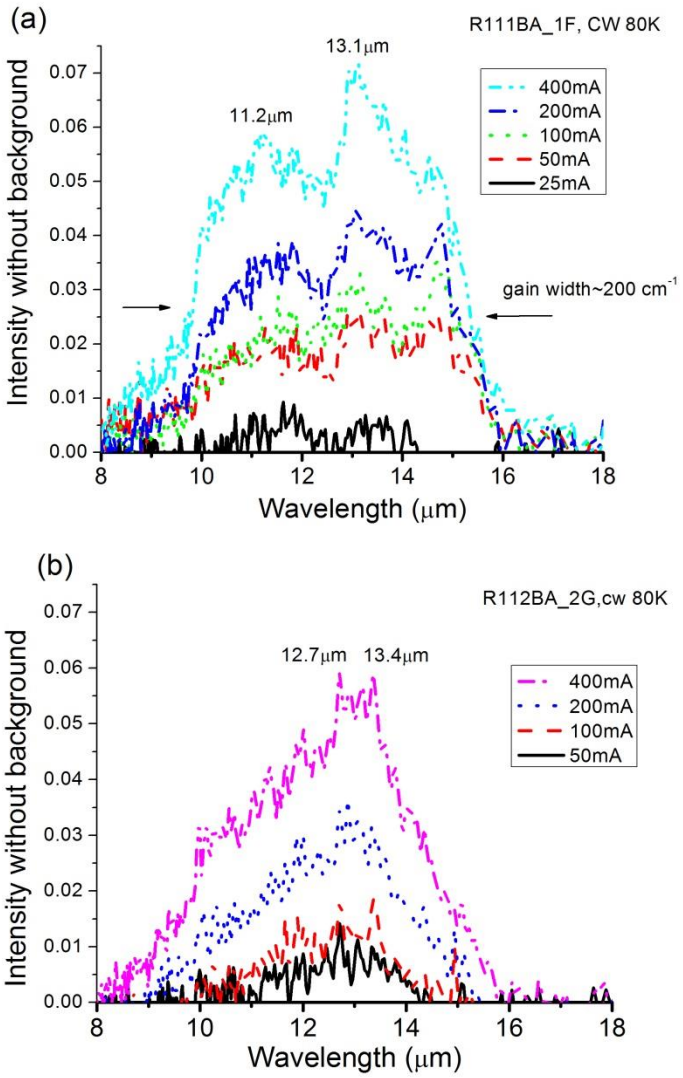


Figure 6-6. Amplified spontaneous emissions at 80K: (a) R111 (b) R112

6.3 Room temperature IC lasers beyond 6 μm

6.3.1 Active region design

For emission at different wavelengths, a simple practice is to vary the thickness of the InAs layers in the "W" shaped active QWs. To counter the Stark-shift effect under an external voltage, the first InAs QW is usually 15%~20% thicker than the second InAs QW. As shown in Figure 6-7(a), the transition energy is nearly linear decreasing with the InAs thickness. The slope, or the tuning rate, is 0.02 eV/Å, in

excellent agreement with the experimental value, 0.18 μm/1 Å, at the lasing wavelengths of 2.97~4.16 μm [185].

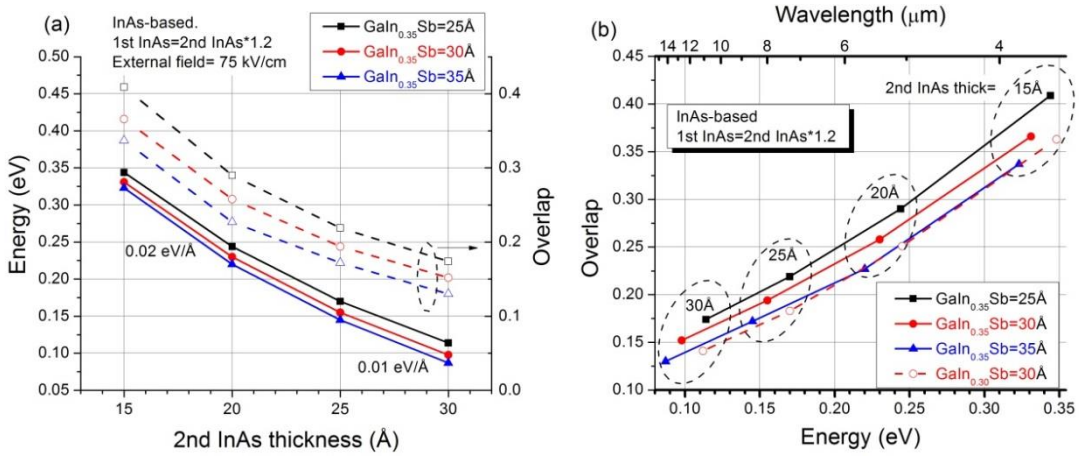


Figure 6-7. (a) Transition energy and overlap integral as a function of InAs QW thickness; (b) The intensity of overlap integral as a function of transition energy.

Another way of tuning the emission wavelength is to change the thickness of the GaInSb layer. However, as shown in Figure 6-7(a), the shift is much less efficient, mainly due to the much shallower potential well for holes and the much larger effective mass [186]. In addition, the GaInSb is not only the QW for holes, but also acts as a potential barrier to the electron QW. As the thickness of GaInSb layer decreases, the hole energy is elevated away from the band gap, but the electron energy is approaching the band gap because the electron sees a thinner barrier. As a result, the transition energy changes very slowly. The change may not be monotonous, as reported in [98].

The overlap integrals of the electron and hole wave functions, calculated by equation (2.3.22), is a figure of merit for the transition oscillator strength. Figure 6-7(b) shows the overlap as a function of the transition energy with layer thicknesses. Generally, the overlap decreases as the energy is reduced. This is in good agreement with the photon reflectance measurement [98]. However, my calculation predicts a

decreasing of overlap for a lower indium content in GaInSb, which is opposite to the results reported in [98]. This discrepancy is not well understood yet. Anyway, the research group at the University of Wurzburg used $GaIn_{0.24}Sb$ with thicker InAs layers and GaInSb layer for their long-wavelength IC lasers [70,71] [179].

Note that the GaInSb layer is under compressive strain on either GaSb or InAs substrate. On the positive side, a larger compressive strain can lift the heavy hole energy with a slightly reduced effective mass. This could partially alleviate the asymmetry of between very light conduction band mass and the heavy valence band mass, leading to a significantly reduced threshold current [17]. On the downside, the strain will set a critical thickness after which the material may have undesirable dislocation and defects.

Two preliminary structures, R139 and R141, are designed with different active regions. The hole QW in the active region is 29-Å-thick $Ga_{0.7}In_{0.3}Sb$ for R139 and 28-Å-thick $Ga_{0.65}In_{0.35}Sb$ for R141, both of which are designed with a wavelength around 6.5 μm. They both have four InAs QWs in the electron injector and two GaSb QWs in the hole injector. Two of InAs QWs in the electrons are doped at 3.3×10^{18} cm^{-3} for R139 and 1.2×10^{18} cm^{-3} for R141. The wavefunction alignments in both structures are similar to Figure 6-2.

6.3.2 Waveguide simulation

The introduction of InAs/AlSb superlattice (SL) into the waveguide of InAs-based IC lasers at λ~4.6 μm, has significantly improved the optical confinement, resulting in reduced threshold current densities and CW operation above room

temperature [73]. This strategy should be applicable to IC lasers at a longer wavelength, *e.g.,* λ>6 μm.

For compensation of the reduced wavefunction overlap and larger loss at a longer wavelength, the cascade stage number is increased to 18 for the aforementioned two structures. The InAs/AlSb SL with a thickness of about 1 μm was inserted between the cladding layer and InAs spacer layer. Figure 6-8 shows the waveguide structure and modal profile for R139. The top cladding layer is only 35-nm-thick n^{++} InAs, and the 200-nm-thick SiO_2 layer also serves as the top cladding layer. In contrast, R141 has the top InAs cladding layer of 0.8 μm. The doping in the cladding layers for both structures are the same at 1.0×10^{19} cm^{-3}, which results in a complex refractive index of 2.5466-0.0261i. Note that the plasmon wavelength at this doping concentration is 9.5 μm. The refractive index for SiO_2 at 6.5 μm is 1.212-0.0066i. So SiO_2 seems to be a better waveguide material than highly doped InAs as long as there's no leakage of the optical wave. As a result, R139 has a waveguide loss of 8.7 cm^{-1} and a core confinement factor of 36.6%, while R141 has a waveguide loss of 9.7 cm^{-1} and a core confinement factor of 35.6%. From the perspective of the waveguide, R139 is better than R141.

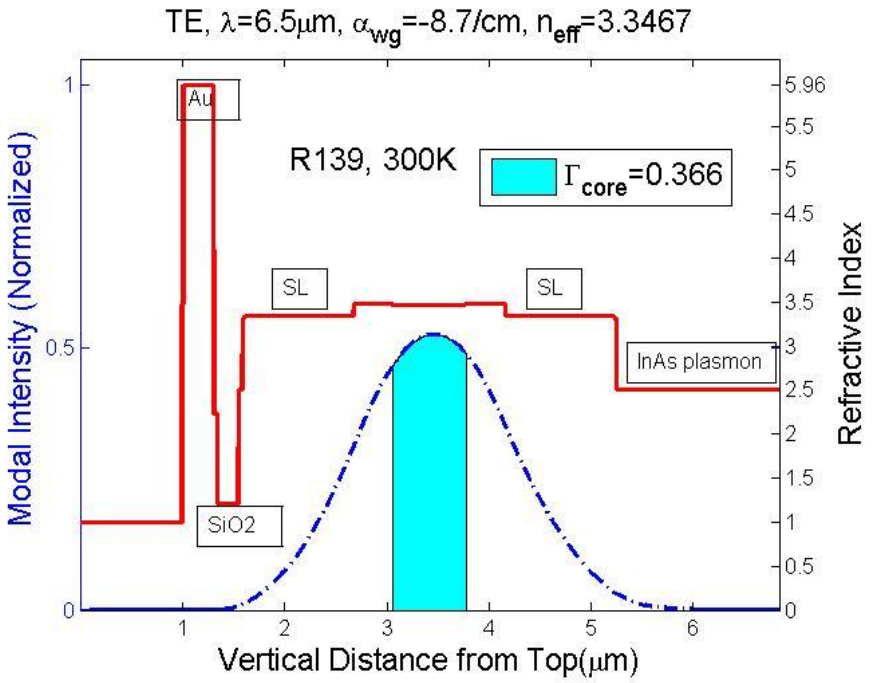


Figure 6-8. Waveguide structure and optical profile for R139.

6.3.3 Experimental results

Both structures are grown by MBE on InAs substrates and routinely processed into broad-area lasers without any special treatment. Figure 6-9 (a) shows the pulse spectra for R139 broad-area lasers with operating temperature up to 324 K at 6.4 μm. Figure 6-9 (a) shows their temperature dependence of wavelength and threshold current densities in pulsed and CW modes. Due to the huge amount of heat accumulated in the broad ridge, the maximum CW operating temperature is around 200 K, with a threshold current density of 120 A/cm^2, as shown in Figure 6-9(b).

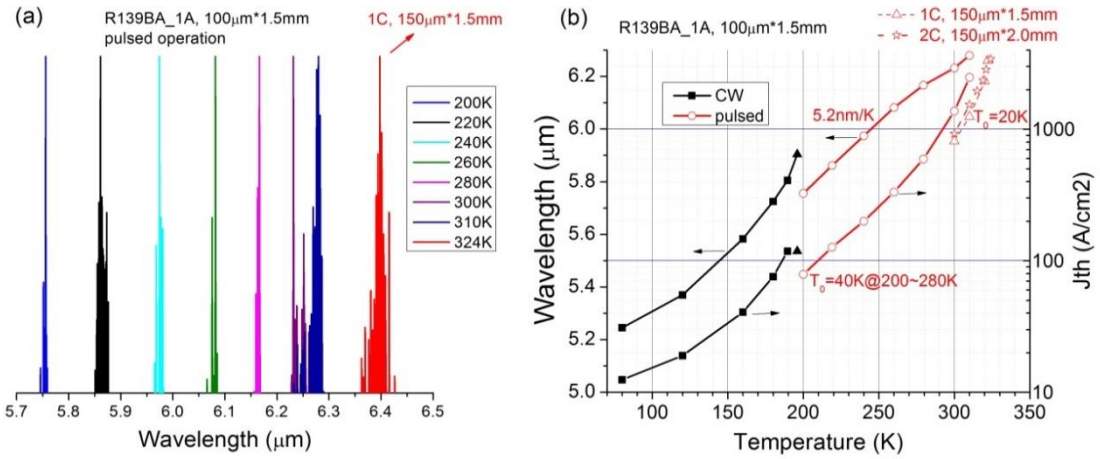


Figure 6-9. Characteristics for R139: (a) pulsed spectra; (b) temperature dependence of wavelength and threshold current densities.

In pulsed operation, the threshold current density at 300 K is as low as 810 A/cm^2 for a 150 μm×2 mm laser from R139. Although a competing group reports a value of 588 A/cm^2 at 293 K, they use a larger number of stages, *i.e.*,22 [187]. And they also achieved T_0 of 41 K, similar to our value of 40 K. At temperatures above 300K, our laser exhibited a rapid increase in the threshold current density with T_0 of 20 K. Note that our laser may work at higher temperature due to the current limit of the current source. At the maximum operating temperature, the threshold current density is as high as 3.4 kA/cm^2, which is an indication of good material quality and structure endurance.

For R141, the lasing wavelength drifted to 6.7 μm due to the intrinsic uncertainty of MBE growth. Although with a wavelength of 0.4 μm longer, the broad-area lasers from R141 can still work up to 310 K. As shown in Figure 6-10, the threshold current densities for R141 is about 20~30% larger than that for R139. The waveguide simulation predicts a ~10% larger threshold for R141. The discrepancy may

come from the wavelength dependent optical loss, the difference in the doping concentration in the injector as well as the top cladding layers.

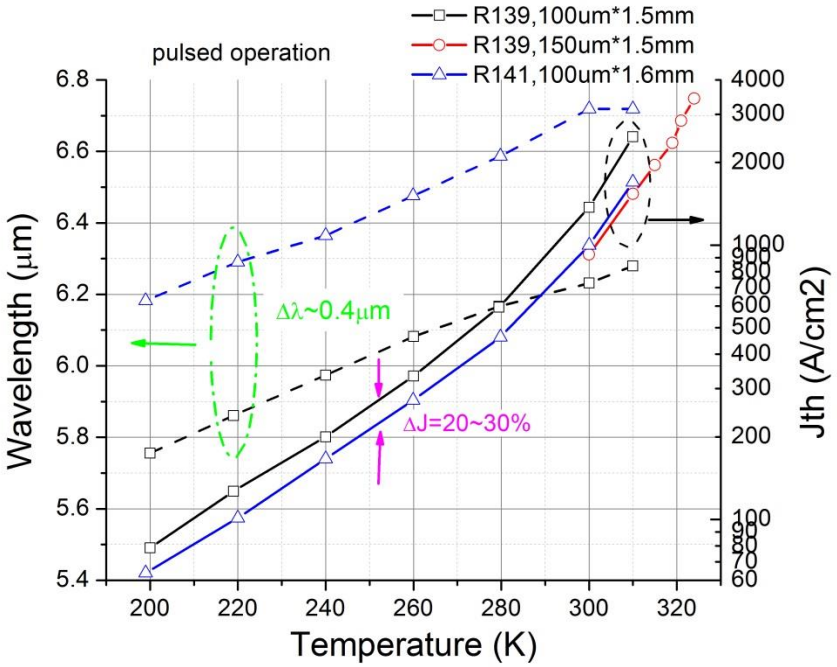


Figure 6-10. Temperature dependence of wavelength and threshold current densities for R139 and R141

6.3.4 Further optimization and current status

These preliminary results are encouraging. There is ample room for further improvement. Several design guidelines on injection, waveguide, and active region have been discussed in section 5.2.1, section 5.3.1 and section 6.3.1, respectively.

For the active region, the optimal indium content is still in doubt, despite a lower indium content suggested by the photon reflectance measurement [98]. The optimal doping is 1.0×10^{12} cm^{-2} divided by the layer thickness, and the InAs QW most adjacent to the active region should be undoped to avoid excessive loss.

For waveguide design, the total thickness of cascade region and the spacer layer is the wavelength in the medium. The choice of optimal thickness of SL layer can be

seen in Figure 6-11, where the confinement factor for the cascade region is nearly linearly decreasing with the increasing SL layer. For a reduction of the high loss in the n^{++} InAs cladding layer, the doping centration is intentionally controlled under 1.0×10^{19} cm^{-3}. Note that the core index is somewhat smaller than R139, possibly due to highly doped InAs.

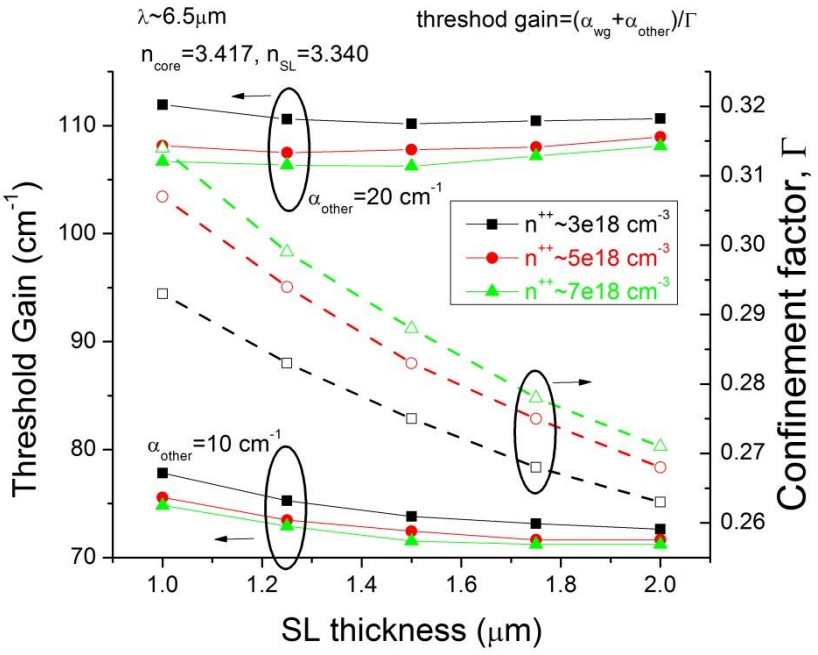


Figure 6-11. Waveguide optimization for 6.5 μm InAs-based IC lasers.

Our results are compared with state-of-the-art InAs-based IC lasers, as shown in Figure 6-12. At present, the 6-7 μm region is under intensive research. CW operation at 0 C was achieved very recently [179]. It's very promising that the CW operation at room temperature will be realized in the near future.

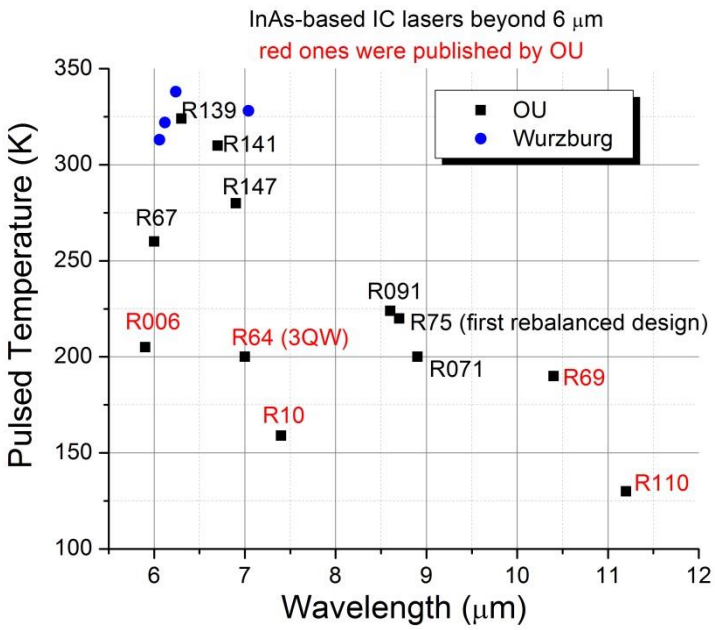


Figure 6-12. Maximum pulsed operating temperatures for IC lasers beyond 6 μm

Chapter 7 Electrically widely-tunable interband cascade lasers

Electrically-tunable interband cascade lasers are demonstrated with a wide tuning range of about 180 cm^{-1} (22 meV in energy or 900 nm in wavelength) at λ~7 μm. The laser structures are designed such that the Stark effect and heating effect act together to enhance the red-shift of the lasing wavelength with a current injection to achieve wide tunability.

7.1 Tuning mechanisms

Semiconductor mid-infrared lasers with a wide tuning range are highly desirable for many applications such as spectroscopy and biochemical analysis [188]. Although various techniques have been developed, there are mainly three spectral tuning mechanisms.

The first is temperature tuning, based on the temperature dependent band gap and refractive index. Although the lasing wavelengths of both QC lasers and IC lasers are not determined by the bandgap of the materials, the conduction band offset and alignment are still affected by the different temperature coefficients of bandgap between the well material and the barrier material [189]. Furthermore, the injection of current in CW operation can generate a significant amount of heat and significantly raise the laser core temperatures. As a result, a typical temperature tuning coefficient for QC laser is -0.078 cm^{-1}/K at 9.1 μm, with a power tuning coefficient of -1.51 cm^{-1}/W [18]. In contrast, a typical temperature tuning coefficient for an IC laser is 1.88 nm/K (or ~1.4 cm^{-1}/K) around 3.7 μm [185], which is much larger than that in QC lasers due to their

distinct natures of transition. For emission at 4.3 μm, the temperature tuning coefficient is about -1.0 cm^{-1}/K for IC lasers [190] and about -0.47 cm^{-1}/K for QC lasers [191].

The tuning of single-mode distributed feedback lasers is via the change of refractive index in temperature, which results in a very limited current-tuning range (*e.g.*,~10 cm^{-1}) [192]. This will be further discussed in Chapter 9.

The second is grating tuning, such as external cavity (EC) lasers. The tuning range of 120 cm^{-1} and 110 cm^{-1} were achieved from a λ~8 μm QC laser [193] and a λ~3.2 μm IC laser [194], respectively. However, these lasers are cumbersome and require well-aligned external optics components. The long-term mechanical and vibrational stability also make their practical application less appealing.

The third is electric field tuning, based on the Stark effect in a diagonal transition [195]. The Stark shift of the transition energy ΔE can be written as [196]

$$\Delta E = -e \cdot \overrightarrow{z_{e-h}} \cdot \overrightarrow{\Delta F} = -\frac{dE}{dF}\Delta F = -\frac{dE}{dU}\Delta U \tag{7.1}$$

where -*e* is single electron charge, ΔF is the change of the electric field *F* across the active region. The displacement of the probability centroids from electron to hole, $z_{e\text{-}h}$, actually determines electric field tuning rate dE/dF or voltage tuning rate dE/dU .

The diagonal transition in QC lasers can have a large $\left|\overrightarrow{z_{e-h}}\right|$, *e.g.*, ~100 Å, which resulted in a broad tuning range of over 220 cm^{-1} for intersubband electroluminescence [196].

However, the real challenge of Stark effect tuning is that the electric field is difficult to change after the laser reaches its threshold. As clearly shown in Figure 2-1, the gain coefficient is nearly "locked" at the laser's threshold value, and the carrier

densities are almost pinned at the threshold. The solution is how to break or shift the balance between the modal gain and the optical loss.

One approach is to use multiple sections [196,197] or heterogeneous active regions [198], to create different threshold gains for different sections or different wavelengths. The injection of current is through multiple contacts. An extensive tuning range can be achieved, which comes at the cost of higher threshold current and additional device fabrication steps.

Two different designs on the band structure of QC lasers had achieved voltage tuning ranges of about 100 cm^{-1}. The first design, dubbed "two-step" coupling [199], is to insert a coupling state to slow down the stimulated emission above the threshold. The slow down, however, implies an internal leakage path in the structure. The second design, based on a bound-to-continuum active region with a relatively short and wide miniband, were able to maintain a wall plug efficiency up to 11.5% [200].

For IC lasers, a large Stark tuning is possible because the spatially diagonal transition in the type-II QW allows a large nonzero dipole moment $\overrightarrow{z_{e-h}}$. Until this work, there had been only one report on electrically tunable IC lasers with a blue-shift tuning range of 120 cm^{-1} (15 meV in energy or 120 nm in wavelength) near 3.3 μm at 80 K [201]. The active region of these tunable IC lasers consists one InAs QW and one GaInSb QW, as shown in Figure 7-1(a). At a common electric field of 75 kV/cm, the dipole moment $z_{e\text{-}h}$ is -24.6 Å, doubled the value of -12.8 Å for the mainstreamed W-like active region with two InAs QWs as shown in Figure 7-1(b).

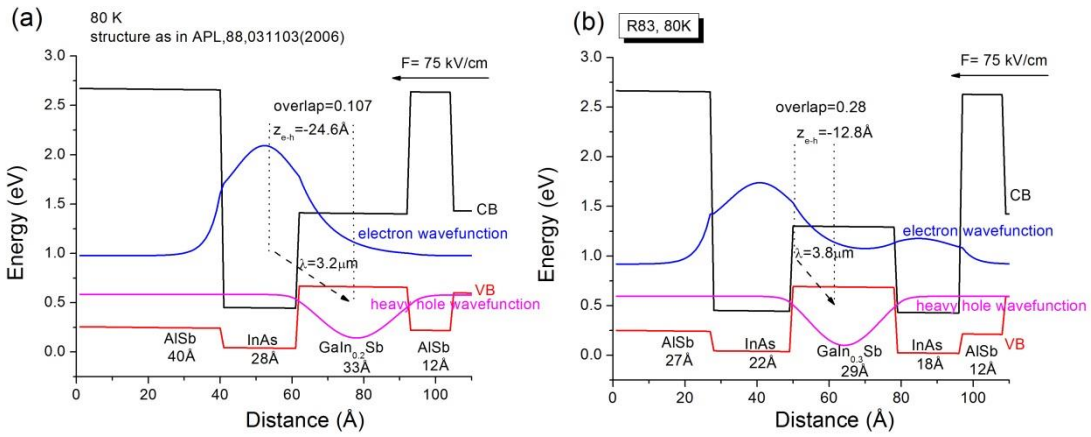


Figure 7-1. Transition dipole and wavefunction overlap for active regions with one QW(a) and two QWs(b)

Gregory Belenky's group attributed the unpinning of carriers to "charge accumulation layers," [201] or "collection quantum wells" in a type-I QW laser [202], which was made by the insertion of thick tunneling barriers. Carrier transport was slowed down, and hence, carriers continue to accumulate after the threshold, leading to a continuous increase of the voltage across the active region. However, their claim is dubious. Because the growth of carrier in "charge accumulation layers" doesn't necessary lead to the growth of carrier in the active region. In addition, the AlSb barrier before the W-like active region can also be thickened to achieve the function of "charge accumulation layers." But they didn't have such experiments to draw a fair conclusion.

In my understanding, the large tuning range for the one-InAs-QW active region is because of the quick shift of the balance between gain and loss. As shown in Figure 7-1, the wavefunction overlap factors for single QW and W-QW are 0.107 and 0.28, respectively. The latter is more than twice larger than the previous one. Consequently, the differential gain of the single QW active region is much smaller than the W-like active region. The thicker AlSb layer actually decreases the injection efficiency and

increases the free carrier absorption loss due to the accumulated carriers on the injector side. These factors together with a large dipole moment result in a large tuning range.

Note that the two structures in Figure 7-1 both have a negative value of dipole moment. So the possible Stark-tuning would be a blue-shift of the transition energy, which is opposite to the red-shift trend caused by the heating effect in a CW operation. For practical application, it's more desirable to have a red-shift version of the Stark tuning.

7.2 Tunable IC laser structures and modeling

A 15-stage laser structure R64 features an active region consisting of three InAs QWs separated by two GaInSb QWs. The specific thickness of each layer is shown in Figure 7.2. Other parts of the laser structure are similar to R48 (discussed in section 5.1.1), in which the injection region is not heavily doped, and the waveguide is asymmetric.

This structure was initially designed to improve the laser operation temperature by increasing the wavefunction overlap. However, it turned out that the maximum overlap for such structure was acutely sensitive to the electric field. For example, the maximum overlap is achieved at the electric field of about 32 kV/cm, as shown in Figure 7-2, when the fundamental heavy hole wavefunction spreads over the two GaInSb QWs and the fundamental electron wavefunction spreads over the three InAs QWs. Under the electric field with the direction from right to left, the hole carriers tend to move to the left GaInSb QW and the electron carriers tend to shift to the middle and the right InAs QWs. As the electric field increases, the heavy hole wavefunction

quickly concentrates on the left GaInSb QW due to its large effective mass. This results in a sharply decreased overlap with the electron wavefunction. So a larger overlap for the 3-InAs-QWs [203], possibly achieved in a narrow range of electric field, is not sustainable for the practical operation. On the contrary, a significantly reduced overlap factor, together with a positive dipole moment, is useful to obtain a red-shift Stark tuning.

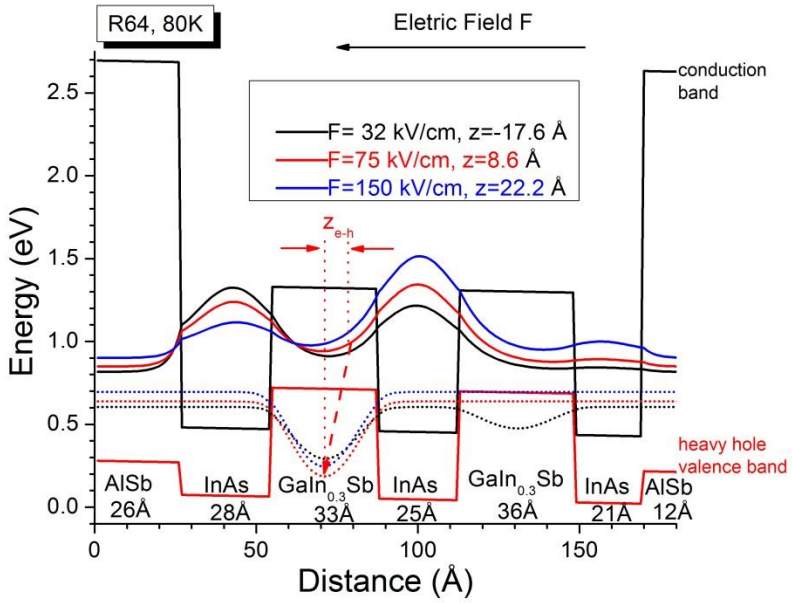


Figure 7-2. Fundamental electron and hole wavefunctions in the active region of structure R64 at different electric fields. The band alignment and the diagonal interband transition (dashed arrow) are plotted at F=75 kV/cm.

Figure 7-3 compares the transition energy as a function of the electric field for typical active regions with a different number of InAs QWs. The threshold electric field is assumed to be 75 kV/cm for all the structures. We can see from the figure that the absolute electric field tuning rate is in the range of 0.05~0.13 meV/(kV/cm). The field tuning rate is essentially the dipole moment according to equation (7.1). For example, dE/dF= -0.10 meV(kV/cm) corresponds to a dipole moment of -e·10 Å. However, the

electric field is not readily measurable. Hence, the field tuning rate is further converted to the voltage tuning rate. The electric field is nearly uniform according to SVM [165] across one cascade stage of about 400 Å. Consequently, a dipole moment of -e·10 Å corresponds to a voltage tuning rate of -0.025·e per stage.

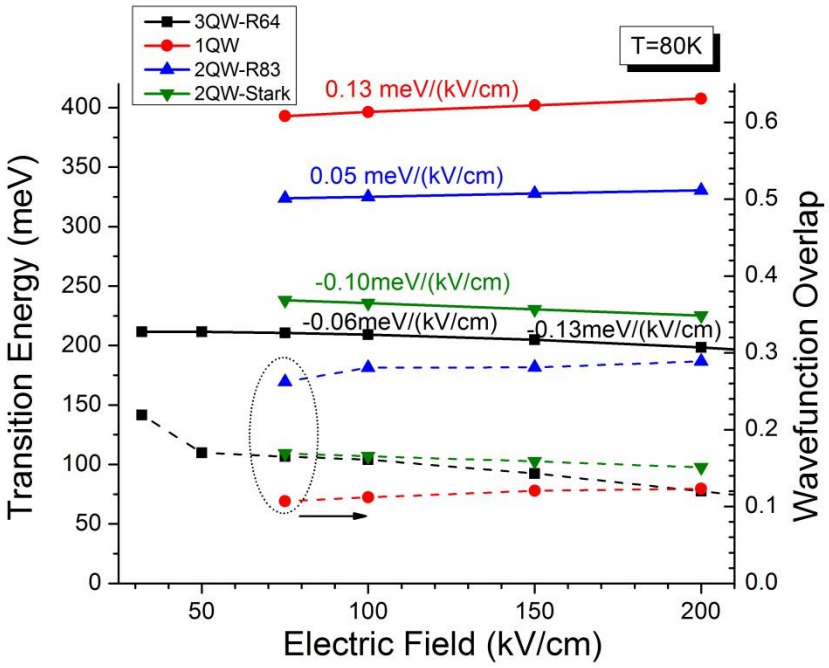


Figure 7-3. Calculated fundamental transition energy and wavefunction overlap as a function of electric field for active regions with different numbers of QWs.

The comparison of the wavefunction overlap for different active region designs, shown in Figure 7-3, is actually more important. According to the calculation, the overlap factor is much smaller for lasers with one or three QWs in the active region, compared to the conventional two-QW laser represented by R83. This is the key point that shifts the balance between gain and loss.

As shown in Figure 2-1, no real laser can achieve 100% of carrier pinning above the threshold. The extra injection above the threshold will unavoidably introduce more loss until a new balance is reached. A recent publication also shows the carrier density dependence of the internal loss [108]. However, some lasers seem to have "good" carrier pinning, because their large differential modal gain can quickly compensate the additional loss above the threshold. These lasers also have "good" performances such as

high-temperature operation and high output power. For the same reason, "bad" carrier pinning is due to the small differential modal gain and significant differential optical loss. So the carrier concentration has to increase significantly.

For laser structure R64 with three InAs QWs, the low transition energy implies relatively large free-carrier absorption loss compared to λ=3~4 μm lasers. In addition, the small overlap even decreases with the electric field. These factors together help to unpin the carrier density in the active region.

According to the above analysis, a red-shift Stark tuning should also be possible for the active region with two InAs QWs. Another new design, labeled as "2QW-stark" in Figure 7-3, increases the thickness of the right QW of R83 from 18 Å to 27 Å. Calculation predicts this design can achieve similar tuning effect as R64.

7.3 Experimental results

The narrow ridge lasers fabricated from R64 had threshold current densities of about 10 A/cm^2 at 80 K with an emission wavelength of 6.7 μm. The low threshold current density indicates a small transparency carrier density. A 20 μm× 2.1 mm laser worked at temperatures up to 188 K in CW mode, and up to 200 K in pulsed mode at 7.6 μm. The maximum pulsed operating temperature is the same as reference [201].Figure 7-4 shows the CW *L-I-V* characteristics of this laser. The slope efficiency at 80 K is only 0.1 W/A per facet, corresponding to a differential quantum efficiency of 10% per stage. Assuming an injection efficiency of 100%, the upper bound of the internal loss is 45 cm^{-1}. The large internal loss is partially due to a not optimized waveguide structure, which was originally designed for a shorter wavelength laser similar to R48. As a result, the characteristic temperature T_0 of 30.5 K is lower than typical values (*e.g.*, 40~50 K) for IC lasers based on a 2-InAs-QW active region .

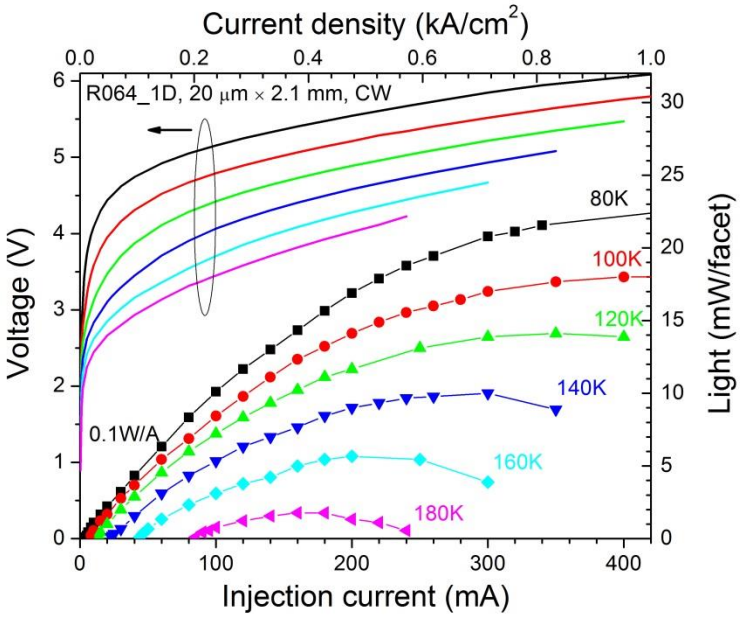


Figure 7-4. Light-current-voltage characteristics of an IC laser (20 μm × 2.1 mm) from R064 at different heat-sink temperatures in CW mode.

The electrical tuning spectra for a 15 μm×1.8 mm laser at 80 K were shown in Figure 7-5. The lasing wavelength was red-shifted by about 180 cm^{-1} (22 meV or 900 nm) with a tuning rate of -0.49 cm^{-1}/mA. The tuning range covered over 12.5% of the center emission wavelength, indicating a significant electrical tuning. The change of voltage was 3V, which was a strong evidence of carrier unpinning. Note that the voltage efficiency at threshold is 77%, which may imply some possible carrier leakage.

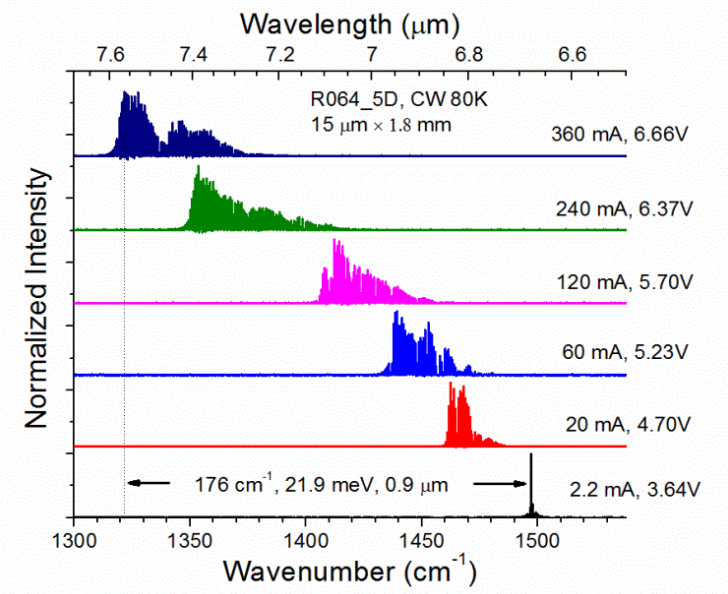


Figure 7-5. CW tuning spectra of an IC laser (15 μm × 1.8 mm) from R64 at 80 K.

For separation of Stark effect from heating effect, the tuning spectra of a 15 μm×1.8 mm device were measured at 80 K in CW mode and pulsed (0.5 μs×1 kHz) mode. As shown in Figure 7-6, the tuning ranges at the same current range are 100 cm^{-1} (13 meV) in CW mode and 85 cm^{-1} (11 meV) in pulsed mode, respectively. This indicates that the wavelength shift was mainly due to the Stark effect. For this device, the change of voltage is 2.44 V, corresponding to a voltage tuning rate of -2.3 cm^{-1}/V per stage or -0.4·e per stage. This value, however, is more than one order of magnitude larger than the calculated value of -0.02·e per stage. The discrepancy is possible due to the extreme non-uniformity of the electric field in the active region[165].

Lasing wavelengths of six other devices with ridge width ranging from 15 to 40 μm were also examined at 80 K. It was found that they have similar red-shift wavelengths with a tuning coefficient of ~0.13 cm^{-1}/ (A/cm^2).

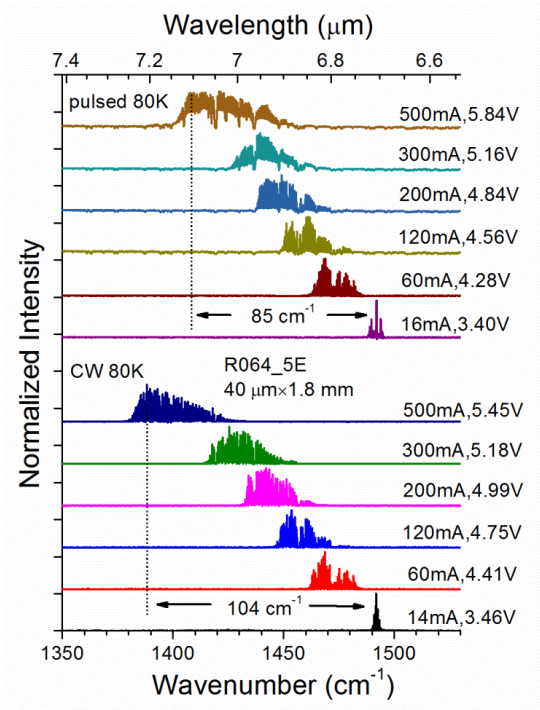


Figure 7-6. Tunable lasing spectra of an IC laser in both CW and pulsed modes.

Another wafer R094, with a nominally identical laser structure to R64, was grown separately after more than one year to verify the reproducibility. Both narrow-ridge and broad-area devices made from wafer R094 were also investigated. These devices had threshold current densities of 10~14 A/cm^2 at 80 K, similar to devices from R64. They all exhibited substantial red-shift wavelength with increasing injection current in both CW and pulsed modes. For example, a 20-μm-wide and 3-mm-long IC laser at 80 K had a lasing wavelength redshift of 132 cm^{-1} in CW mode and 99 cm^{-1} in the pulsed mode when current was increased from 8 mA to 450 mA.

The tuning rates and ranges are summarized in Table 7-1. There are two units are tuning rates. While Stark effect is better represented by the unit of cm^{-1}/(A/cm^2), the heating effect is better represented by the unit of cm^{-1}/mA.

Table 7-1. Tuning ranges and rates at 80 K

Wafer	Active region Thickness in Å InAs/GaInSb/InAs (/GaInSb/InAs)	λ in μm at 80 K	Device size		cw/ pulse	Tuning range cm^{-1}	Tuning rate $cm^{-1}/(A/cm^2)$	Tuning rate cm^{-1}/mA	Maximum current mA
			width μm	length mm					
R064	28/33/25/**36/21**	6.67	15	1.8	cw	180	-0.13	-0.49	360
			40	1.8	cw	100	-0.16	-0.22	500
					pulse	85	-0.10	-0.14	500
R094		6.89	20	3.0	cw	130	-0.20	-0.34	400
					pulse	99	-0.13	-0.22	450

In summary, electrically tunable interband cascade lasers were demonstrated with a wide tuning span about 180 cm^{-1} (22 meV or 900 nm) near 7 μm. The wide wavelength tuning was achieved by making the red-shift Stark effect that's consistent with the heating effect of current injection.

Chapter 8 InAs-based single-mode distributed feedback interband cascade lasers

InAs-based single-mode distributed feedback (DFB) interband cascade (IC) lasers were investigated. The DFB grating was patterned using interference lithography to etch through the thin top cladding into the top spacer layer of the IC laser structure. Single-mode emission with a side mode suppression ratio of ~30 dB was obtained in CW operation at temperatures up to 180 K near 4.5 μm. A total tuning range of 16 nm was achieved for a single device, with a temperature rate of 0.4 nm/K and a current-tuning rate of 0.016 nm/mA. The impact of the DFB grating on device performance was evaluated and discussed in comparison with Fabry-Perot (FP) lasers.

8.1 Introduction

Interband cascade lasers have emerged as efficient mid-infrared (IR) laser sources with low power consumption in the wavelength range from 3 to 6 μm. For emission beyond 6 μm, IC lasers have been developed on InAs substrates instead of GaSb substrates that have been used for shorter wavelengths. By replacing the InAs/AlSb superlattice (SL) cladding layers with highly-doped n^+-InAs layers, InAs-based IC lasers can have improved optical confinement and thermal dissipation, which has enabled continuous wave (CW) operation at wavelengths as long as 11 μm [68]. InAs-based IC lasers can also cover the 3-6 μm wavelength region for the detection of significant molecules such as CO_2 (4.3 μm), OCS (4.5 μm), CO (4.6 μm), N_2O (4.8 μm), and NO (5.2 μm). For sensitive detection, single-mode IC lasers are highly desirable in tunable laser absorption spectroscopy [60]. In this work, we report the first

demonstration of InAs-based single-mode DFB IC lasers, which are tunable over 16 nm near 4.5 μm.

8.2 Laser structure and fabrication

The laser structure comprises 8 cascade stages and was grown in a molecular beam epitaxy (MBE) system on a n^+-type InAs (001) epi-ready substrate (wafer R087). The growth rate of InAs for this IC laser structure was 0.22 ML/s and 0.66 ML/s for cascade stages and other regions, respectively [105]. The cascade structure is similar to Figure 5-5 (a). The waveguide features in asymmetric claddings [161]. The upper portion of the waveguide consists of a 35-nm-thick n^+-InAs top contact layer (doped with Si to 1.0×10^{19} cm^{-3}) and a 1.5-μm-thick top InAs spacer layer just above the cascade region. The thin top cladding layer was capped by a top dielectric layer (SiO_2 or Si_3N_4 layer) and a metal contact layer for current injection. The preliminary research suggested that the device performance was comparable to IC lasers with symmetric cladding structure [69]. As a result, this thin top cladding simplifies both the growth and the fabrication of a top DFB grating with a large coupling coefficient.

Part of the wafer was first processed into broad-area (150-μm-wide) mesa FP devices (without a DFB grating) by contact lithography and wet chemical etching. The etching depth reached the bottom spacer to prevent lateral current spreading. A 200-nm-thick, 90-μm-wide SiO_2 insulating layer was deposited on the center of the mesa stripe, followed by the sputter deposition of a 30/300-nm-thick and 100-μm-wide Ti/Au layer for current injection. The laser geometry is the same as that in Figure 3 of reference [68]. The processed wafer was cleaved into laser bars with cavity lengths of 1 to 2 mm,

the facets were left uncoated, and the devices were mounted epi-side-up. The laser bars were placed on the cold finger of a cryostat for measurements in CW and pulsed modes. A representative 2-mm-long device worked at temperatures up to 231 K in CW mode near 4.9 μm (~4.3 μm at 80 K) and 320 K in pulsed mode near 5.0 μm. This result will be compared to our DFB lasers later. At 300 K, this device had a threshold current density of 600 A/cm^2 with a voltage efficiency of about 70 %. These characteristics indicate that the material quality is good enough for DFB lasers.

After the F-P devices had been evaluated, two pieces of the wafer were processed separately into DFB lasers. The DFB grating period Λ was designed according to the standard first-order grating equation:

$$\lambda = 2\Lambda N_{eff} \tag{8.1}$$

where N_{eff} is the effective index of the laser waveguide and λ is the lasing wavelength. The FP laser mentioned above had a temperature tuning rate of ~3.2 nm/K, which is mainly determined by the temperature dependence of the bandgap of InAs. In contrast, the wavelength tuning of a DFB laser is dominated by the less temperature sensitive effective index ($dN_{eff}/dT=2\text{-}3\times10^{-4}$). This means, a DFB laser only operates in a relatively narrow temperature window with single mode output. In order to have a reasonable tuning range, we chose an intermediate CW operating temperature of 160 K and the corresponding target wavelength of 4.45 μm. Our waveguide simulation gave an effective index of ~3.4 for our laser structure. Using equation (8.1), the designed grating period was 655 nm. For simplification of the device fabrication in this initial attempt, gratings were only covered by metal without a dielectric layer.

Our calculation gave a coupling coefficient κ of ~12 cm^{-1} for a rectangular grating with a depth of 100 nm and a duty cycle of 50%. If the cavity length is chosen to be 1 mm, the DFB laser will have a decent coupling strength (*i.e.*, $|\kappa L| > 1$). Note that in this case both index coupling and loss coupling exist because of the large imaginary part of the refractive index for a metal. The calculation detail for coupling coefficient is given in Appendix D.

The DFB gratings were defined using interference lithography and Cl_2 reactive ion etching (RIE) to etch the grating into the top InAs spacer layer. These DFB samples were processed into broad-area lasers by contact photolithography and wet chemical etching. Without a dielectric layer on the top, 30/300 nm-thick Ti/Au layers were directly deposited on the top grating for complex coupling as well as electrical injection. After substrate lapping and deposition of metal back contact, the wafers were cleaved into laser bars with cavity lengths of 1.0 to 2.0 mm without facet coatings. These DFB lasers were mounted epi-side-up on a Cu-heatsink for measurement. SEM images of typical lasers with magnifications in critical regions are shown in Figure 8-1. As shown in Figures 8-1(a) and (d), the top metal strips were narrower than the mesa width to prevent the metal contact layer from covering to the sidewall. However, the majority of the grating regions are covered by these metal layers to ensure sufficient complex index coupling. As seen in Figure 8-1(b), the grating stripes are essentially periodic, although there are some variations and non-uniformity. In Figure 8-1(c), 8 cascade stages can be identified below the top spacer layer. The surface DFB gratings have a depth of about 100 nm and a duty cycle of about 35%. The grating period is 595±5 nm for the first and 640±5 nm for the second laser samples, respectively.

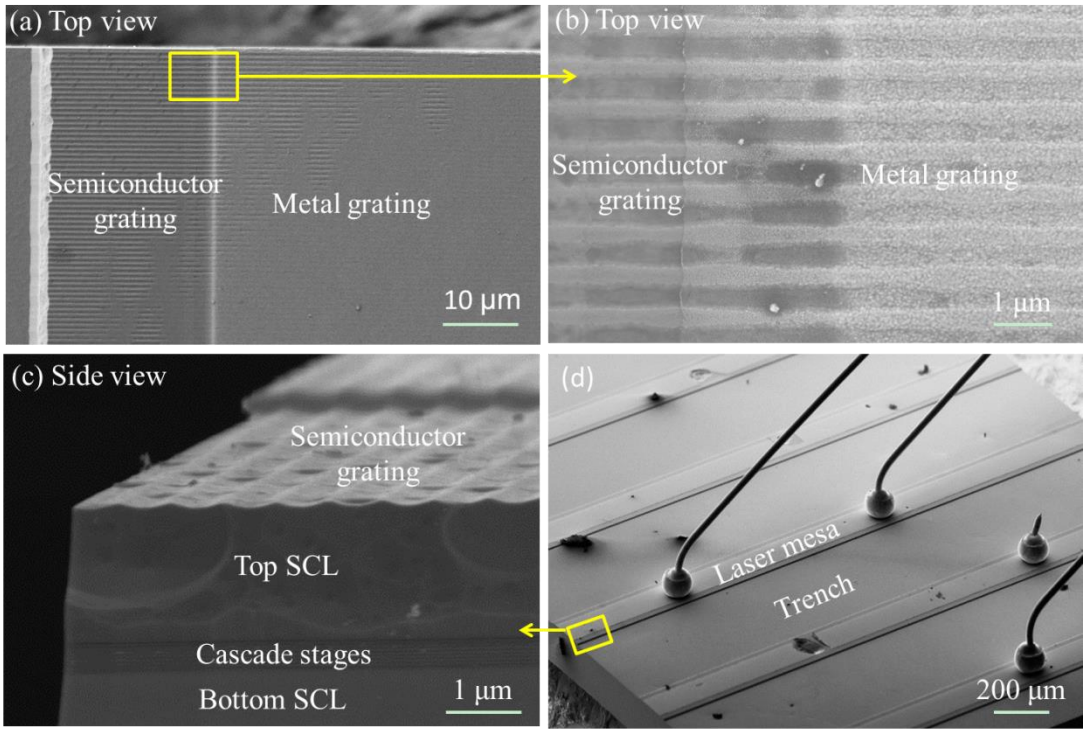


Figure 8-1. SEM images of IC laser structures. (a) a top view of the DFB gratings covered with metal and without metal; (b) zoom-in of the DFB grating; (c) cross-sectional view of a laser; (d) the top view of a laser mesa

8.3 Experimental results and discussion

Lasers from the first DFB sample (grating period of ~ 595 nm) had a grating period that was significantly shorter than the designed value (655 nm). As expected, most of these lasers didn't achieve single-mode operation, because the wavelength determined by DFB grating was not aligned with the gain peak (see Figure 8-2 (a) for device C). Two emission peaks separated by about 140 nm could be identified in the sub-threshold spectra. The two distinctive dips around 4.3 μm are due to the strong absorption of CO_2. Nevertheless because of material non-uniformity and/or variations of the DFB grating period, a few lasers did exhibit single-mode operation at 80 K. The gain peak for these devices happened to be close to the wavelength of the DFB mode, as

shown in Figure 8-2(b) for device H. The electroluminescent intensity for device H increased with current at a much slower rate than that for device C (without lasing at the DFB wavelength), which resulted in a higher threshold current density (7 A/cm^2 vs. 4 A/cm^2). The differences in lasing wavelength and threshold current density between device C and H in a wide temperature range are shown later in Figure 8-7.

Nevertheless, when device H lased at 80 K, the side modes around the DFB mode (4.204 μm) were suppressed with a side mode suppression ratio (SMSR) of about 30 dB (see inset in Figure 8-2 (b)). For device H, the single DFB-mode behavior was observed at 90 K with an SMSR larger than 20 dB, which existed simultaneously with FP modes at a relatively high current. By equation (1), the effective index is 3.53±0.04 at 80 K. The extracted effective index is greater than the value calculated for this laser layer structure. Currently, this is not understood: it may be related to the complex refractive index in metals and/or variations and non-uniformity of the DFB gratings.

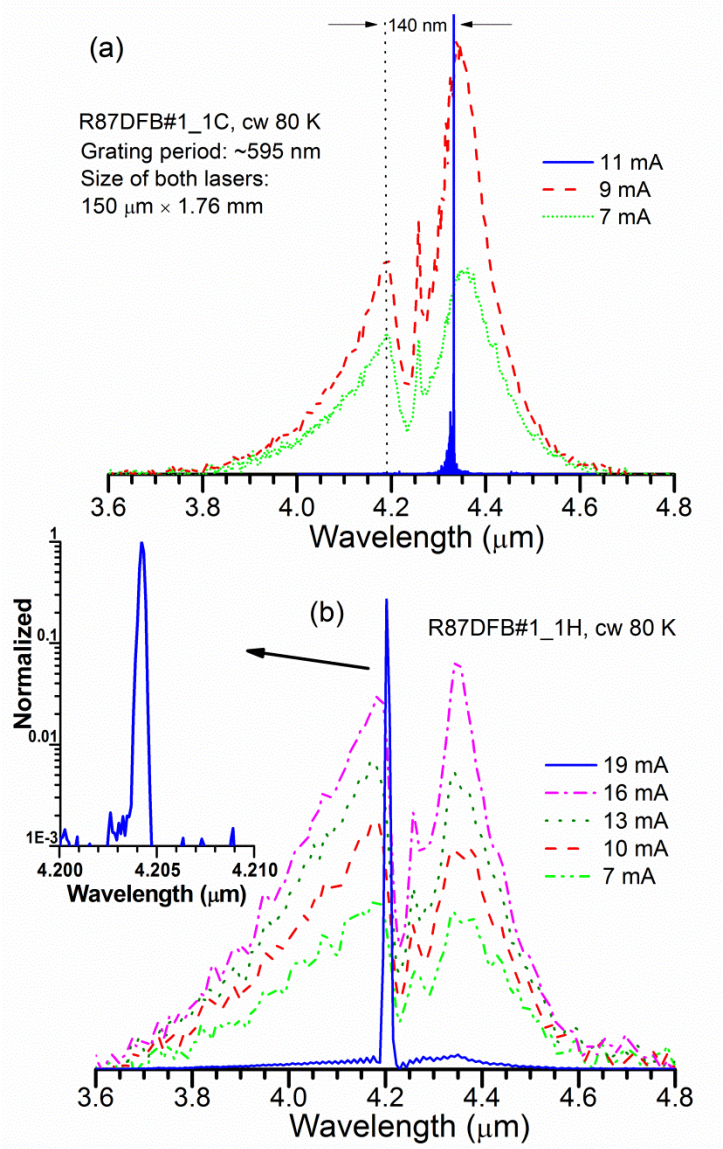


Figure 8-2. Emission spectra from two IC lasers at 80 K for several currents using (a) device C; and (b) device H, with a log intensity plot inset.

The devices from the second DFB processing (grating period ~640 nm) exhibited robust single-mode lasing at higher temperatures. The sub-threshold emission spectra for a 1-mm-long device at several temperatures (see Fig. 8-3) indicate that the laser gain spectrum is well aligned with the DFB mode at a wavelength of 4.48 μm at temperatures above 150 K. Above threshold, mode-hop-free single-frequency lasing emission was observed in a temperature range of 150 to 180 K. Figure 8-4 shows

typical current-tuning spectra for this device at 160 K, with an SMSR of about 30 dB (inset Figure 8-4) and a current-tuning rate of 16 nm/A in the vicinity of 4.49 μm. The tuning range was about 2 nm when the current was raised from 140 to 260 mA at 160 K. This tuning is mainly attributed to Joule heating from the current injection. Hence, the tuning range can be expanded with more current injection.

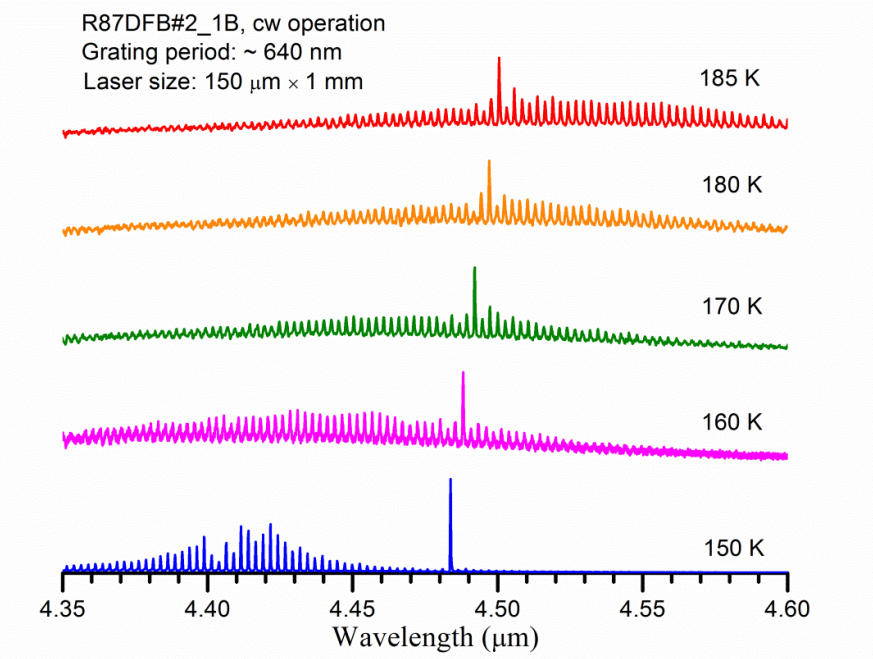


Figure 8-3. Sub-threshold emission spectra from a BA DFB IC laser with a grating period of ~640-nm at heat-sink temperatures ranging from 150 to 180 K.

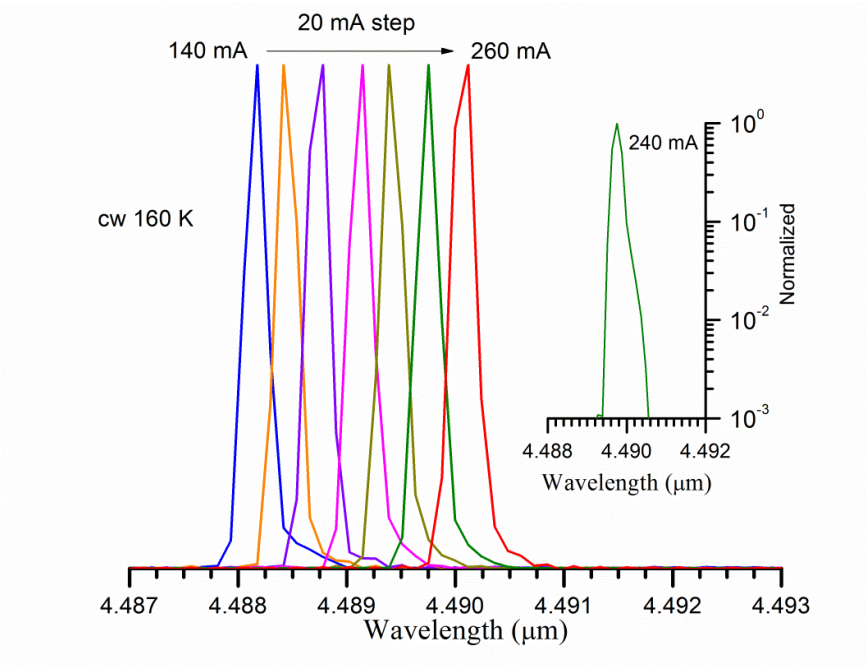


Figure 8-4. Current-tuning CW spectra of a DFB IC laser at 160 K, with an SMSR of ~ 30 dB. Inset is a plot of intensity with log scale.

Figure 8-5 shows the single-mode lasing wavelengths measured for the same laser operating at various heat-sink temperatures and injection currents. The extracted average temperature-tuning rate of the DFB wavelengths was 0.4 nm/K. Together with the current tuning, a total tuning range of 16 nm was achieved for a single device, which is sufficient for covering a specific absorption spectral region of a gas molecule. The current-voltage-light characteristics in the corresponding temperature range are shown in Figure 8-6. The output power emitted from one facet of the device was larger than 1 mW at 180 K near 4.50 μm, which can satisfy the requirement for in situ sensing. Note that the collected power was only calibrated for the 10% transmission loss from the CaF_2 window of the cryostat, without considering the beam divergence. Hence, the actual output power should be higher than the reported values in Figure 8-6. The inset to Figure 8-6 shows the lasing spectra of the DFB mode at temperatures from 150 to

180 K. The extracted effective index is 3.54±0.04 (with a temperature coefficient of $dN_{eff}/dT = 3.6 \times 10^{-4}$), which is again higher than our previous estimate based on the laser layer structure. These experimental values will help improve the accuracy in our future design.

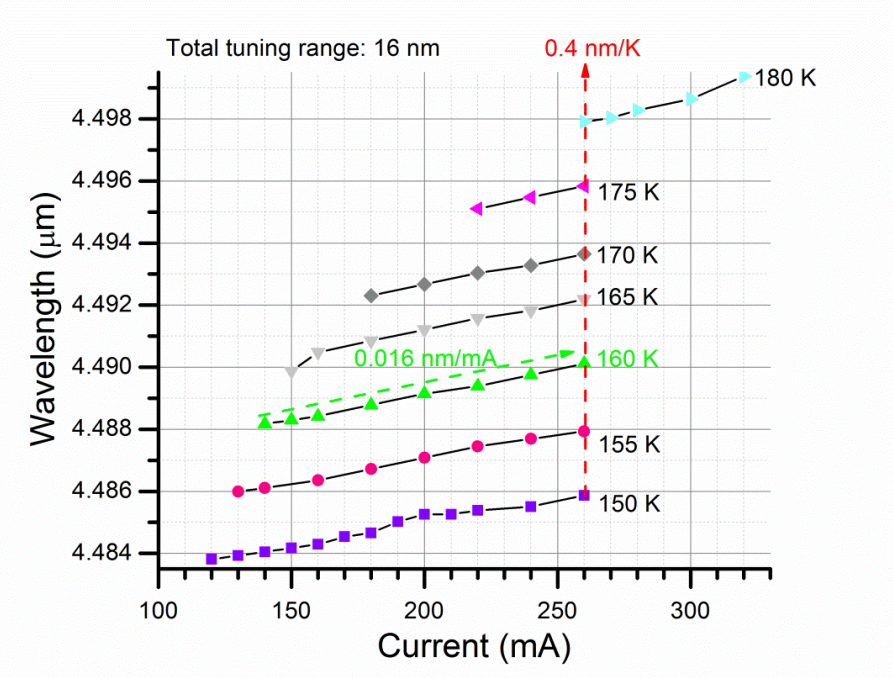


Figure 8-5. Single-mode lasing wavelengths of a DFB IC laser as a function of injected currents at T=150-180 K. A total tuning range of 16 nm was achieved for a single device. Only data with spectral SMSR larger than 20 dB are included.

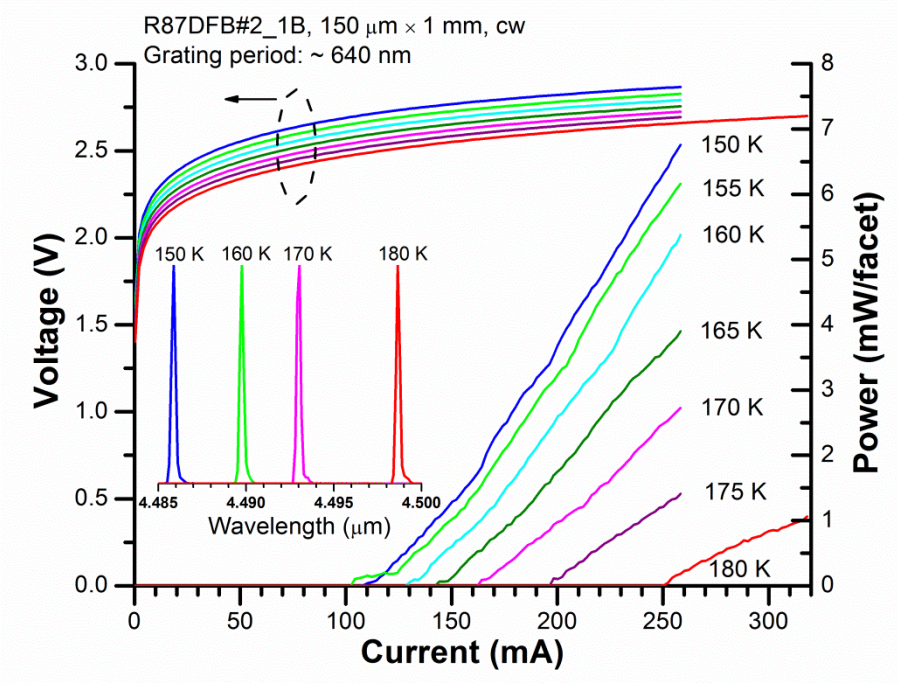


Figure 8-6. Current-voltage-light characteristics for a DFB IC laser. Inset: CW spectra at temperatures from 150 to 180 K.

Above 185 K, the DFB mode became detuned away from the gain spectral peak, and the laser operation was based on FP modes. If the DFB grating is designed to aim at a longer wavelength (corresponding to the higher temperature) and narrow-ridge lasers are fabricated from this same wafer, CW operation based on the DFB mode near 5 μm would be achievable at ~ 270 K. The expected device performance from this InAs-based wafer is still no better than the state-of-the-art GaSb-based DFB IC lasers. However, these results are helpful to assess how InAs-based DFB lasers can be further developed.

For the evaluation of how DFB gratings can affect the device performance, InAs-based IC lasers with and without DFB gratings are compared in terms of their temperature dependent lasing wavelengths and threshold current densities in CW mode, as shown in Figure 8. Because the longer cavity length will lead to less mirror loss and

thus lower threshold current density, the FP lasers with two different lengths are plotted as reference lasers (open symbols). As shown in Figure 8, the devices that lased with DFB modes (solid symbols) had a higher threshold current densities than the FP lasers. The laser with the widest DFB tuning range had the highest threshold current density. This suggests that the optical coupling with the grating contributed significant extra loss, or the differential gain at the DFB wavelength is smaller, or a combination of both factors. Also shown in Figure 8-7, a laser with the grating but no lasing at the DFB mode (half-filled symbol), had a similar threshold current density as the FP laser of a similar length. This implies that the metal grating itself might not have a substantial effect on the FP laser performance, which is worth further exploration. A way to further eliminate the possible loss is to insert a dielectric layer between the grating and the metal, which will be investigated in our future effort.

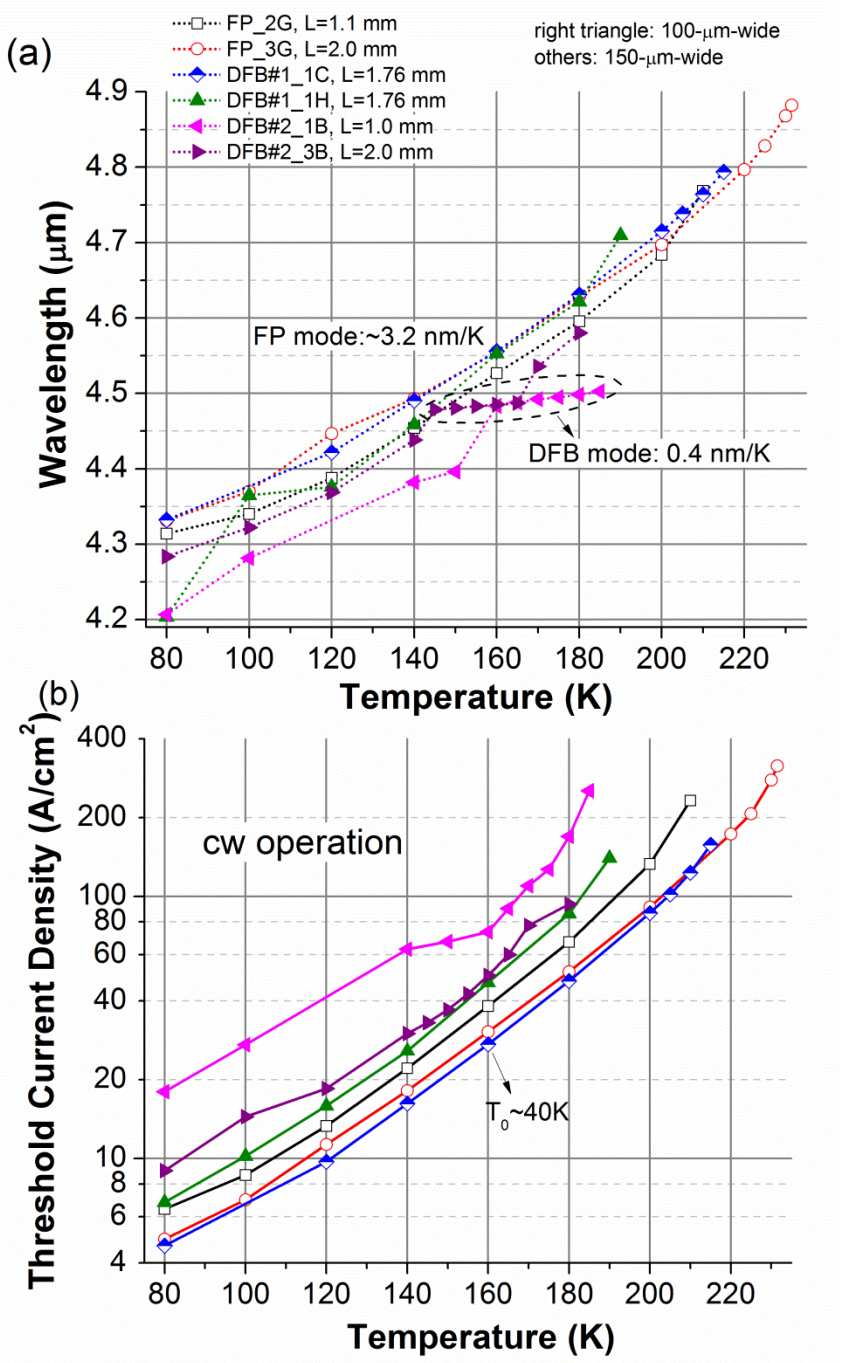


Figure 8-7. Comparison of FP and DFB lasers at various operating temperatures: (a) CW lasing wavelengths and (b) threshold current densities

Chapter 9 Conclusions and future work

9.1 Dissertation summary

The topic of this dissertation is to achieve high performance for InAs-based IC lasers. This subject is broken into several parts, including far-field patterns, high-temperature operation, long-wavelength operation, wide tunability and single frequency mode operation.

Double-lobe far-field pattern was observed on a single-plasmon waveguide IC laser, which was attributed to the first-order lateral mode. By placing metal contacts near the edges of the ridge, the higher-order spatial mode can be suppressed, resulting in a better beam quality.

A series of design innovations, including "shortened injector" and "carrier rebalancing," have brought down the threshold current densities of IC lasers. The double cladding waveguide, consisting of superlattice and highly doped InAs, has significantly improved the optical confinement of InAs-based IC lasers. CW operations above room temperature have been achieved at wavelengths of 4.6~4.8 μm. The threshold current density, 247 A/cm^2 at 300 K in pulsed mode, is the lowest ever reported among semiconductor mid-infrared lasers at similar wavelengths. The pulsed operating temperature is as high as 377 K.

With the hybrid waveguide mentioned above, the lasing wavelength reaches 6.4 μm at 324 K. While a large waveguide loss at 11 μm is because the lasing wavelength overlaps with the plasmon wavelength for the heavily-doped InAs, the unexpected high

threshold current density may be related to the negative differential resistance discovered from the current-voltage characteristics.

By increasing the number of InAs active QWs to three, a large repeatable electrical tunable range of 180 cm^{-1} (or 900 nm in wavelength) are achieved near $\lambda \sim 7$ μm. Significant tuning by Stark effect requires a substantial change of the applied electrical field, which is usually inhibited by the carrier density pinned at the threshold level. However, by trading the modal gain, the carrier pinning is loosened. As a result, the electrical field after threshold continuously increased and significant tuning was achieved.

The InAs-based IC lasers with the asymmetric cladding layers are readily suitable for the fabrication of top DFB grating which would have a strong coupling with the optical wave. The grating was patterned using interference lithography to etch directly into the top spacer layer. Single-mode emission with a side mode suppression ratio of 30 dB is obtained in continuous wave operation at temperatures up to 180 K near 4.5 μm. A total tuning range of 16 nm could be achieved for a single device, with typical temperature- and current-tuning rates of 0.4 nm/K and 0.016 nm/mA, respectively. However, the grating had introduced quite a lot of loss and degraded device performance compared with Fabry–Perot lasers. A further optimization of the structure and a better grating fabrication are required to improve the DFB laser performance.

9.2 Future work

InAs-based IC lasers have been developed as a complementary counterpart of GaSb-based IC lasers at a longer wavelength. Up to now, the overall performance of InAs-based IC lasers is still lagging behind, which is due to the increased optical loss at longer wavelengths, reduced optical confinement, less optimized laser structure and less uniform material quality grown on the InAs substrate.

Although significant advancements have been made by several research groups in recent years, there is ample room for design improvement. Practically, it's impossible to exhaust all the design possibilities. Instead, inspiration comes from a better understanding of the fundamental physics behind a working laser. Such understanding and derived models also require prompt feedback from reliable and repeatable experimental results.

Specifically, I have several ideas for future work.

1st, the injection region can be further optimized. The concept of "carrier rebalancing" [51] should be revisited. The authors' latest claim is still that the hole density will outnumber the electron density in the active region if no extrinsic doping is introduced [34]. However, the scanning voltage microscopy (SVM) measurements [204] showed the electron and hole carriers are always approximately balanced, no matter the electron injector is heavily-doped (as in wafer R84) or lightly-doped (as in wafer R94). Furthermore, SVM measurements found no net charge distribution in the electron injector in both cases. This means that the ionized electrons are still bound to their donors without participating in the active region.

In my understanding, these ionized electrons, provided by the heavily-doped InAs, play an essential role in the Bernard-Duraffourg inversion, as in equation (2.2.5). These electron carriers help to elevate the common quasi-Fermi level [34] on both sides of the semi-metallic interface. The real effects of the heavily-doped injection are: (1) decrease the transparency carrier density, as shown in Figure 5-3; (2) smooth the carrier transport process from the semi-metallic interface to the active region. A comprehensive model, considering Poisson equation, rate equation, and Schrödinger equation, should be able to describe these two effects quantitatively and obtain the optimal doping concentration for the different injector. A set of design structures is to be grown accordingly to validate this idea.

2nd, the active region can be further optimized. A drawback of the type-II active QWs is the reduced wavefunction overlap, which leads to reduced modal gain. This becomes more severe at a longer wavelength because the InAs QWs are made wider to decrease the eigenenergy. One way to compensate the reduced modal gain is to increase the number of cascade stages, which, however, also increase the thermal resistance that is detrimental to the CW operation. Another way is to change the composition of GaInSb hole QW. It's still not conclusive whether increasing the indium content would reduce the overlap as suggested by [98]. From the perspective of the strain effect, the increased indium content will introduce more compressive strain and reduce the effective mass of heavy hole, which partially alleviate the asymmetry of between very light conduction band mass and the heavy valence band mass. Furthermore, a quaternary material is probably a better substitute for GaInSb [98].

3rd, the mechanism of the negative differential resistance in long-wavelength IC lasers is not year clear. As suggested by Figure 5-3, the material gain is independent of the lasing wavelength. So the dramatically increased threshold current is a combined effect of increased optical loss, reduced carrier lifetime and reduced wavefunction overlap. The material is inversely linear dependent on the temperature. So at low temperature, all the IC lasers at different wavelengths should have very similar threshold current densities. This is verified by the experimental results that even the λ~9 μm laser with a size of 20 μm×1 mm had a threshold current density as low as 10 A/cm^2 at 80 K [180]. However, the λ~11 μm laser had a threshold current density as high as 72 A/cm^2 at 80 K, as shown in Figure 6-4. Although the closeness of the lasing wavelength and the plasmon wavelength due to heavy doping brought significant free-carrier absorption loss, another unknown mechanism, related to the negative differential resistance as shown in Figure 6-5, could be the culprit. Similar negative differential resistance had also been observed for devices at even longer wavelengths. A more systematical study to identify the possible mechanism would give some hints on the design of long-wavelength IC lasers.

4th, there's no systematical study on carrier leakage in the cascade stage. For QC lasers based on the inter-subband transition, carrier leakage is a severe problem due to insufficient band offset. For IC lasers, carrier leakage is still possible because AlSb is an indirect bandgap material. The conduction band electrons, due to thermal spread or high injection, will gain enough energy to leak into the satellite valley of AlSb. In addition, the hole leakage into the AlSb valence band had been discovered by SVM [165]. Furthermore, a recent study by NRL yielded internal efficiency of 70%~90% at

room temperature, which they expected the value to be close to 100% [108]. In my understanding, this is because they didn't account the "injection efficiency" degraded by the carrier leakage. What's more, their shorter-wavelength laser performed worse, which is also possibly due the carrier leakage.

5th, a semi-empirical model for the temperature dependence of the threshold current density is both scientifically interesting and technologically useful. The threshold current densities of most interband lasers are known to be sensitive to the operation temperature, which can be fitted into an exponential relation since 1968 [205]:

$$J_{th}(T) = J_0 \exp(\frac{T}{T_0}) \tag{9.1}$$

where T_0 is the characteristic temperature, and J_0 is the threshold current density extrapolated to T=0 K.

It's quite surprising that such a simple empirical relation is valid through a wide range of temperature for various interband lasers, although T_0 is not a constant when the temperature approaches the maximum operating temperature. Hence, the characteristic temperature T_0 are routinely reported in the literature.

Early research on GaAs-based InGaAs QW lasers at 0.86 μm reported an extremely high T_0 value of 437 C [206], which was overinflated due to its high threshold current density of 3 kA/cm^2. T_0 quickly dropped to 200 K for a λ~1.17 μm InGaAs QW laser with a J_{th} of 130 A/cm^2 [207]. For λ~1.5 μm lasers on different substrates, T_0 all dived below 100 K [208-210]. For λ=3~5 μm state-of-the-art IC lasers, T_0 is in the range of 45-60 K [34,73,211]. These statistics seem to imply some common mechanism.

Early researchers tried to relate this strong temperature dependence of threshold to the temperature-dependent shift of the Fermi level [212] [213] and density of states [214], or the band tailing effect on the differential gain [215,216]. Later, the temperature dependence of threshold current in different types of lasers have been extensively investigated [208,217-226]. However, due to various detailed scenarios, these individual case studies are not readily applicable to a more general case. So far, the low T_0 value is not quantitatively explained.

Figure 9-1 shows a preliminary model of temperature-dependent threshold current, which is able to predict reasonable T_0 values. Basically, the gain coefficient is calculated from the Fermi's Golden rule. The Auger coefficient and temperature dependent optical loss are empirical values from the literature. The details of this model require more refinements and deserve a future publication. Anyway, a generalized model of $J_{th}(T)$ is of vital importance to the interband laser community.

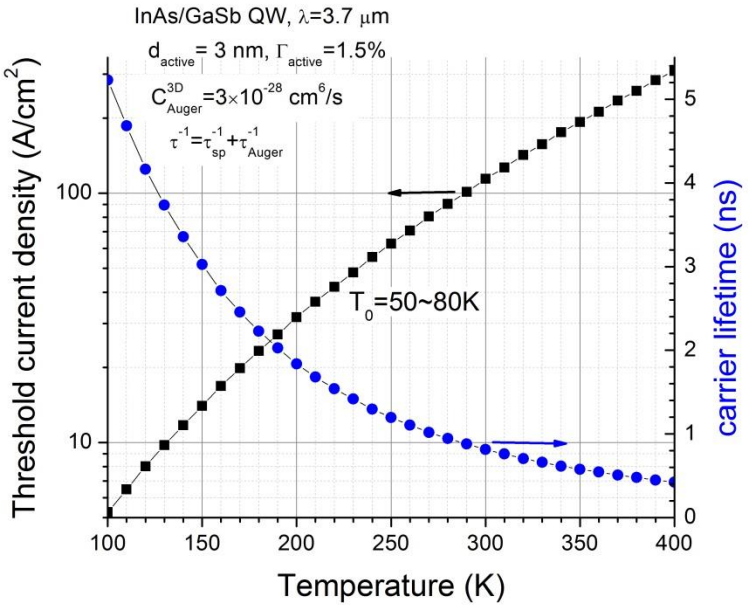


Figure 9-1.Calculated temperature dependency of the threshold current density

References

[1] http://www.nobelprize.org/nobel_prizes/physics/laureates/1964/

[2] J. P. Gordon, H. J. Zeiger, and C. H. Townes, "Molecular Microwave Oscillator and New Hyperfine Structure in the Microwave Spectrum of NH3", Phys. Rev., **95**, 282 (1954).

[3] J. P. Gordon, H. J. Zeiger, and C. H. Townes, "The Maser--New Type of Microwave Amplifier, Frequency Standard, and Spectrometer", Phys. Rev., **99**, 1264 (1955).

[4] A. Einstein, "On the quantum theory of radiation", Physikalische Zeitschrift, **18**, 121 (1917).

[5] Y.-M. Mu and R. Q. Yang, "Modeling of InAs/GaInSb/AlSb type-II mid-IR interband cascade lasers", SPIE,**3625**, 811 (1999)

[6] T. H. Maiman, "Stimulated optical radiation in ruby", Nature, **187**, 493 (1960).

[7] A. L. Schawlow and C. H. Townes, "Infrared and Optical Masers", Phys. Rev., **112**, 1940 (1958).

[8] H. Kroemer, "A proposed class of hetero-junction injection lasers", Proceedings of the IEEE, **51**, 1782 (1963).

[9] Z. I. Alferov and R. F. Kazarinov, "Semiconductor laser with electric pumping", Russian Patent No. 181737 (1963).

[10] Z. I. Alferov, V. M. Andreev, D. Z. Garbuzov, Y. V. Zhilyaev, E. P. Morozov, E. L. Portnoi, and V. G. Trofim, "Investigation of the influence of the AlAs/GaAs heterostructure parameters on the laser threshold and the realization of continuous emission at the room temperature", Fiz. Tekh. Poluprovodn., **4**, 1826 (1970).

[11] I. Hayashi, M. B. Panish, P. W. Foy, and S. Sumski, "Junction lasers which operate continuously at room temperature", Appl. Phys. Lett., **17**, 109 (1970).

[12] R. D. Dupuis, P. D. Dapkus, N. Holonyak, E. A. Rezek, and R. Chin, "Room - temperature laser operation of quantum - well $Ga_{(1-x)}Al_xAs$ - GaAs laser diodes grown by metalorganic chemical vapor deposition", Appl. Phys. Lett., **32**, 295 (1978).

[13] W. T. Tsang, "A graded - index waveguide separate - confinement laser with very low threshold and a narrow Gaussian beam", Appl. Phys. Lett., **39**, 134 (1981).

[14] W. T. Tsang, "Extremely low threshold (AlGa)As graded - index waveguide separate - confinement heterostructure lasers grown by molecular beam epitaxy", Appl. Phys. Lett., **40**, 217 (1982).

[15] H. K. Choi and C. A. Wang, "InGaAs/AlGaAs strained single quantum well diode lasers with extremely low threshold current density and high efficiency", Appl. Phys. Lett., **57**, 321 (1990).

[16] R. L. Williams, M. Dion, F. Chatenoud, and K. Dzurko, "Extremely low threshold current strained InGaAs/AlGaAs lasers by molecular beam epitaxy", Appl. Phys. Lett., **58**, 1816 (1991).

[17] E. Yablonovitch and E. O. Kane, "Reduction of lasing threshold current density by the lowering of valence band effective mass", Lightwave Technology, Journal of, **4**, 504 (1986).

[18] 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).

[19] C. K. N. Patel, "Laser based in-situ and standoff detection of chemical warfare agents and explosives", Proc. SPIE,**7484**, 748402 (2009)

[20] M. Taslakov, V. Simeonov, and H. van den Bergh, "Line-of-sight data transmission system based on mid IR quantum cascade laser", Proc. SPIE,**6877**, 68770F (2008)

[21] http://atoptics.com

[22] M. Tacke, *Lead–salt lasers*. (2001), pp.547.

[23] R. Klann, T. Höfer, R. Buhleier, T. Elsaesser, and J. W. Tomm, "Fast recombination processes in lead chalcogenide semiconductors studied via transient optical nonlinearities", J. Appl. Phys., **77**, 277 (1995).

[24] P. C. Findlay, C. R. Pidgeon, R. Kotitschke, A. Hollingworth, B. N. Murdin, C. J. G. M. Langerak, A. F. G. van der Meer, C. M. Ciesla, *et al.*, "Auger recombination dynamics of lead salts under picosecond free-electron-laser excitation", Phys. Rev. B, **58**, 12908 (1998).

[25] I. Vurgaftman, W. W. Bewley, C. L. Canedy, K. Chul Soo, K. Mijin, C. D. Merritt, J. Abell, and J. R. Meyer, "Interband Cascade Lasers With Low Threshold Powers and High Output Powers", IEEE J. Select. Topics Quantum Electron., **19**, 1200210 (2013).

[26] M. Eibelhuber, T. Schwarzl, S. Pichler, W. Heiss, and G. Springholz, "Near room temperature continuous-wave laser operation from type-I interband transitions at wavelengths beyond 4 mu m", Appl. Phys. Lett., **97**, 061103 (2010).

[27] F. Zhao, H. Wu, A. Majumdar, and Z. Shi, "Continuous wave optically pumped lead-salt mid-infrared quantum-well vertical-cavity surface-emitting lasers", Appl. Phys. Lett., **83**, 5133 (2003).

[28] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum Cascade Laser", Science, **264**, 553 (1994).

[29] J. Faist, M. Beck, T. Aellen, and E. Gini, "Quantum-cascade lasers based on a bound-to-continuum transition", Appl. Phys. Lett., **78**, 147 (2001).

[30] D. Hofstetter, M. Beck, T. Aellen, and J. Faist, "High-temperature operation of distributed feedback quantum-cascade lasers at 5.3 μm", Appl. Phys. Lett., **78**, 396 (2001).

[31] Y. Bai, N. Bandyopadhyay, S. Tsao, S. Slivken, and M. Razeghi, "Room temperature quantum cascade lasers with 27% wall plug efficiency", Appl. Phys. Lett., **98**, 181102 (2011).

[32] J. Faist, "Continuous-Wave, Room-Temperature Quantum Cascade Lasers", Opt. Photon. News, **17**, 32 (2006).

[33] R. Q. Yang, "Interband cascade (IC) lasers", Chap. 12, in *Semiconductor Lasers: fundamentals and applications*, edited by Alexei Baranov and Eric Tournié (Woodhead Publishing, Cambridge, UK, 2013); and references therein.

[34] I. Vurgaftman, R. Weih, M. Kamp, J. R. Meyer, C. L. Canedy, C. S. Kim, M. Kim, W. W. Bewley, *et al.*, "Interband cascade lasers", J. Phys. D-Appl. Phys., **48**, 123001 (2015).

[35] R. Q. Yang and S. S. Pei, "Novel type-II quantum cascade lasers", J. Appl. Phys., **79**, 8197 (1996).

[36] J. P. van der Ziel and W. T. Tsang, "Integrated multilayer GaAs lasers separated by tunnel junctions", Appl. Phys. Lett., **41**, 499 (1982).

[37] J. C. Garcia, E. Rosencher, P. Collot, N. Laurent, J. L. Guyaux, B. Vinter, and J. Nagle, "Epitaxially stacked lasers with Esaki junctions: A bipolar cascade laser", Appl. Phys. Lett., **71**, 3752 (1997).

[38] S. G. Patterson, G. S. Petrich, R. J. Ram, and L. A. Kolodziejski, "Continuous-wave room temperature operation of bipolar cascade laser", Electron. Lett., **35**, 395 (1999).

[39] J. K. Kim, E. Hall, O. Sjölund, and L. A. Coldren, "Epitaxially-stacked multiple-active-region 1.55μm lasers for increased differential efficiency", Appl. Phys. Lett., **74**, 3251 (1999).

[40] W. J. Siskaninetz, J. E. Ehret, T. N. Dang, J. E. V. Nostrand, J. A. Lott, and T. R. Nelson, in *Electron. Lett.* (Institution of Engineering and Technology, 2002), Vol. 38, pp. 1259.

[41] R. Q. Yang and Y. Qiu, "Bipolar cascade lasers at emitting wavelengths near 2 μm", Appl. Phys. Lett., **83**, 599 (2003).

[42] F. Akyol, S. Krishnamoorthy, and S. Rajan, "Tunneling-based carrier regeneration in cascaded GaN light emitting diodes to overcome efficiency droop", Appl. Phys. Lett., **103**, 081107 (2013).

[43] R. Q. Yang, "Infrared laser based on intersubband transitions in quantum wells ", 7th International Conference on Superlattice, Microstructure and Microdevices,(1994)

[44] R. Q. Yang, "Infrared laser based on intersubband transitions in quantum wells", Superlattices Microstruct., **17**, 77 (1995).

[45] C.-H. Lin, R. Q. Yang, D. Zhang, S. J. Murry, S. S. Pei, A. A. Allerman, and S. R. Kurtz, "Type-II interband quantum cascade laser at 3.8 μm", Electron. Lett., **33**, 598 (1997).

[46] J. R. Meyer, I. Vurgaftman, R. Q. Yang, and L. R. Ram-Mohan, "Type-II and type-I interband cascade lasers", Electron. Lett., **32**, 45 (1996).

[47] J. R. Meyer, C. A. Hoffman, F. J. Bartoli, and L. R. Ram-Mohan, "Type-II quantum-well lasers for the mid-wavelength infrared", Appl. Phys. Lett., **67**, 757 (1995).

[48] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Mid-IR vertical-cavity surface-emitting lasers", IEEE J. Quantum Electron., **34**, 147 (1998).

[49] L. J. Olafsen, E. H. Aifer, I. Vurgaftman, W. W. Bewley, C. L. Felix, J. R. Meyer, D. Zhang, C. H. Lin, *et al.*, "Near-room-temperature mid-infrared interband cascade laser", Appl. Phys. Lett., **72**, 2370 (1998).

[50] A. Bauer, M. Dallner, M. Kamp, S. Hofling, L. Worschech, and A. Forchel, "Shortened injector interband cascade lasers for 3.3- to 3.6-μm emission", Opt. Eng., **49**, 111117 (2010).

[51] I. Vurgaftman, W. W. Bewley, C. L. Canedy, C. S. Kim, M. Kim, C. D. Merritt, J. Abell, J. R. Lindle, *et al.*, "Rebalancing of internally generated carriers for mid-infrared interband cascade lasers with very low power consumption", Nat. Commun., **2**, 585 (2011).

[52] R. Q. Yang, J. L. Bradshaw, J. D. Bruno, J. T. Pham, D. E. Wortman, and R. L. Tober, "Room temperature type-II interband cascade laser", Appl. Phys. Lett., **81**, 397 (2002).

[53] C. J. Hill, B. Yang, and R. Q. Yang, "Low-threshold interband cascade lasers operating above room temperature", Physica E: Low-dimensional Systems and Nanostructures, **20**, 486 (2004).

[54] J. S. Yu, A. Evans, S. Slivken, S. R. Darvish, and M. Razeghi, "Temperature dependent characteristics of $\lambda \sim 3.8$ μm room-temperature continuous-wave quantum-cascade lasers", Appl. Phys. Lett., **88**, 251118 (2006).

[55] N. Bandyopadhyay, Y. Bai, S. Tsao, S. Nida, S. Slivken, and M. Razeghi, "Room temperature continuous wave operation of $\lambda \sim 3$--3.2 μm quantum cascade lasers", Appl. Phys. Lett., **101**, 241110 (2012).

[56] R. Q. Yang, C. J. Hill, and B. H. Yang, "High-temperature and low-threshold midinfrared interband cascade lasers", Appl. Phys. Lett., **87**, 151109 (2005).

[57] K. Mansour, Y. Qiu, C. J. Hill, A. Soibel, and R. Q. Yang, "Mid-infrared interband cascade lasers at thermoelectric cooler temperatures", Electron. Lett., **42**, 1034 (2006).

[58] R. Q. Yang, C. J. Hill, K. Mansour, Y. M. Qiu, A. Soibel, R. E. Muller, and P. M. Echternach, "Distributed feedback mid-IR interband cascade lasers at thermoelectric cooler temperatures", Ieee Journal of Selected Topics in Quantum Electronics, **13**, 1074 (2007).

[59] C. Webster, "Low Upper Limit to Methane Abundance on Mars", Science, **342**, 355 (2013).

[60] C. R. Webster, P. R. Mahaffy, S. K. Atreya, G. J. Flesch, M. A. Mischna, P.-Y. Meslin, K. A. Farley, P. G. Conrad, *et al.*, "Mars methane detection and variability at Gale crater", Science, **347**, 415 (2015).

[61] M. Kim, C. L. Canedy, W. W. Bewley, C. S. Kim, J. R. Lindle, J. Abell, I. Vurgaftman, and J. R. Meyer, "Interband cascade laser emitting at λ=3.75 μm in continuous wave above room temperature", Appl. Phys. Lett., **92**, 191110 (2008).

[62] R. Weih, M. Kamp, and S. Hofling, "Interband cascade lasers with room temperature threshold current densities below 100 A/cm^2", Appl. Phys. Lett., **102**, 231123 (2013).

[63] W. W. Bewley, C. L. Canedy, C. S. Kim, M. Kim, C. D. Merritt, J. Abell, I. Vurgaftman, and J. R. Meyer, "High-power room-temperature continuous-wave mid-infrared interband cascade lasers", Opt. Express, **20**, 20894 (2012).

[64] C. L. Canedy, J. Abell, C. D. Merritt, W. W. Bewley, C. S. Kim, M. Kim, I. Vurgaftman, and J. R. Meyer, "Pulsed and CW performance of 7-stage interband cascade lasers", Opt. Express, **22**, 7702 (2014).

[65] M. Kim, W. W. Bewley, C. L. Canedy, C. S. Kim, C. D. Merritt, J. Abell, I. Vurgaftman, and J. R. Meyer, "High-power continuous-wave interband cascade lasers with 10 active stages", Opt. Express, **23**, 9664 (2015).

[66] Z. Tian, R. Q. Yang, T. D. Mishima, M. B. Santos, R. T. Hinkey, M. E. Curtis, and M. B. Johnson, "InAs-based interband cascade lasers near 6 μm", Electron. Lett., **45**, 48 (2009).

[67] R. Q. Yang, L. Li, L. Zhao, Y. Jiang, Z. Tian, H. Ye, R. T. Hinkey, C. Niu, *et al.*, "Recent progress in development of InAs-based interband cascade lasers", Proc. SPIE, **8640**, paper 86400Q (2013).

[68] L. Li, H. Ye, Y. Jiang, R. Q. Yang, J. C. Keay, T. D. Mishima, M. B. Santos, and M. B. Johnson, "MBE-grown long-wavelength interband cascade lasers on InAs substrates", J. Cryst. Growth, **426**, 369 (2015).

[69] Z. Tian, Y. Jiang, L. Li, R. T. Hinkey, Z. Yin, R. Q. Yang, T. D. Mishima, M. B. Santos, *et al.*, "InAs-Based Mid-Infrared Interband Cascade Lasers Near 5.3 μm", IEEE J. Quantum Electron., **48**, 915 (2012).

[70] M. Dallner, S. Hofling, and M. Kamp, "Room-temperature operation of InAs-based interband-cascade-lasers beyond 6 μm", Electron. Lett., **49**, 286 (2013).

[71] M. Dallner, F. Hau, S. Höfling, and M. Kamp, "InAs-based interband-cascade-lasers emitting around 7 μm with threshold current densities below 1 kA/cm^2 at room temperature", Appl. Phys. Lett., **106**, 041108 (2015).

[72] Y. Jiang, L. Li, H. Ye, R. Yang, T. D. Mishima, M. B. Santos, and M. B. Johnson, "Continuous-wave operation of InAs-based interband cascade lasers above room temperature", CLEO: 2015 Postdeadline Paper Digest,JTh5A.8 (2015)

[73] L. Li, Y. Jiang, H. Ye, R. Q. Yang, T. D. Mishima, M. B. Santos, and M. B. Johnson, "Low-threshold InAs-based interband cascade lasers operating at high temperatures", Appl. Phys. Lett., **106**, 251102 (2015).

[74] J. T. Verdeyen, *Laser Electronics*, 3rd ed. (Prentice Hall, 1995).

[75] 郭长志, *速率方程理论*. (2009).

[76] M. G. A. Bernard and G. Duraffourg, "Laser Conditions in Semiconductors", phys. stat. sol. (b), **1**, 699 (1961).

[77] K. J. Vahala and C. E. Zah, "Effect of doping on the optical gain and the spontaneous noise enhancement factor in quantum well amplifiers and lasers studied by simple analytical expressions", Appl. Phys. Lett., **52**, 1945 (1988).

[78] 郭长志, *量子理论*. (2008).

[79] R. H. Yan, S. W. Corzine, L. A. Coldren, and I. Suemune, "Corrections to the Expression for Gain in GaAs", IEEE J. Quantum Electron., **26**, 213 (1990).

[80] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III--V compound semiconductors and their alloys", J. Appl. Phys., **89**, 5815 (2001).

[81] R. Q. Yang, L. Li, and Y. Jiang, "Interband Cascade Lasers: From Original Concept to Practical Devices", Progress in Physics, **34**, 169 (2014).

[82] L. F. Luo, R. Beresford, and W. I. Wang, "Interband tunneling in polytype GaSb/AlSb/InAs heterostructures", Appl. Phys. Lett., **55**, 2023 (1989).

[83] K. F. Longenbach, L. F. Luo, and W. I. Wang, "Resonant interband tunneling in InAs/GaSb/AlSb/InAs and GaSb/InAs/AlSb/GaSb heterostructures", Appl. Phys. Lett., **57**, 1554 (1990).

[84] H. Kitabayashi, T. Waho, and M. Yamamoto, "Resonant interband tunneling current in InAs/AlSb/GaSb/AlSb/InAs double barrier diodes", J. Appl. Phys., **84**, 1460 (1998).

[85] R. Q. Yang and J. M. Xu, "Population inversion through resonant interband tunneling", Appl. Phys. Lett., **59**, 181 (1991).

[86] H. Ohno, L. Esaki, and E. E. Mendez, "Optoelectronic devices based on type II polytype tunnel heterostructures", Appl. Phys. Lett., **60**, 3153 (1992).

[87] K. Ohtani and H. Ohno, "Intersubband electroluminescence in InAs/GaSb/AlSb type-II cascade structures", Appl. Phys. Lett., **74**, 1409 (1999).

[88] K. Faist, D. Hofstetter, M. Beck, T. Aellen, M. Rochat, and S. Blaser, "Bound-to-continuum and two-phonon resonance, quantum-cascade lasers for high duty cycle, high-temperature operation", IEEE J. Quantum Electron., **38**, 533 (2002).

[89] A. A. Allerman, R. M. Biefeld, and S. R. Kurtz, "InAsSb‐based mid‐infrared lasers (3.8‐3.9 μm) and light‐emitting diodes with AlAsSb claddings and semimetal electron injection, grown by metalorganic chemical vapor deposition", Appl. Phys. Lett., **69**, 465 (1996).

[90] R. Q. Yang, "Novel Concepts and Structures for Infrared Lasers", Chap. 2, in *Long Wavelength Infrared Emitters Based on Quantum Wells and Superlattices*, edited by M. Helm (Gordon & Breach Pub., Singapore, 2000); and references therein.

[91] E. O. Kane, "Band structure of indium antimonide", J. Phys. Chem. Solids, **1**, 249 (1957).

[92] G. Liu and S.-L. Chuang, "Modeling of Sb-based type-II quantum cascade lasers", Phys. Rev. B, **65**, 165220 (2002).

[93] S. R. White and L. J. Sham, "Electronic Properties of Flat-Band Semiconductor Heterostructures", Phys. Rev. Lett., **47**, 879 (1981).

[94] E. O. Kane, "Zener tunneling in semiconductors", J. Phys. Chem. Solids, **12**, 181 (1960).

[95] J. R. Söderström, E. T. Yu, M. K. Jackson, Y. Rajakarunanayake, and T. C. McGill, "Two‐band modeling of narrow band gap and interband tunneling devices", J. Appl. Phys., **68**, 1372 (1990).

[96] R. Q. Yang and J. M. Xu, "Bound and quasibound states in leaky quantum wells", Phys. Rev. B, **46**, 6969 (1992).

[97] R. Q. Yang and J. M. Xu, "Analysis of transmission in polytype interband tunneling heterostructures", J. Appl. Phys., **72**, 4714 (1992).

[98] F. Janiak, G. Sek, M. Motyka, K. Ryczko, J. Misiewicz, A. Bauer, S. Hofling, M. Kamp, *et al.*, "Increasing the optical transition oscillator strength in GaSb-based type II quantum wells", Appl. Phys. Lett., **100**, 231908 (2012).

[99] E. T. Yu, J. O. McCaldin, and T. C. McGill, "Band Offsets in Semiconductor Heterojunctions", Chap., in *Solid State Physics*, edited by Ehrenreich Henry and Turnbull David (Academic Press, 1992); and references therein.

[100] C. G. Van de Walle, "Band lineups and deformation potentials in the model-solid theory", Phys. Rev. B, **39**, 1871 (1989).

[101] F. Goos and H. Hanchen, "Concerning the permeation of the total reflective light in the weak medium", Annalen der Physik, **43**, 383 (1943).

[102] L. Lewin, "Obliquity-factor correction to solid-state radiation patterns", J. Appl. Phys., **46**, 2323 (1975).

[103] J. H. C. Casey, M. B. Panish, and J. L. Merz, "Beam divergence of the emission from double-heterostructure injection lasers", J. Appl. Phys., **44**, 5470 (1973).

[104] C. L. Canedy, G. I. Boishin, W. W. Bewley, C. S. Kim, I. Vurgaftman, M. Kim, J. R. Lindle, J. R. Meyer, *et al.*, "Correlating growth conditions with photoluminescence and lasing properties of mid-IR antimonide type II "W" structures", Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures, **22**, 1575 (2004).

[105] H. Ye, L. Li, R. T. Hinkey, R. Q. Yang, T. D. Mishima, J. C. Keay, M. B. Santos, and M. B. Johnson, "MBE growth optimization of InAs (001) homoepitaxy", J. Vac. Sci. Technol. B, **31**, 03C135 (2013).

[106] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multilayers: I. Misfit dislocations", J. Cryst. Growth, **27**, 118 (1974).

[107] B. R. Bennett, "MBE growth and characterization of mismatched InGaAs and InAlAs layers on InP",PhD dissertation, MIT, 1993.

[108] C. D. Merritt, W. W. Bewley, C. S. Kim, C. L. Canedy, I. Vurgaftman, J. R. Meyer, and M. Kim, "Gain and loss as a function of current density and temperature in interband cascade lasers", Appl. Opt., **54**, F1 (2015).

[109] W. Shockley, "The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors", Bell Syst. Tech. J., **28**, 435 (1949).

[110] E. F. Schubert, *Light-emitting Diodes*. (Cambridge University Press, Cambridge, United Kingdom, 2003).

[111] M. Versleijen, P. Kuindersma, K. Giok-Djan, and L. Meuleman, "Accurate analysis of dc electrical characteristics of 1.3 μm DCPBH laser diodes", IEEE J. Quantum Electron., **23**, 925 (1987).

[112] T. L. Paoli and P. A. Barnes, "Saturation of the junction voltage in stripe‐geometry (AlGa)As double‐heterostructure junction lasers", Appl. Phys. Lett., **28**, 714 (1976).

[113] H. S. Sommers, "Algorithm for measuring the internal quantum efficiency of individual injection lasers", Appl. Phys. Lett., **32**, 547 (1978).

[114] A. R. Reisinger, R. N. Roberts, S. R. Chinn, and T. H. Myers Ii, "Photoluminescence of infrared-sensing materials using an FTIR spectrometer", Review of Scientific Instruments, **60**, 82 (1989).

[115] J. Shao, W. Lu, X. Lu, F. Yue, Z. Li, S. Guo, and J. Chu, "Modulated photoluminescence spectroscopy with a step-scan Fourier transform infrared spectrometer", Review of Scientific Instruments, **77**, 063104 (2006).

[116] Y. G. Zhang, Y. Gu, K. Wang, X. Fang, A. Z. Li, and K. H. Liu, "Fourier transform infrared spectroscopy approach for measurements of photoluminescence and electroluminescence in mid-infrared", Review of Scientific Instruments, **83**, 053106 (2012).

[117] J. Shao, W. Lu, F. Yue, X. Lu, W. Huang, Z. Li, S. Guo, and J. Chu, "Photoreflectance spectroscopy with a step-scan Fourier-transform infrared

spectrometer: Technique and applications", Review of Scientific Instruments, **78**, 013111 (2007).

[118] B. W. Hakki and T. L. Paoli, "cw degradation at 300[degree]K of GaAs double-heterostructure junction lasers. II. Electronic gain", J. Appl. Phys., **44**, 4113 (1973).

[119] B. W. Hakki and T. L. Paoli, "Gain spectra in GaAs double - heterostructure injection lasers", J. Appl. Phys., **46**, 1299 (1975).

[120] A. Soibel, K. Mansour, Y. Qiu, C. J. Hill, and R. Q. Yang, "Optical gain, loss, and transparency current in high performance mid-infrared interband cascade lasers", J. Appl. Phys., **101**, 093104 (2007).

[121] C. Sirtori, P. Kruck, S. Barbieri, H. Page, J. Nagle, M. Beck, J. Faist, and U. Oesterle, "Low-loss Al-free waveguides for unipolar semiconductor lasers", Appl. Phys. Lett., **75**, 3911 (1999).

[122] W. W. Bewley, J. R. Lindle, C. L. Canedy, M. Kim, C. S. Kim, D. C. Larrabee, I. Vurgaftman, and J. R. Meyer, "Gain, loss, and internal efficiency in interband cascade lasers emitting at λ = 3.6-4.1 μm", J. Appl. Phys., **103**, 013114 (2008).

[123] B. A. Ikyo, I. P. Marko, A. R. Adams, S. J. Sweeney, C. L. Canedy, I. Vurgaftman, C. S. Kim, M. Kim, *et al.*, "Temperature dependence of 4.1 μm mid-infrared type II ``W" interband cascade lasers", Appl. Phys. Lett., **99**, 021102 (2011).

[124] C. Zhou, B. Cui, I. Vurgaftman, C. L. Canedy, C. S. Kim, M. Kim, W. W. Bewley, C. D. Merritt, *et al.*, "Thermal conductivity tensors of the cladding and active layers of interband cascade lasers", Appl. Phys. Lett., **105**, 261905 (2014).

[125] D. G. Cahill and R. O. Pohl, "Thermal conductivity of amorphous solids above the plateau", Phys. Rev. B, **35**, 4067 (1987).

[126] C. A. Evans, V. D. Jovanovic, D. Indjin, Z. Ikonic, and P. Harrison, "Investigation of thermal effects in quantum-cascade lasers", IEEE J. Quantum Electron., **42**, 859 (2006).

[127] H. K. Lee and J. S. Yu, "Thermal analysis of short wavelength InGaAs/InAlAs quantum cascade lasers", Solid-State Electronics, **54**, 769 (2010).

[128] Y. M. Ali and L. C. Zhang, "Relativistic heat conduction", International Journal of Heat and Mass Transfer, **48**, 2397 (2005).

[129] L. Wei, A.-Z. Li, Y.-G. Zhang, and Y.-Y. Li, "The Self-Heating Effect of Quantum Cascade Lasers Based on a Spectroscopic Method", Chinese Physics Letters, **26**, 084206 (2009).

[130] W. G. Scheibenzuber and U. T. Schwarz, "Fast self-heating in GaN-based laser diodes", Appl. Phys. Lett., **98**, 181110 (2011).

[131] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "High-power/low-threshold type-II interband cascade mid-IR laser-design and modeling", IEEE Photonics Technol. Lett., **9**, 170 (1997).

[132] C. Sirtori, J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, "Quantum Cascade Laser with Plasmon-Enhanced Wave-Guide Operating at 8.4 μm Wavelength", Appl. Phys. Lett., **66**, 3242 (1995).

[133] Z. Tian, C. Chen, R. Q. Yang, T. D. Mishima, M. B. Santos, J. C. Keay, M. B. Johnson, and J. F. Klem, "InAs-based plasmon-waveguide interband cascade lasers", Proc. SPIE,**7616**, 76161B (2010)

[134] C. Sirtori, C. Gmachl, F. Capasso, J. Faist, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, "Long-wavelength ($\lambda \approx$ 8-11.5 μ m) semiconductor lasers with waveguides based on surface plasmons", Opt. Lett., **23**, 1366 (1998).

[135] K. Unterrainer, R. Colombelli, C. Gmachl, F. Capasso, H. Y. Hwang, A. M. Sergent, D. L. Sivco, and A. Y. Cho, "Quantum cascade lasers with double metal-semiconductor waveguide resonators", Appl. Phys. Lett., **80**, 3060 (2002).

[136] M. Bahriz, G. Lollia, A. N. Baranov, and R. Teissier, "High temperature operation of far infrared (λ~20μm) InAs/AlSb quantum cascade lasers with dielectric waveguide", Opt. Express, **23**, 1523 (2015).

[137] E. D. Palik, *Handbook of Optical Constants of Solids*. (Elsevier, 1998).

[138] P. P. Paskov, "Refractive indices of InSb, InAs, GaSb, $InAs_xSb_{1-x}$, and $In_{1-x}Ga_xSb$: Effects of free carriers", J. Appl. Phys., **81**, 1890 (1997).

[139] Y. B. Li and et al., "Infrared reflection and transmission of undoped and Si-doped InAs grown on GaAs by molecular beam epitaxy", Semicond. Sci. Technol., **8**, 101 (1993).

[140] S. Shokhovets, O. Ambacher, and G. Gobsch, "Conduction-band dispersion relation and electron effective mass in III-V and II-VI zinc-blende semiconductors", Phys. Rev. B, **76**, 125203 (2007).

[141] D. M. Caughey and R. E. Thomas, "Carrier mobilities in silicon empirically related to doping and field", Proceedings of the IEEE, **55**, 2192 (1967).

[142] C. M. Maziar and M. S. Lundstrom, "Caughey-Thomas parameters for electron mobility calculations in GaAs", Electron. Lett., **22**, 565 (1986).

[143] M. Sotoodeh, A. H. Khalid, and A. A. Rezazadeh, "Empirical low-field mobility model for III--V compounds applicable in device simulation codes", J. Appl. Phys., **87**, 2890 (2000).

[144] A. N. Baranov, personal communications (2010).

[145] R. T. Hinkey, Z. Tian, R. Q. Yang, T. D. Mishima, and M. B. Santos, "Reflectance spectrum of plasmon waveguide interband cascade lasers and observation of the Berreman effect", J. Appl. Phys., **110**, 043113 (2011).

[146] K. Ohtani and H. Ohno, "An InAs-Based Intersubband Quantum Cascade Laser", Jpn.J. Appl. Phys., **41**, L1279 (2002).

[147] K. Ohtani and H. Ohno, "InAs/AlSb quantum cascade lasers operating at 10 μm", Appl. Phys. Lett., **82**, 1003 (2003).

[148] K. Ohtani, K. Fujita, and H. Ohno, "Mid-infrared InAs/AlGaSb superlattice quantum-cascade lasers", Appl. Phys. Lett., **87**, 211113 (2005).

[149] K. Ohtani, Y. Moriyasu, H. Ohnishi, and H. Ohno, "Above room-temperature operation of InAs/AlGaSb superlattice quantum cascade lasers emitting at 12 mu m", Appl. Phys. Lett., **90**, 261112 (2007).

[150] K. Ohtani, H. Sakuma, and H. Ohno, "InAs-based quantum cascade light emitting structures containing a double plasmon waveguide", J. Cryst. Growth, **251**, 718 (2003).

[151] G. H. B. Thompson and P. Kirkby, "(GaAl)As lasers with a heterostructure for optical confinement and additional heterojunctions for extreme carrier confinement", IEEE J. Quantum Electron., **9**, 311 (1973).

[152] P. Y. Yu and M. Cardona, "Temperature Coefficient of the Refractive Index of Diamond- and Zinc-Blende-Type Semiconductors", Phys. Rev. B, **2**, 3193 (1970).

[153] G. B. Hocker and W. K. Burns, "Mode dispersion in diffused channel waveguides by the effective index method", Applied Optics, **16**, 113 (1977).

[154] I. Vurgaftman, W. W. Bewley, R. E. Bartolo, C. L. Felix, M. J. Jurkovic, J. R. Meyer, M. J. Yang, H. Lee, *et al.*, "Far-field characteristics of mid-infrared angled-grating distributed feedback lasers", J. Appl. Phys., **88**, 6997 (2000).

[155] https://en.wikipedia.org/wiki/Gaussian_beam

[156] O. Svelto, *Principles of Lasers*, 5th ed. (Springer US, 2010).

[157] Y. Jiang, L. Li, H. Ye, R. Q. Yang, T. D. Mishima, M. B. Santos, M. B. Johnson, D. J. Feng, *et al.*, "InAs-Based Single-Mode Distributed Feedback Interband Cascade Lasers", IEEE J. Quantum Electron., **51**, 2300307 (2015).

[158] C. S. Kim, M. Kim, W. W. Bewley, J. R. Lindle, C. L. Canedy, J. Abell, I. Vurgaftman, and J. R. Meyer, "Corrugated-sidewall interband cascade lasers with single-mode midwave-infrared emission at room temperature", Appl. Phys. Lett., **95**, 231103 (2009).

[159] C. S. Kim, M. Kim, W. W. Bewley, J. R. Lindle, C. L. Canedy, J. A. Nolde, D. C. Larrabee, I. Vurgaftman, *et al.*, "Broad-stripe, single-mode, mid-IR interband

cascade laser with photonic-crystal distributed-feedback grating", Appl. Phys. Lett., **92**, 071110 (2008).

[160] W. W. Bewley, C. S. Kim, C. L. Canedy, C. D. Merritt, I. Vurgaftman, J. Abell, J. R. Meyer, and M. Kim, "High-power, high-brightness continuous-wave interband cascade lasers with tapered ridges", Appl. Phys. Lett., **103**, 111111 (2013).

[161] G. M. Smith, D. V. Forbes, R. M. Lammert, and J. J. Coleman, "Metallization to asymmetric cladding separate confinement heterostructure lasers", Appl. Phys. Lett., **67**, 3847 (1995).

[162] Y. Jiang, L. Li, Z. Tian, R. T. Hinkey, R. Q. Yang, T. D. Mishima, M. B. Santos, M. B. Johnson, *et al.*, "Room-Temperature InAs-based Interband Cascade Lasers", CLEO,CF3K.1 (2012)

[163] Y. Jiang, L. Li, Z. Tian, H. Ye, L. Zhao, R. Q. Yang, T. D. Mishima, M. B. Santos, *et al.*, "Electrically widely tunable interband cascade lasers", J. Appl. Phys., **115**, 113101 (2014).

[164] I. Vurgaftman, W. W. Bewley, C. L. Canedy, C. Kim, M. Kim, J. J. Ryan Lindle, C. D. Merritt, J. Abell, *et al.*, "Mid-IR Type-II Interband Cascade Lasers", IEEE J. Select. Topics Quantum Electron., **17**, 1435 (2011).

[165] R. S. Dhar, L. Li, H. Ye, S. G. Razavipour, X. Wang, R. Q. Yang, and D. Ban, "Nanoscopically resolved dynamic charge-carrier distribution in operating interband cascade lasers", Laser Photon. Rev., **2**, 224 (2015).

[166] S. M. Shank, J. A. Varriano, and G. W. Wicks, "Single quantum well GaAs/AlGaAs separate confinement heterostructure lasers with n - type modulation doped cores", Appl. Phys. Lett., **61**, 2851 (1992).

[167] I. Vurgaftman and J. R. Meyer, "High-Temperature Interband Cascade Lasers", Patent No. US 20100097690A1 (2010).

[168] I. Vurgaftman, J. R. Meyer, C. L. Canedy, W. W. Bewley, J. R. Lindle, C.-s. Kim, and M. Kim, "High-temperature interband cascade lasers", Patent No. US 8,125,706 B2 (2012).

[169] I. Vurgaftman, "Gain medium design for IC laser", Patent No. US 2013/0003770A1 (2013).

[170] I. Vurgaftman and et al., "Mid-infrared interband cascade lasers operating at ambient temperatures", New J. Phys., **11**, 125015 (2009).

[171] C. S. Kim, M. Kim, J. Abell, W. W. Bewley, C. D. Merritt, C. L. Canedy, I. Vurgaftman, and J. R. Meyer, "Mid-infrared distributed-feedback interband cascade lasers with continuous-wave single-mode emission to 80 [degree]C", Appl. Phys. Lett., **101**, 061104 (2012).

[172] J. Devenson, D. Barate, O. Cathabard, R. Teissier, and A. N. Baranov, "Very short wavelength (lambda = 3.1--3.3 mu m) quantum cascade lasers", Appl. Phys. Lett., **89**, 191115 (2006).

[173] O. Cathabard, R. Teissier, J. Devenson, J. C. Moreno, and A. N. Baranov, "Quantum cascade lasers emitting near 2.6 μ m", Appl. Phys. Lett., **96**, 141110 (2010).

[174] P. Laffaille, J. C. Moreno, R. Teissier, M. Bahriz, and A. N. Baranov, "High temperature operation of short wavelength InAs-based quantum cascade lasers", AIP Advances, **2**, 022119 (2012).

[175] H. F. Lockwood, H. Kressel, H. S. Sommers, and F. Z. Hawrylo, "An efficient large optical cavity injection laser", Appl. Phys. Lett., **17**, 499 (1970).

[176] E. F. Pecora, H. Sun, L. Dal Negro, and T. D. Moustakas, "Deep-UV optical gain in AlGaN-based graded-index separate confinement heterostructure", Opt. Mater. Express, **5**, 809 (2015).

[177] S. Kawakami and S. Nishida, "Characteristics of a doubly clad optical fiber with a low-index inner cladding", IEEE J. Quantum Electron., **10**, 879 (1974).

[178] W. Bewley, C. Canedy, C. S. Kim, M. Kim, J. R. Lindle, J. Abell, I. Vurgaftman, and J. Meyer, "Ridge-width dependence of midinfrared interband cascade laser characteristics", Opt. Eng., **49**, 111116 (2010).

[179] M. Dallner, J. Scheuermann, L. Nähle, M. Fischer, J. Koeth, S. Höfling, and M. Kamp, "InAs-based distributed feedback interband cascade lasers", Appl. Phys. Lett., **107**, 181105 (2015).

[180] Z. Tian, L. Li, H. Ye, R. Q. Yang, T. D. Mishima, M. B. Santos, and M. B. Johnson, "InAs-based interband cascade lasers with emission wavelength at 10.4 μm", Electron. Lett., **48**, 113 (2012).

[181] K. Le, "Design of type-II interband cascade lasers for terahertz emission", Optical and Quantum Electronics, **40**, 973 (2008).

[182] G. D. Shen, D. X. Xu, M. Willander, and G. V. Hansson, "The origin of the temperature dependence in resonant tunneling transport", High Speed Semiconductor Devices and Circuits, 1991., Proceedings IEEE/Cornell Conference on Advanced Concepts in,84 (1991)

[183] J. F. Chen, M. C. Wu, L. Yang, and A. Y. Cho, "InAs/AlSb/GaSb single‐barrier interband tunneling diodes with high peak‐to‐valley ratios at room temperature", J. Appl. Phys., **68**, 3040 (1990).

[184] D. Li, J. Shao, L. Tang, C. Edmunds, G. Gardner, M. J. Manfra, and O. Malis, "Temperature-dependence of negative differential resistance in GaN/AlGaN resonant tunneling structures", Semicond. Sci. Technol., **28**, 074024 (2013).

[185] A. Bauer, F. Langer, M. Dallner, M. Kamp, M. Motyka, G. Sek, K. Ryczko, J. Misiewicz, *et al.*, "Emission wavelength tuning of interband cascade lasers in the 3--4 mm spectral range", Appl. Phys. Lett., **95**, 251103 (2009).

[186] M. Motyka, K. Ryczko, G. Sęk, F. Janiak, J. Misiewicz, A. Bauer, S. Höfling, and A. Forchel, "Type II quantum wells on GaSb substrate designed for laser-based gas sensing applications in a broad range of mid infrared", Optical Materials, **34**, 1107 (2012).

[187] S. Höfling, R. Weih, M. Dallner, J. Scheuermann, M. von Edlinger, L. Nähle, M. Fischer, J. Koeth, *et al.*, "Mid-infrared (~2.8 μm to ~7.1 μm) interband cascade lasers", SPIE,**9550**, 95500F (2015)

[188] K. Wörle, F. Seichter, A. Wilk, C. Armacost, T. Day, M. Godejohann, U. Wachter, J. Vogt, *et al.*, "Breath Analysis with Broadly Tunable Quantum Cascade Lasers", Anal. Chem., **85**, 2697 (2013).

[189] M. Motyka, G. Sek, K. Ryczko, J. Misiewicz, T. Lehnhardt, S. Hofling, and A. Forchel, "Optical properties of GaSb-based type II quantum wells as the active region of midinfrared interband cascade lasers for gas sensing applications", Appl. Phys. Lett., **94**, 251901 (2009).

[190] C. L. Canedy, W. W. Bewley, M. Kim, C. S. Kim, J. A. Nolde, D. C. Larrabee, J. R. Lindle, I. Vurgaftman, *et al.*, "High-temperature interband cascade lasers emitting at λ= 3.6-4.3 μm", Appl. Phys. Lett., **90**, 181120 (2007).

[191] J. S. Yu, A. Evans, S. Slivken, S. R. Darvish, and M. Razeghi, "Short wavelength λ~4.3 μm high-performance continuous-wave quantum-cascade lasers", IEEE Photonics Technol. Lett., **17**, 1154 (2005).

[192] J. Faist, C. Gmachl, F. Capasso, C. Sirtori, D. L. Sivco, J. N. Baillargeon, and A. Y. Cho, "Distributed feedback quantum cascade lasers", Appl. Phys. Lett., **70**, 2670 (1997).

[193] A. Mohan, A. Wittmann, A. Hugi, S. Blaser, M. Giovannini, and J. Faist, "Room-temperature continuous-wave operation of an external-cavity quantum cascade laser", Opt. Lett., **32**, 2792 (2007).

[194] D. Caffey, T. Day, C. S. Kim, M. Kim, I. Vurgaftman, W. W. Bewley, J. R. Lindle, C. L. Canedy, *et al.*, "Performance characteristics of a continuous-wave compact widely tunable external cavity interband cascade lasers", Opt. Express, **18**, 15691 (2010).

[195] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and C. A. Burrus, "Band-Edge Electroabsorption in Quantum Well Structures: The Quantum-Confined Stark Effect", Phys. Rev. Lett., **53**, 2173 (1984).

[196] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, "Laser action by tuning the oscillator strength", Nature, **387**, 777 (1997).

[197] A. Muller, M. Beck, J. Faist, U. Oesterle, and M. Ilegems, "Electrically tunable, room-temperature quantum-cascade lasers", Appl. Phys. Lett., **75**, 1509 (1999).

[198] C. Gmachl, D. L. Sivco, R. Colombelli, F. Capasso, and A. Y. Cho, "Ultra-broadband semiconductor laser", Nature, **415**, 883 (2002).

[199] Y. Yao, K. J. Franz, X. Wang, J.-Y. Fan, and C. Gmachl, "A widely voltage-tunable quantum cascade laser based on ``two-step" coupling", Appl. Phys. Lett., **95**, 021105 (2009).

[200] A. Bismuto, R. Terazzi, M. Beck, and J. Faist, "Electrically tunable, high performance quantum cascade laser", Appl. Phys. Lett., **96**, 141105 (2010).

[201] S. Suchalkin, M. V. Kisin, S. Luryi, G. Belenky, F. J. Towner, J. D. Bruno, C. Monroy, and R. L. Tober, "Widely tunable type-II interband cascade laser", Appl. Phys. Lett., **88**, 031103 (2006).

[202] N. Le Thomas, N. T. Pelekanos, and Z. Hatzopoulos, "Tunable laser diodes by Stark effect", Appl. Phys. Lett., **83**, 1304 (2003).

[203] M. Motyka, K. Ryczko, M. Dyksik, G. Sęk, J. Misiewicz, R. Weih, M. Dallner, S. Höfling, *et al.*, "On the modified active region design of interband cascade lasers", J. Appl. Phys., **117**, 084312 (2015).

[204] R. S. Dhar, "Nanometer Probing of Operating Nano-Photonic Devices",Dissertation, University of Waterloo, 2014.

[205] J. I. Pankove, "Temperature dependence of emission efficiency and lasing threshold in laser diodes", IEEE J. Quantum Electron., **4**, 119 (1968).

[206] R. Chin, N. Holonyak, B. A. Vojak, K. Hess, R. D. Dupuis, and P. D. Dapkus, "Temperature dependence of threshold current for quantum ‐ well $Al_xGa_{1-x}As$ ‐ GaAs heterostructure laser diodes", Appl. Phys. Lett., **36**, 19 (1980).

[207] N. Tansu, C. Ying-Lan, T. Takeuchi, D. P. Bour, S. W. Corzine, M. R. T. Tan, and L. J. Mawst, "Temperature analysis and characteristics of highly strained InGaAs-GaAsP-GaAs (λ> 1.17 μm) quantum-well lasers", IEEE J. Quantum Electron., **38**, 640 (2002).

[208] S. R. Bank, L. L. Goddard, M. A. Wistey, H. B. Yuen, and J. S. Harris, "On the temperature sensitivity of 1.5 μm GaInNAsSb lasers", IEEE J. Select. Topics Quantum Electron., **11**, 1089 (2005).

[209] T. Garrod, D. Olson, M. Klaus, C. Zenner, C. Galstad, L. Mawst, and D. Botez, "50% continuous-wave wallplug efficiency from 1.53 μ m-emitting broad-area diode lasers", Appl. Phys. Lett., **105**, 071101 (2014).

[210] L. Cerutti, A. Castellano, J. B. Rodriguez, M. Bahri, L. Largeau, A. Balocchi, K. Madiomanana, F. Lelarge, *et al.*, "GaSb-based composite quantum wells for laser diodes operating in the telecom wavelength range near 1.55-μm", Appl. Phys. Lett., **106**, 101102 (2015).

[211] Y. Jiang, L. Li, R. Q. Yang, J. A. Gupta, G. C. Aers, E. Dupont, J.-M. Baribeau, X. Wu, *et al.*, "Type-I interband cascade lasers near 3.2 μm", Appl. Phys. Lett., **106**, 041117 (2015).

[212] J. I. Pankove, "Absorption Edge of Impure Gallium Arsenide", Phys. Rev., **140**, A2059 (1965).

[213] G. Lasher and F. Stern, "Spontaneous and Stimulated Recombination Radiation in Semiconductors", Phys. Rev., **133**, A553 (1964).

[214] G. C. Dousmanis, H. Nelson, and D. L. Staebler, "Temperature dependence of threshold current in GaAs Lasers", Appl. Phys. Lett., **5**, 174 (1964).

[215] F. Stern, "Effect of Band Tails on Stimulated Emission of Light in Semiconductors", Phys. Rev., **148**, 186 (1966).

[216] M. Pilkuhn, H. Rupprecht, and S. Blum, "Effect of temperature on the stimulated emission from GaAs p-n junctions", Solid-State Electronics, **7**, 905 (1964).

[217] H. C. Casey, "Room ‐ temperature threshold ‐ current dependence of GaAs ‐ $Al_xGa_{1-x}As$ double ‐ heterostructure lasers on x and active ‐ layer thickness", J. Appl. Phys., **49**, 3684 (1978).

[218] H. C. Casey, "Temperature dependence of the threshold current density in InP‐$Ga_{0.28}In_{0.72}As_{0.6}P_{0.4}$ (λ =1.3 μm) double heterostructure lasers", J. Appl. Phys., **56**, 1959 (1984).

[219] J. O’ Gorman, A. F. J. Levi, S. Schmitt‐Rink, T. Tanbun‐Ek, D. L. Coblentz, and R. A. Logan, "On the temperature sensitivity of semiconductor lasers", Appl. Phys. Lett., **60**, 157 (1992).

[220] J. O’Gorman, A. F. J. Levi, T. Tanbun‐Ek, D. L. Coblentz, and R. A. Logan, "Temperature dependence of long wavelength semiconductor lasers", Appl. Phys. Lett., **60**, 1058 (1992).

[221] E. P. O’Reilly and M. Silver, "Temperature sensitivity and high temperature operation of long wavelength semiconductor lasers", Appl. Phys. Lett., **63**, 3318 (1993).

[222] H. Temkin, D. Coblentz, R. A. Logan, J. P. van der Ziel, T. Tanbun‐Ek, R. D. Yadvish, and A. M. Sergent, "High temperature characteristics of InGaAsP/InP laser structures", Appl. Phys. Lett., **62**, 2402 (1993).

[223] S. R. Johnson and T. Tiedje, "Temperature dependence of the Urbach edge in GaAs", J. Appl. Phys., **78**, 5609 (1995).

[224] S. Seki, H. Oohasi, H. Sugiura, T. Hirono, and K. Yokoyama, "Dominant mechanisms for the temperature sensitivity of 1.3 μm InP‐based strained‐layer multiple‐quantum‐well lasers", Appl. Phys. Lett., **67**, 1054 (1995).

[225] W. Fang, M. Hattendorf, S. L. Chuang, J. Minch, C. S. Chang, C. G. Bethea, and Y. K. Chen, "Analysis of temperature sensitivity in semiconductor lasers using gain and spontaneous emission measurements", Appl. Phys. Lett., **70**, 796 (1997).

[226] S. R. Chinn, P. S. Zory, and A. R. Reisinger, "A model for GRIN-SCH-SQW diode lasers", IEEE J. Quantum Electron., **24**, 2191 (1988).

Appendix A: Fourier transform in lock-in amplifier

A lock-in amplifier (LIA) is a very useful electronic equipment to extract the pulsed signal from the noisy background. However, the reading on the display is different from the input signal. It's important to figure out their relation so as to restore the original input signal.

Assume a square pulse wave f(t), as shown in Figure A-1. It has a peak-to-peak height of P, a period of T and a duty cycle of $c=T_1/T$. A period function can be decomposed into the sum of a set of sine and cosine functions by Fourier series expansion:

$$f(t)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\left[a_n\cos(n\omega t)+b_n\sin(n\omega t)\right] \tag{A.1}$$

where $w=2\pi/T$ is the circular frequency. The coefficient a_n and b_n are defined by:

$$\begin{aligned} a_n &= \frac{2}{T}\int_0^T f(t)\cos\left(\frac{2n\pi}{T}t\right)dt=\frac{2}{T}\int_0^{T_1} P\cos(n\omega t)dt, n=0,1,2,\cdots \\ b_n &= \frac{2}{T}\int_0^T f(t)\sin\left(\frac{2n\pi}{T}t\right)dt=\frac{2}{T}\int_0^{T_1} P\sin(n\omega t)dt, n=1,2,\cdots \end{aligned} \tag{A.2}$$

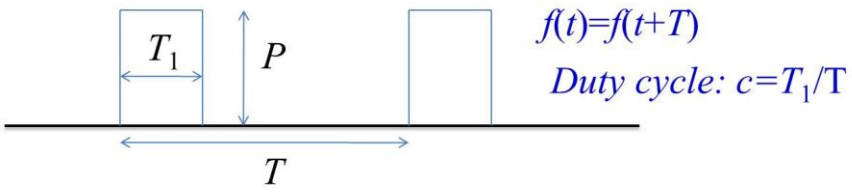


Figure A-1. a squared pulsed wave

The zero order and first order of Fourier coefficient can be explicitly given by:

$$a_0=\frac{2}{T}\int_0^{T_1} P\cos 0dt=\frac{2}{T}T_1P=2Pc \tag{A.3}$$

$$a_1 = \frac{2}{T}\int_0^{T_1} P\cos\left(\frac{2\pi}{T}t\right)dt = \frac{P}{\pi}\sin(2\pi c) = \frac{2P}{\pi}\sin(\pi c)\cos(\pi c)$$
$$b_1 = \frac{2}{T}\int_0^{T_1} P\sin\left(\frac{2\pi}{T}t\right)dt = \frac{P}{\pi}[1-\cos(2\pi c)] = \frac{2P}{\pi}\sin^2(\pi c) \quad \text{(A.4)}$$

Because LIA only extracts the first harmonic component in the signal, we pack the first order term together:

$$a_1\cos(\omega t) + b_1\sin(\omega t) = \frac{2P}{\pi}\sin(\pi c)\cos(\pi c - \omega t) = \frac{2P}{\pi}\sin(\pi c)\sin(\omega t + \theta_{sig}) \quad \text{(A.5)}$$

From the manual of the LIA, we know the internal reference signal is $V_L\sin(\omega_{ref}t + \theta_{ref})$.The output of the phase-sensitive detector (PSD) is the product of input signal and internal signal:

$$V_{psd} = \frac{2P}{\pi}\sin(\pi c)\sin(\omega t + \theta_{sig})V_L\sin(\omega_{ref}t + \theta_{ref}) \quad \text{(A.6)}$$

If the internal reference signal is triggered to the source frequency of the input signal, they have the same frequency ($\omega = \omega_{ref}$). With a low pass filter, the frequency-independent value (DC signal) can be output as:

$$X = \frac{P}{\pi}\sin(\pi c)V_L\cos(\theta_{sig} - \theta_{ref}) \quad \text{(A.7)}$$

Note that X is phase dependent. By changing phase by $\pi/2$, the output from another PSD is :

$$Y = \frac{P}{\pi}\sin(\pi c)V_L\sin(\theta_{sig} - \theta_{ref}) \quad \text{(A.8)}$$

We are most interested in the phase-independent amplitude:

$$R = \sqrt{X^2 + Y^2} = \frac{P}{\pi}\sin(\pi c)V_L \quad \text{(A.9)}$$

The ultimate output in the display of LIA is actually a root mean square (rms) value:

$$V_{rms} = \frac{R}{\sqrt{2}} = \frac{P}{\sqrt{2}\pi}\sin(\pi c)V_L \qquad \text{(A.10)}$$

Then the only unknown factor is V_L, which can be determined by a simple measurement. For this purpose, a resistor of 50 Ω driven by a current source with 50% duty cycle is used. The peak-peak value obtained in an oscilloscope, as shown in the x-axis of Figure A-2, is the voltage input to the LIA. The reading in LIA is plotted as the y-axis value. A predicted curve using V_L=2 is in excellent agreement with the experimental value.

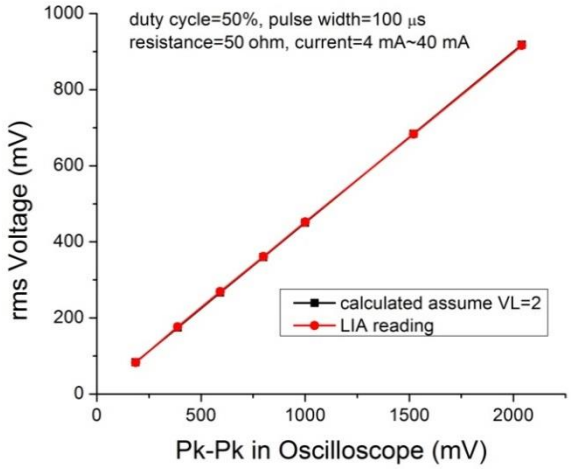


Figure A-2. Determination of V_L by equation (A.10)

In conclusion, the relation between the input voltage P and the LIA display V_{rms} is:

$$V_{rms} = \frac{\sqrt{2}P}{\pi}\sin(\pi c) = 0.45P\sin(\pi c) \qquad \text{(A.11)}$$

Appendix B: Light collecting efficiency due to beam divergence

The laser beam is very divergent along the fast axis due to the small size of active QW. As shown in Figure B-1, the full divergence angle is 95° in the vertical direction, almost twice the receiving angle of the active element inside a power meter. This severely decreases the light collecting efficiency. An accurate value of the laser output power is very critical to extracting the most fundamental properties such as internal efficiency and internal loss, as discussed in section 3.3.1. This appendix aims to estimate the light collecting efficiency in our current setup.

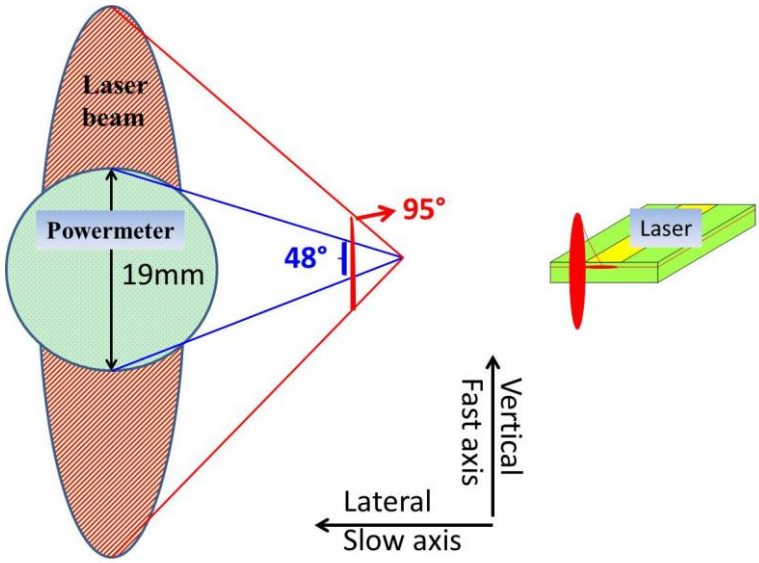


Figure B-1. Elliptical laser beam struck on the round sensor of a power meter

Figure B-2 shows a thermal pile power meter and an MCT detector that are used in our lab to collect the laser power. The power meter is put as close as to the cryostat's optical window the maximize the power collection. A divergence half angle is calculated to be 24° in the vertical direction. In contrast, the MCT detector is much more sensitive to a rather high response. Its active element is quite small, with a diameter of 1 mm. If it's put as close as the cryostat window, the far-field patterns can

be measured. The receiving half angle for this MCT detector is 1.5°. As a result, the resolution of the far-field measurement can't be better than 3°.

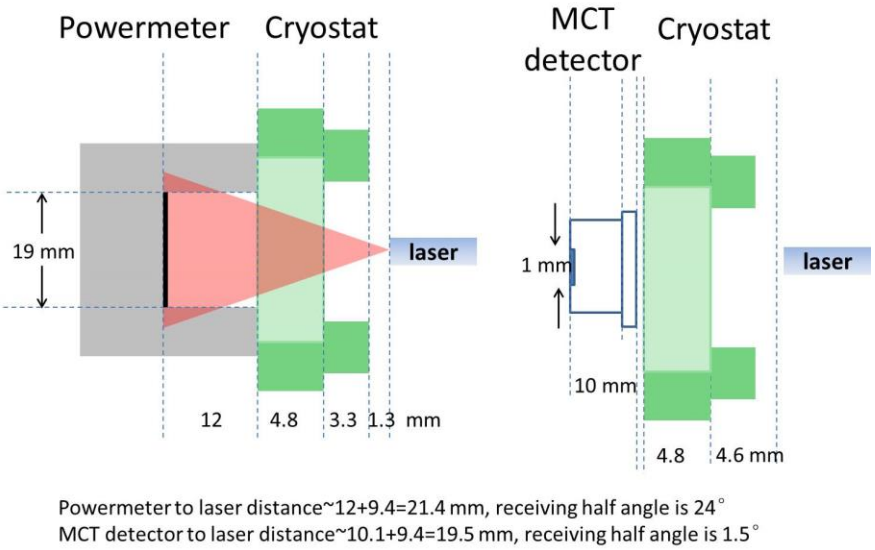


Figure B-2. Receiving angle of power meter and MCT detector

If the far-field pattern is measured by the translation method, as in Figure 4-7(a), attention must be paid to the finite field of view of the MCT detector. The detailed geometric structure of the MCT detector is shown in Figure B-3. Due to the finite size of the detector window, when the oblique incidence angle is larger than 33°, the light can't fully reach the 1×1 mm^2 sensitive element. If the incident angle is greater than 38°, the light will be totally blocked by the opaque material. This means that the translation is not suitable to measure the vertical far-field pattern because the divergence half angle of a typical laser is larger than 45°. We note that the optical window of the cryostat has an emitting half angle of about 65°, which is much bigger than the divergence half angle along the fast axis.

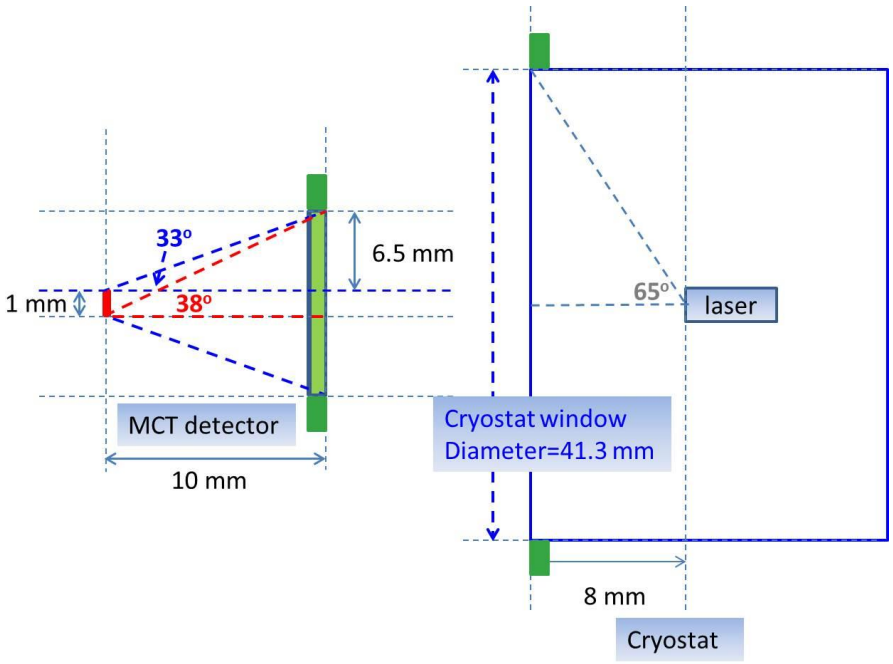
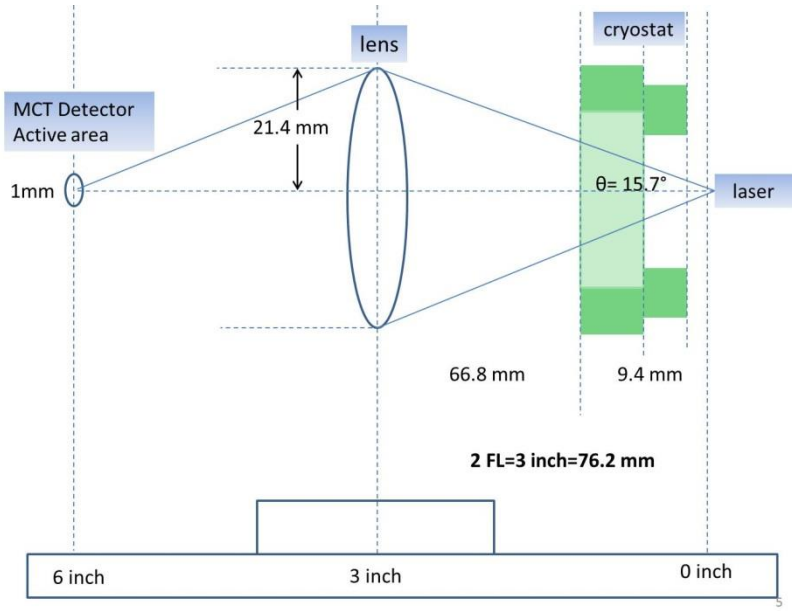


Figure B-3. Internal structure of the MCT detector

In the collection of laser output power in pulsed operation, a lens is used to focus the light to the tiny active area of MCT detector, as shown in Figure B-4. According to the geometry optics, the lens must put two times focal lengths away from the laser source. As a result, the receiving half angle for the lens is only 15.7°, and the collecting efficiency is much smaller than that of the power meter near the cryostat. This also causes some difficulty in the optical alignment for a brand new test.



B-4. Experimental setup for collecting laser power in pulsed operation

After we examine blind areas of all the setups, the measured far-field patterns can be used to estimate the collecting efficiency. Figure B-5 shows a far-field pattern measured on a narrow ridge device R083NR_1F. We can immediately recognize that the data was collected by the translation method because the light intensity decay to zero near 38°. Although this far-field measurement is not desirable, we could still get a rough estimate of the collecting efficiency. With a simple integral, we can calculate that 72% of the light is collected by the lens, and 83% of the light is collected directly by the power meter.

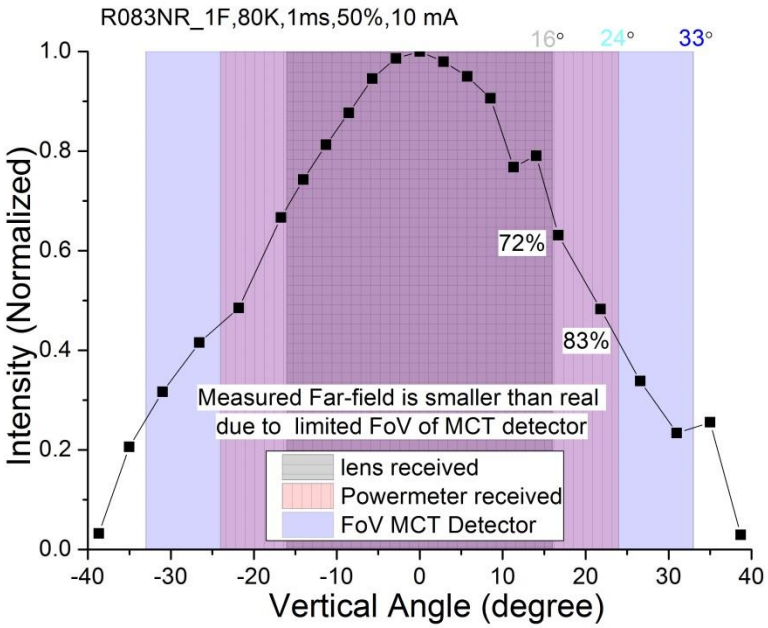


Figure B-5. Collecting efficiency along the fast axis

Along the lateral direction, the collecting efficiencies are expected to be higher due to smaller divergence. At low injection current, the collecting efficiencies for lens and power meter are 89% and 96%, respectively. But as the current increased, the beam became significantly wider, as shown in Figure B-6. The collection efficiency of lens decreased sharply to 71%. The collection efficiency of power meter decreased slowly due to its larger receiving angle. This means that the degradation of beam quality is also one of the reason for a decreasing power slope efficiency at large injection.

A further measurement is required to get a more accurate estimation of the collection efficiency. Anyway, this appendix serves as a good starting point.

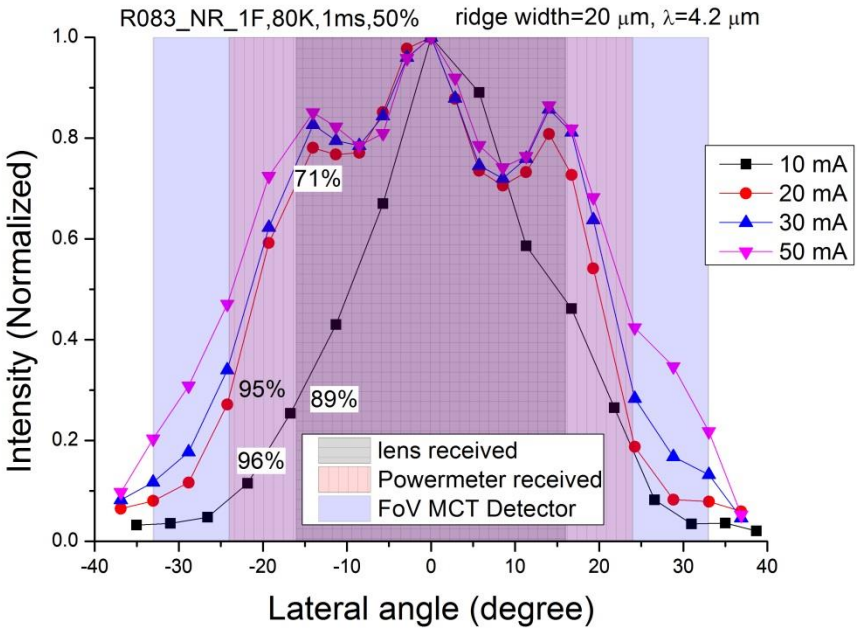


Figure B-6. Collecting efficiency along the slow axis

Appendix C: Plasmonic waveguide

As mentioned at the end of section 4.1.1, the terminology "plasmon waveguide" in InAs-based QC [1] or IC lasers [2] was often used to refer to "plasmon-enhanced waveguide" that consists of highly doped InAs as the cladding layer. This may confuse readers because "plasmon waveguide" usually refers to "plasmonic waveguide" or "surface plasmon waveguide". To avoid such confusion, I think it's better to fix the name "plasmon-enhanced waveguide" without any abbreviation. Because in principle, "plasmon-enhanced waveguide" is still optical waveguide that confines optical wave instead of surface plasmon wave.

To fully explain the surface plasmon wave, I will also clarify some concepts with similar spelling: plasma, plasmon, surface plasmon, and surface plasmon polariton.

Plasma can be understood as an ionized gas. It is the fourth fundamental state of matter after solid, liquid and gas. A plasma can be created by heating a gas or subjecting it to a strong electromagnetic field. Actually, plasma is the most abundant matter in our universe, most of which is in the intergalactic regions and in stars, including the Sun. Typical examples of plasma in our daily life are lightning, plasma display, fluorescent lamps and plasma globe.

Plasmon is the quasi-particle of the collective oscillation of plasma, based on the "free electron model", also known as "Drude-Sommerfeld model." Similar to the photon, the plasmon also exhibits wave-particle duality [3]. The "particle" picture of **bulk plasmon** has been discussed by a damped oscillator in 4.1.2. The "wave" picture of **surface plasmon** [4] is presented as follows.

Assumed a slab waveguide without the core layer, as shown in Figure C-1, an electromagnetic wave propagates along the interface between layer 1 with a dielectric constant of ε_1 and layer 2 with a dielectric constant of ε_2. Due to the total internal reflection, the evanescent fields will decay exponentially away from the interface.

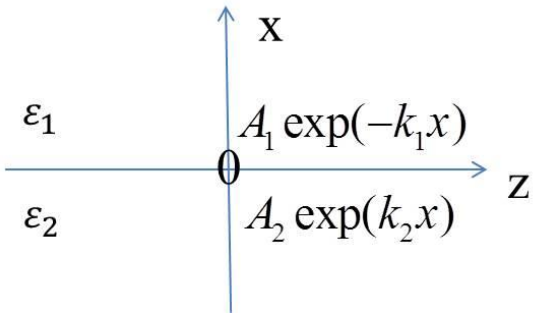


Figure C-1. Surface wave propagating along the interface x=0

Specifically, for TE modes:

$$E_y = A_1 \cdot \exp(-k_{1x} \cdot x),\ x \geq 0 \qquad \text{(C-1a)}$$

$$E_y = A_2 \cdot \exp(k_{2x} \cdot x),\ x \leq 0 \qquad \text{(C-1b)}$$

For TM modes:

$$H_y = A_1 \cdot \exp(-k_{1x} \cdot x),\ x \geq 0 \qquad \text{(C-1a)}$$

$$H_y = A_2 \cdot \exp(k_{2x} \cdot x),\ x \leq 0 \qquad \text{(C-1b)}$$

The wavevectors k_{1x} and k_{2x} are related to the propagating constant β by:

$$\beta^2 = k_{1x}^2 + k_0^2 \varepsilon_1 \qquad \text{(C-3a)}$$

$$\beta^2 = k_{2x}^2 + k_0^2 \varepsilon_2 \qquad \text{(C-3b)}$$

where $k_0=2\pi/\lambda$ is the wavevector in the vacuum. The exponential attenuation constants k_{1x} and k_{2x} are both positive numbers to ensure the exponential decay.

If the surface wave is TE polarized, the continuity condition requires that E_y and $\frac{\partial E_y(x)}{\partial x}$ are both continuous at x=0:

$$A_1=A_2 \tag{C-4a}$$

$$-k_{1x}\cdot A_1=k_{2x}\cdot A_2 \tag{C-4b}$$

Because the exponential attenuation constants k_{1x} and k_{2x} should be both positive, equations (C-4a,b) can never be satisfied.

If the surface wave is TM polarized, the continuity condition requires that H_y and $\frac{1}{\varepsilon}\frac{\partial H_y(x)}{\partial x}$ are both continuous at x=0:

$$A_1=A_2 \tag{C-5a}$$

$$-k_{1x}\cdot A_1/\varepsilon_1=k_{2x}\cdot A_2/\varepsilon_2 \tag{C-5b}$$

The equations (C-5a,b) can be satisfied only if the two dielectric constants are in the opposite sign, *i.e.*, $\varepsilon_1\cdot\varepsilon_2<0$. The solution, a TM polarized surface wave, is termed "**surface plasmon wave**". By solving equations (C-3a,b) and (C-5a,b) together, the propagating constant of the surface plasmon wave is:

$$\beta=\sqrt{\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1+\varepsilon_2}}k_0 \tag{C-6}$$

In order to have a positive real part of the propagating constant (so the surface plasmon wave can have a real phase velocity to transmit signal), we will require $\varepsilon_1 + \varepsilon_2 < 0$. This condition is quite easy to be satisfied, *e.g.*, the interface between air and a metal. In this scenario, $\varepsilon_1 = 1$, and the dielectric constant of metal can be calculated by a Drude model (as discussed in section 4.1.2):

$$\varepsilon_2 \approx 1 - \omega_p^2 / \omega^2 \tag{C-7}$$

where the plasma frequency ω_p is ~9 eV for typical metals such as gold and silver. For visible and infrared regions, the photon energies are smaller than 3 eV. So the condition $\varepsilon_1 + \varepsilon_2 < 0$ is satisfied.

If we substitute equation (C-7) into equation (C-6), replace $k_0 = \frac{\omega}{c}$, and define $k_p \equiv \frac{\omega_p}{c}$, the energy-wavevector dispersion relation of surface plasmon wave can be obtained as:

$$(\frac{\omega}{\omega_p})^2 = \frac{1}{2} + (\frac{\beta}{k_p})^2 - \frac{1}{2}\left[1 + 4(\frac{\beta}{k_p})^4\right]^{\frac{1}{2}} \tag{C-8}$$

Figure C-2 compares the energy dispersion relations of photons in metal, air and at the air/metal interface. The cut-off frequency of a surface plasmon wave is $\omega = \omega_p / \sqrt{2}$, which is equivalent to the condition $\varepsilon_1 + \varepsilon_2 < 0$. As we see, when the frequency of the surface plasmon wave approaches the cut-off frequency, its wavevector becomes much longer than the electromagnetic wave in air with the same energy. This means the wavelength of surface plasmon wave can be much shorter than

the regular light, which can be used to significantly improve the imaging resolution and overcome the diffraction limit.

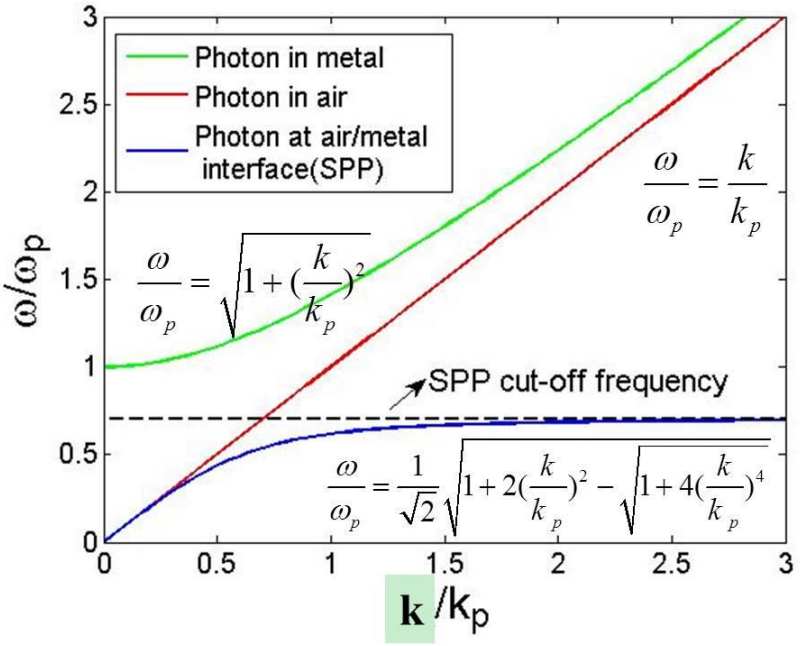


Figure C-2. Energy dispersion relations of photons in metal, air and at the air/metal interface

However, the downside of the longer wavevector is that the surface plasmon wave is not easy to excite or emit due to the momentum mismatch, even if its frequency is away from the cut-off frequency. Generally, there are two approaches to compensate the momentum mismatch for the generation of surface plasmon wave [5]. The first one is to use a prism coupler and take advantage of an evanescent wave, as shown in Figure C-3(a). The second one is to use a grating coupler, as shown in Figure C-3(b). Once the photon is coupled with the surface plasmon, this combination of photon and surface plasmon is called surface plasmon polariton (SPP).

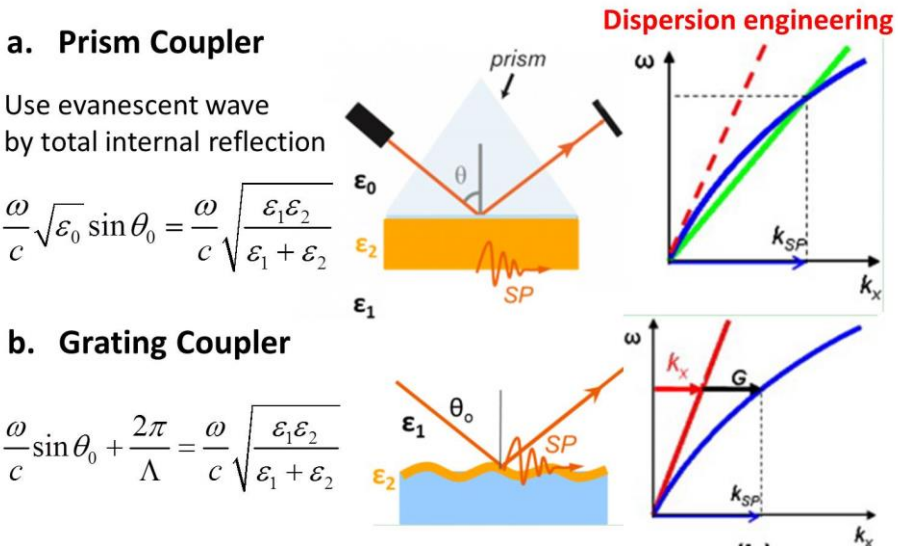


Figure C-3. Two ways to generate surface plasmon waves [5]

Actually, the propagating constant of a surface plasmon wave is usually a complex number. Consequently, the surface plasmon wave suffers significant propagation loss and has a finite propagation length. Instead of being used as a good waveguide, the surface plasmon wave is widely used to construct various biosensors [6] or optical fiber sensors [7]. Because the incident light intensity can be significantly reduced when the momentum matching is satisfied.

References

[1] K. Ohtani and H. Ohno, "An InAs-Based Intersubband Quantum Cascade Laser", Jpn.J. Appl. Phys., **41**, L1279 (2002).

[2] Z. Tian, R. Q. Yang, T. D. Mishima, M. B. Santos, R. T. Hinkey, M. E. Curtis, and M. B. Johnson, "InAs-based interband cascade lasers near 6 μm", Electron. Lett., **45**, 48 (2009).

[3] R. Kolesov, B. Grotz, G. Balasubramanian, R. J. Stohr, A. A. L. Nicolet, P. R. Hemmer, F. Jelezko, and J. Wrachtrup, "Wave-particle duality of single surface plasmon polaritons", Nat Phys, **5**, 470 (2009).

[4] P. Yeh, *Optical waves in layered media*. (Wiley Online Library, 1988).

[5] E. Wijaya, C. Lenaerts, S. Maricot, J. Hastanin, S. Habraken, J.-P. Vilcot, R. Boukherroub, and S. Szunerits, "Surface plasmon resonance-based biosensors: From the development of different SPR structures to novel surface functionalization strategies", Current Opinion in Solid State and Materials Science, **15**, 208 (2011).

[6] W. M. Mullett, E. P. C. Lai, and J. M. Yeung, "Surface Plasmon Resonance-Based Immunoassays", Methods, **22**, 77 (2000).

[7] B. D. Gupta and R. K. Verma, "Surface Plasmon Resonance-Based Fiber Optic Sensors: Principle, Probe Designs, and Some Applications", Journal of Sensors, **2009**, 12 (2009).

Appendix D: Coupling coefficient in DFB lasers

The coupling coefficient is an important physical quantity describing the degree to which the oppositely going waves convert into each other and thus the amount of feedback per unit length provided by the structure.

In Kogelnik and Shank's simple couple-wave model [1], they assume a spatial modulation of the complex refractive index with a cosine shape:

$$\tilde{n}(z) = \tilde{n} + \tilde{n}_1 \cos(2\beta_0 z) \tag{D-1}$$

where $\tilde{n}$ is the average value of the refractive index of the medium, $\tilde{n}_1$ is the amplitude of the spatial modulation, $\beta_0 = \frac{2\pi}{\lambda_B} = \frac{2\pi}{2n_{eff}\Lambda}$ is the Bragg wavevector and Λ is the period of the grating. If the perturbation of the refractive index is small, *i.e.*, $\tilde{n}_1 << \tilde{n}$, the coupling coefficient is simply written as [1]:

$$\kappa = \frac{\pi}{\lambda_0}\tilde{n}_1 \tag{D-2}$$

However, this only applies to the "weak-coupling" scenario. For more general cases with more realistic grating shapes, the coupling coefficient is given by [2-4]:

$$\kappa = \frac{k_0}{2\tilde{n}_{eff}} \frac{\sin(\pi dc)}{\pi} \Delta\varepsilon \cdot \Gamma_{grating} \tag{D-3}$$

where dc is the duty cycle, $\Gamma_{grating}$ is the confinement factor of the grating. For index grating, $\Delta\varepsilon = n_1^2 - n_2^2$; for gain grating, $\Delta\varepsilon = \tilde{n}_{eff}\Delta g / k_0$.

An example for calculation of the coupling coefficient is given below. The calculated structure is a metal grating consisting of InAs and gold, as shown in Figure D-1(real object see Figure 8-1). For InAs/gold, the difference of dielectric constant is 796+185i, a large number that can significantly enhance the coupling. For example, if the confinement factor of the grating can reach 0.01%, the optimal coupling coefficient will be ~50 cm^{-1}.

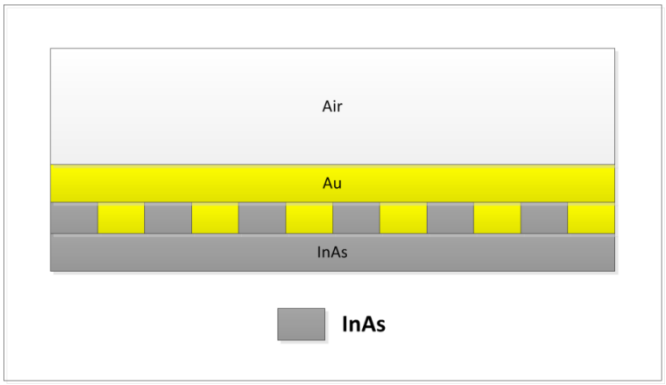


Figure D-1 Schematic of metal grating

The effective index and the confinement factor in equation (D-3) are calculated on the unperturbed structure, where the grating region is replaced by their average value. Then the coupling coefficient can be obtained by equation (D-3).

Figure D-2 shows the calculated coupling coefficient as a function of the grating duty cycle for three different gratings. At small duty cycle, the semiconductor gratings have larger coupling coefficients because of the larger confinement factor of the grating. Nevertheless, the large contrast of refractive index in metal grating makes it quickly catch up and even exceed the other two at larger duty cycle.

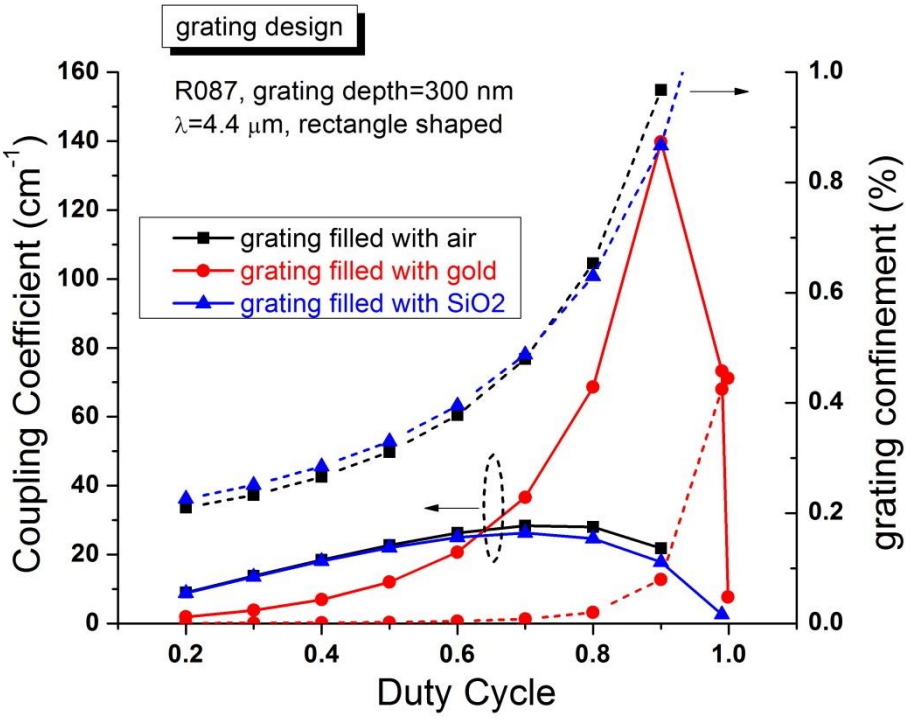


Figure D-2 Schematic of metal grating

References

[1] H. Kogelnik and C. V. Shank, "Coupled-Wave Theory of Distributed Feedback Lasers", J. Appl. Phys., **43**, 2327 (1972).

[2] K. David, J. Buus, G. Morthier, and R. Baets, "Coupling coefficients in gain-coupled DFB lasers: inherent compromise between coupling strength and loss", IEEE Photonics Technol. Lett., **3**, 439 (1991).

[3] W. Streifer, D. Scifres, and R. Burnham, "Coupling coefficients for distributed feedback single- and double-heterostructure diode lasers", IEEE J. Quantum Electron., **11**, 867 (1975).

[4] J. Carroll, J. Whiteaway, and D. Plumb, *Distributed feedback semiconductor lasers*. (The Institution of Electrical Engineers, London, United Kingdom, 1998).

Appendix E: Publication list

1) Z. Yin, **Y. Jiang**, Z. Tian, R. Q. Yang, T. D. Mishima, M. B. Santos, and M. B. Johnson, "Far-Field Patterns of Plasmon Waveguide Interband Cascade Lasers", IEEE J. Quantum Electron., **47**, 1414 (2011).

2) **Y. Jiang**, L. Li, Z. Tian, R. T. Hinkey, R. Q. Yang, T. D. Mishima, M. B. Santos, M. B. Johnson, et al., "Room-Temperature InAs-based Interband Cascade Lasers", paper CF3K.1, at The Conference on Lasers and Electro-Optics (CLEO), (2012).

3) L. Li, Z. Tian, **Y. Jiang**, H. Ye, R. Q. Yang, T. D. Mishima, M. B. Santos, and M. B. Johnson, "Interband Cascade Lasers at Long Wavelengths", CLEO, paper CF3K.2 (2012).

4) L. Li, L. Zhao, **Y. Jiang**, R. Q. Yang, J. C. Keay, T. D. Mishima, M. B. Santos, and M. B. Johnson, "Single-waveguide dual-wavelength interband cascade lasers", Appl. Phys. Lett., **101**, 171118 (2012).

5) Z. Tian, **Y. Jiang**, L. Li, R. T. Hinkey, Z. Yin, R. Q. Yang, T. D. Mishima, M. B. Santos, *et al.*, "InAs-Based Mid-Infrared Interband Cascade Lasers Near 5.3 μm", IEEE J. Quantum Electron., **48**, 915 (2012).

6) R. Q. Yang, L. Li, L. Zhao, **Y. Jiang**, Z. Tian, H. Ye, R. T. Hinkey, C. Niu, *et al.*, "Recent progress in development of InAs-based interband cascade lasers", Proc. SPIE, **8640**, paper 86400Q (2013).

7) **Y. Jiang**, L. Li, Z. Tian, H. Ye, L. Zhao, R. Q. Yang, T. D. Mishima, M. B. Santos, *et al.*, "Electrically widely tunable interband cascade lasers", J. Appl. Phys., **115**, 113101 (2014).

8) R. Q. Yang, L. Li, and **Y. Jiang**, "Interband Cascade Lasers: From Original Concept to Practical Devices", Progress in Physics, **34**, 169 (2014).

9) L. Li, H. Ye, **Y. Jiang**, R. Q. Yang, J. C. Keay, T. D. Mishima, M. B. Santos, and M. B. Johnson, "MBE-grown long-wavelength interband cascade lasers on InAs substrates", J. Cryst. Growth, **426**, 369 (2015).

10) **Y. Jiang**, L. Li, R. Q. Yang, J. A. Gupta, G. C. Aers, E. Dupont, J.-M. Baribeau, X. Wu, et al., "Type-I interband cascade lasers near 3.2 μm", Appl. Phys. Lett., 106, 041117 (2015).

11) J. Gupta, G. Aers, E. Dupont, J.-M. Baribeau, X. Wu, **Y. Jiang**, L. Li, R. Yang, et al., "Type-I Interband Cascade Lasers Near 3.2μm", CLEO, paper STu2G.3 (2015)

12) **Y. Jiang**, L. Li, H. Ye, R. Yang, T. D. Mishima, M. B. Santos, and M. B. Johnson, "Continuous-wave operation of InAs-based interband cascade lasers above room temperature", CLEO, Postdeadline Paper Digest,JTh5A.8 (2015)

13) **Y. Jiang**, L. Li, H. Ye, R. Q. Yang, T. D. Mishima, M. B. Santos, M. B. Johnson, D. J. Feng, *et al.*, "InAs-Based Single-Mode Distributed Feedback Interband Cascade Lasers", IEEE J. Quantum Electron., **51**, 2300307 (2015).

14) L. Li, **Y. Jiang**, H. Ye, R. Q. Yang, T. D. Mishima, M. B. Santos, and M. B. Johnson, "Low-threshold InAs-based interband cascade lasers operating at high temperatures", Appl. Phys. Lett., **106**, 251102 (2015).

15) H. Lotfi, L. Li, L. Lei, **Y. Jiang**, R. Q. Yang, J. F. Klem, and M. B. Johnson, "Short-wavelength interband cascade infrared photodetectors operating above room temperature", J. Appl. Phys., **119**, 023105 (2016).