INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

Optical/Electrothermal Simulation of Integrated Ridge Lasers

by

Gerardo Romo Luevano

**A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of requirements for the degree of**

Master of Applied Science

**Department of Electronics
Carleton University
Ottawa, Ontario, Canada**

August, 2001

© Gerardo Romo Luevano

**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-66895-9

Canada

The undersigned recommend to the Faculty of Graduate Studies
and Research acceptance of the thesis

“Optical/Electrothermal Simulation of Integrated Ridge Lasers”

submitted by Gerardo Romo Luevano, B. Eng.
in partial fulfilment of the requirements for the degree of

Master of Applied Science

Thesis supervisor

Chair, Department of Electronics

Carleton University

August, 2001

Acknowledgments

I would like to thank foremost my thesis supervisor Prof. Tom Smy. Thanks for your encouragement, expertise, guidance and unlimited fruitful discussions, which made all this work a truly fulfilling experience.

I would like to acknowledge the support of the Government of Canada for providing an award through the Government of Canada Award (GCA) Program, without which, this experience would not have been possible. I would also like to thank my program officer, Cristina Frias and other members of the International Council for Canadian Studies (ICCS) for their assistance and efficient administration of the program.

I would like to acknowledge the support of Nortel Networks for providing funding and material for this research, especially to Benoit Reid, for his continuous feedback and interesting discussions throughout the research process.

I would also like to thank the staff and faculty in the Department of Electronics at Carleton University, especially to Prof. David Walkey, my fellow research group members Dritan Celo and Lei Zhang for their help and participation in private discussions, as well as Helen Triantafellou for helping to edit this material.

Last but not least, I am deeply grateful to my wife, family and friends for their help and unconditional support.

Abstract

In this thesis a thorough thermal analysis of a ridge semiconductor laser structure is presented. For this purpose, two approaches are taken: first, a pure thermal examination, which neglects electrical and optical effects, is performed. Next, these physical phenomena are taken into consideration and an opto/electro/thermal self-consistent analysis is carried out. For performing the pure thermal analysis an existing computer program, *Atar*, is used, while for analyzing the laser structure self-consistently, a thermal/optical/electrical model is first developed. The model consists of the solution of the 3D heat equation and the 1D longitudinal carrier and photon rate equations.

The use of the pure thermal model to analyze the laser of interest led to structural alterations that predict a considerable reduction of the laser's active region temperature (17%). This reduction would be accomplished by modifying the current packaging technique from a wire bonded configuration to a flip chip configuration.

By using the self-consistent model, the effect of highly differing facet reflectivities on the longitudinal temperature distribution was assessed. For a laser with left and right facet reflectivities of 0.03 and 0.75 respectively, the self-consistent model predicted a temperature difference of approximately 15% from one end of the laser to the other. Additionally, facet heating in InGaAsP/InP and AlGaAs/GaAs lasers was also investigated. The results suggest that three dimensional heat flow modelling is required to avoid overestimations of the facet temperature. Finally, it was found that InP lasers are more resistant to thermal runaway due to their better thermal properties.

Table of Contents

Acknowledgments	iii
Abstract	iv
List of Figures	viii
List of Tables	xiii
List of Symbols and Abbreviations	xiv
1.0 Introduction	1
1.1 Thesis Motivation	1
1.2 Thesis Objectives	3
1.3 Thesis Organization	4
2.0 Introduction to Laser Operation	5
2.1 Introduction	5
2.2 p-n Junction	6
2.3 Heterojunction Lasers	8
2.3.1 Dielectric Waveguide	11
2.4 Recombination Mechanisms	12
2.4.1 Non-Radiative Recombination	13
2.4.1.1 Auger Recombination	13
2.4.1.2 Recombination at Defects	15
2.4.1.3 Surface Recombination	16
2.4.2 Radiative Recombination	17
2.5 Laser Operating Principles	18
2.5.1 Population Inversion	19
2.5.2 Optical Feedback	20
2.5.3 Optical Gain	21
2.5.4 Threshold Conditions and Laser Losses	23
2.6 Light-Current Characteristics	28
2.7 Principles of Laser Modeling	31

2.7.1	Basic 0-D (Lumped) Model	37
3.0	Pure Thermal Laser Model	41
3.1	Introduction	41
3.2	Distribution of Heat Sources in Semiconductor Lasers	43
3.3	Heat Conduction Equation	49
3.4	Solution of the Heat Conduction Equation	52
3.4.1	Analytical Solution	52
3.4.2	Numerical Solution	55
3.4.2.1	Atar- a 3D Thermal Simulator	55
3.4.2.1.1	3D Model Creation	56
3.4.2.1.2	Extraction of the Thermal Network	57
3.4.3	Comparison Between the Numerical and Analytical Solutions	58
4.0	Self-consistent Laser Model	67
4.1	Introduction	67
4.2	Optical Model	70
4.2.1	Scalar Modes	70
4.2.2	Optical Waveguide Analysis	71
4.2.3	1-D Optical Model	75
4.3	Electrical Model	77
4.3.1	Diffusion of Carriers	77
4.3.2	Non-Diffusion Equation	80
4.4	The overall model	80
4.4.1	Solution Algorithm	86
4.4.2	Convergence Criteria and Stability	89
5.0	Applications of the Laser Model	92
5.1	Introduction	92
5.2	Analysis of a Stand-Alone Symmetric Laser Structure	93
5.2.1	Pure Thermal Approach	95
5.2.2	Assessment of the Importance of Diffusion of Carriers in Self-Consistent Simulations	113
5.2.3	Self-Consistent Approach	119

5.2.4	Self-Consistent vs. Pure Thermal	124
5.3	Analysis of a Stand-Alone Asymmetric Laser Structure.....	125
5.3.1	Self-Consistent Approach	126
5.3.2	Self-Consistent Analysis of the Top Metal	131
5.4	Thermal Runaway Analysis.....	134
5.4.1	Overview	134
5.4.2	Modeling the Heat Flow - Dimensionality	139
5.4.3	Non-Linear Thermal Conductivity and Joule Heating.	147
5.4.4	Comparison Between an InP and a GaAs Laser	150
5.5	Analysis of Semiconductor Lasers in Optical Integrated Environments	152
6.0	Conclusions.....	158
6.1	Future Work.....	160
	References.....	162
 Appendices		
A.	Calculation of the Coefficients of the Analytical Solution of the Heat Equation	165
B.	Temperature Dependence of the Transparency Carrier Density.....	167
C.	Solution of the Diffusion Electrical Equation	169
D.	Solution of the Non-Diffusion Electrical Equation	175

List of Figures

2.1: Heavily doped p - n junction (a) in equilibrium and (b) with forward bias	7
2.2: Diagram illustrating the action of (a) single and (b) double hetero-junction structures in confining the carriers and radiation to the gain region	10
2.3: Dielectric waveguiding in a heterostructure semiconductor laser	11
2.4: Band-to-band Auger process shown schematically	14
2.5: Schematic illustration of (a) spontaneous-emission and (b) stimulated-emission processes	17
2.6: Population of a two level energy system: (a) in thermal equilibrium and (b) after a population inversion has been produced	20
2.7: Radiation passing through a volume element of length Δx and unit cross-sectional area.	23
2.8: Schematic illustration of a semiconductor laser and its associated Fabry-Perot (FP) cavity	26
2.9: Typical light current (L-I) curve for a semiconductor laser	29
2.10: Variation of the carrier population with device current.	38
2.11: Variation of the gain with device current	39
2.12: Laser output power as a function of the applied current	40
3.1: Power balance in diode lasers	43
3.2: Schematical structure of a stripe geometry laser	53
3.3: Thermal parameterization of a typical GaAs/GaAlAs heterostructure laser	59
3.4: Original and discretized domain of the laser structure using Atar	60

3.5: Detailed view of the mesh in the active region and adjacent layers	61
3.6: Comparison between the analytic and Atar solutions of the lateral temperature profiles along the active region	62
3.7: Comparison between the analytic and Atar solutions of the transverse temperature profiles.....	63
3.8: Temperature contours in the vicinity of the active region.....	64
4.1: Mutual interactions among thermal, electrical and optical processes in semiconductor lasers	69
4.2: Simulated optical waveguide.....	72
4.3: Normalized intensity of the fundamental optical mode.....	73
4.4: Contour plot of the fundamental mode	74
4.5: Graphical representation of the self-consistent laser model	83
4.6: Longitudinal carrier density profile.....	84
4.7: Real light-current curve with thermal effects included	86
4.8: Solution Algorithm	87
5.1: Schematical representation of the simulated laser structure	93
5.2: Top view of the simulated structure showing the top metal recess.....	94
5.3: Laser-mounting structure.....	95
5.4: Discretized Atar model of the laser structure	98
5.5: Detailed view of the ridge and active region.....	99
5.6: Temperature profiles for different percentages of power in each heat source (a) longitudinal and (b) transverse.....	100
5.7: Schematical representation of the thermal problem showing the heat flow pattern	102

5.8: Longitudinal temperature profile for different thermal conductivity of the Silicon Dioxide layer	104
5.9: Heat flow pattern in the vicinity of the ridge and active region.	105
5.10: Atar model of the laser structure with the bump representing the wire on the edge	106
5.11: Longitudinal temperature profile for different locations of the bump representing the bonding wire.	107
5.12: Atar model of the laser structure with a stripe on the edge of the chip	108
5.13: Longitudinal temperature profiles for different stripe dimensions	109
5.14: Atar model of the (a) original and (b) modified laser structure.	111
5.15: Transverse temperature profiles of the original and modified laser structures	112
5.16: Typical longitudinal temperature profile for low power laser operation	114
5.17: Corresponding carrier density distributions for the temperature profile of Fig. 5.16 and different carrier mobility	115
5.18: Typical temperature profile for high power laser operation	117
5.19: Carrier density distribution (a) longitudinal and (b) detailed view	118
5.20: Longitudinal carrier density profile	121
5.21: Light intensity distribution of the forward $I^f(z)$ and backward $I^b(z)$ running waves	122
5.22: Heat power contribution of the different sources in the active region (account only for 50% of the total heat power)	123
5.23: Comparison between the self-consistent and pure thermal solutions	125
5.24: Longitudinal carrier density profile	126

5.25: Light intensity distribution of the forward $I^f(z)$ and backward $I^b(z)$ running waves for the asymmetric laser	127
5.26: Heat power contribution of the different sources for the optically asymmetric laser	129
5.27: Longitudinal temperature distribution for the asymmetric laser	130
5.28: Self-consistent temperature profile for different top-metal conditions . . .	132
5.29: Temperature profile for the laser structure of interest for different substrate material	133
5.30: AlGaAs/GaAs laser structure used to study facet heating.	137
5.31: Discretization in the facet region.	138
5.32: Schematical representation of the three different thermal approaches taken to simulate facet heating, (a) 1-D, (b) 2-D and (c) 3-D.	140
5.33: Temperature profiles for different dimensionality of the heat equation.	142
5.34: Atar temperature contour plots of the facet region of the laser. (a) 2D model (b) 3D model	144
5.35: Longitudinal carrier density distribution	145
5.36: Longitudinal photon density distribution.	147
5.37: Effect of Joule heating and non-linear thermal conductivity on the temperature distribution.	149
5.38: Longitudinal temperature distribution for different laser structures	150
5.39: Basic semiconductor laser on an OIE	153
5.40: Detailed view of a laser in an OIE with heat spreader.	154
5.41: Contour plot of the (a) the centered, (b) y -offset and (c) xy -offset structure.	155

5.42: Lateral temperature profile for different positions of the laser structure	156
5.43: Longitudinal temperature profile for different positions of the laser structure	157
C.1: One dimensional discretization of the electrical domain	173

List of Tables

2.1: Definitions of laser parameters	36
2.2: Typical parameter values for a 1.3 μm buried-heterostructure laser	37
3.1: Room-temperature thermal conductivities and thermal diffusivities of III-V binary compounds.	50
3.2: Variation of thermal conductivities of III-V binary compounds with temperature.	51
3.3: Summarized comparison between <i>Atar</i> and the Analytical model	66
5.1: Thermal conductivities used for the laser structure of Fig. 5.1	96
5.2: Parameters used for the self-consistent simulations	120
5.3: Correction values for the AlGaAs/GaAs laser.	139
5.4: Thermal conductivities used for the simulations	139

List of Symbols and Abbreviations

n	Electron density
n_{th}	Electron density at threshold
p	Hole density
n_o	Transparency carrier density
S	Total photon density
I^f	Photon density of the forward travelling wave
I^b	Photon density of the backward travelling wave
t	Time
I	Injected current
V_f	Applied voltage in forward bias
V_{pn}	Voltage drop at the p - n junction
I_{th}	Threshold current
$I(x)$	Beam irradiance
E	Electric field
I_L	Leakage current
J	Current density
R_s	Laser's series resistance
e	Electronic charge
d	Active region thickness

w	Active region width
L	Active region length
$R_{l,2}$	Left and right mirror reflectivities
P_{out}	Output power
$A_{nr}(T)$	Non-radiative recombination coefficient
$B(T)$	Radiative recombination coefficient
$C(T)$	Auger recombination coefficient
Γ	Optical confinement factor
γ	Net absorption coefficient
$g(n, T)$	Optical gain
a	gain constant
G	Round trip gain
D	Diffusion constant
c	Speed of light in vacuum
v_g	Speed of light in the lasing medium
V_{sur}	Surface recombination velocity
β	Spontaneous emission coupling coefficient
β_s	Scalar mode propagation eigenvalue
τ_{ph}	Photon lifetime
τ_e	Carrier recombination lifetime
α_{int}	Internal losses

α_m	Mirror losses
η	Active region refractive index
η_g	Group refractive index
$\eta_{sp, d, i}$	Spontaneous, differential and internal efficiencies
$g_{NR, J, A}$	Non-radiative, Joule and absorption heat power density
T	Temperature
κ	Thermal conductivity
R_{TH}	Thermal resistance
E_g	Energy gap
E_i	Energy of the i th level
E_F	Fermi level
λ	Emission wavelength
ν	Emission frequency
h	Planck's constant
k	Propagation constant
CW	Continuous wave
CD	Catastrophic degradation
LOES	Longitudinal Optical and Electrical Simulator
GG	Gain Guided
IG	Index Guiding
R-K	Runge-Kutta

Chapter 1

Introduction

1.1 Thesis Motivation

In spite of the remarkable progress achieved in the fabrication of semiconductor lasers over the last several years, their continuous wave (CW) performance and integration are still seriously limited by their thermal behavior. Parameters like output power, threshold current, device lifetime and emission wavelength, which are relevant for reliable laser operation, are greatly influenced by the device temperature. In addition, under high power laser operation, excessive heating in the vicinity of the facets can lead to catastrophic optical damage of the device. Thus, the understanding of the thermal behavior of these devices such as the identification of the different heat generation mechanisms, their locations and contributions, as well as the structural optimization of laser chips and their mounting techniques, are key actions toward the improvement of their performance.

To investigate the thermal properties of semiconductor lasers, many numerical and analytical models have been developed, and both purely thermal and self-consistent models incorporating electrical and optical interactions simultaneously, have appeared in the literature. While analytical models are difficult to develop due to many inherent complications discussed later on, most numerical models reduce the simulation domain to a small two-dimensional area around the active region, where the transverse electrical, optical and thermal variations are strong [1], [2]. While this approach provides useful information

about the electrical and optical behavior of the laser, the excessive structural simplifications and thermal boundary conditions applied to the small simulation domain are not adequate from a thermal point of view. Furthermore, these models lack information with respect to the longitudinal distribution of the optical intensity and the way in which this, and the reflectivities of the right and left facets of the laser, affect the longitudinal temperature distribution. Longitudinal spatial variations are specially important when the reflectivity of the right and left facets greatly differ from each other thereby making the laser highly optically asymmetric and the longitudinal variations more important.

When the laser structure is studied for optimization, the temperature and heat flow in the entire simulation domain need to be monitored for an investigation of the effects of material parameters variations and/or structural modifications, which may take place far from the active region. Under this scenario, a larger scale model (3D) is required and the *longitudinal* variations of the optical and electrical parameters need only be modeled as they occur in the same scale as that of the thermal problem ($> 100 \mu\text{m}$). The lateral variations, however, take place over a short distance ($< 2 \mu\text{m}$) and can therefore be averaged when modeling the large scale problem.

The issue of modeling facet heating in semiconductor lasers has also been addressed by several authors, and laser models including longitudinal spatial variations are specially suited for this purpose. In [3], a fully one-dimensional self-consistent model for investigating facet heating was presented. Unfortunately, one-dimensional models greatly overestimate facet heating and, hence, 3D thermal models should be used instead, if more accurate results are desired. Existing three-dimensional thermal models [4] reduce the

simulation domain to a small area in the vicinity of the facets, called the catastrophic degradation (CD) area, thereby limiting the investigation of this effect from a more general perspective that includes the scale of the real thermal domain.

1.2 Thesis Objectives

This thesis is the result of a project carried out in collaboration with the Laser Group of Nortel Networks; as such, the objective behind this work can be stated as the detailed thermal analysis of a ridge laser structure and the proposal of possible structural modifications toward its thermal optimization. The thermal issues of interest were mainly two: *i)* the investigation of the laser structure to understand its thermal behavior and find viable ways to decrease the temperature in the active region of the device, and *ii)* the investigation of the effects of the highly asymmetric optical intensity distribution (which is induced by the differing facet reflectivities) on the longitudinal temperature distribution.

Since no commercially available simulator offered the flexibility that was needed to address this problem from the desired perspective, a 3D pure thermal model needed first to be developed. In addition, thermal effects do not occur in isolation but are rather influenced by the interactions with the electrical and optical phenomena, thus the incorporation of these effects needed also to be targeted. In summary, the objective was to develop a numerical laser model which, in compliance with the principles of laser operation, was capable of performing a thorough 3D thermal analysis of the laser of interest without resorting to excessive structural simplifications and yet capable of dealing with the electrical and optical interactions in a self-consistent way.

As a by-product of the necessity of implementing a longitudinal self-consistent laser model to study the previously mentioned thermal problems, this thesis was also intended to study facet heating to gain a better understanding of the thermal runaway problem.

1.3 Thesis Organization

This thesis is organized in the following manner. Chapter 2 introduces the basic principles of semiconductor laser operation and reviews most of the physical concepts and nomenclature necessary to follow up the material presented in the remaining chapters. A typical approach to modeling semiconductor lasers is also presented with the aid of a lumped model which is often found in the literature. Chapter 3 introduces the concepts related to thermal modeling of semiconductor lasers, and describes the temperature dependent phenomena that determine their thermal behavior. The different heat sources existing in semiconductor lasers are described, and a 3D thermal tool for investigating laser heating is introduced. Chapter 4 presents a self-consistent laser model comprised of the solution of the 3D thermal equation and the 1D longitudinal electrical and optical rate equations. A detailed solution algorithm used to solve these coupled equations is also presented. In Chapter 5, the laser structure of interest is analyzed by using both the pure thermal model described in Chapter 3 and the self-consistent model of Chapter 4. Additionally, the self-consistent model is utilized to investigate facet heating in two different laser structures and the results are compared. Finally, in Chapter 6, conclusions and recommendations for future work are provided.

Chapter 2

Introduction to Laser Operation

2.1 Introduction

The purpose of this chapter is to review the basic principles of semiconductor laser operation. Most of the physical concepts introduced here will be referred to in successive chapters of this thesis. The main equations that describe the electrical and optical behavior are introduced and these equations form the basis of the models to be described. At the end of the chapter, a basic introduction to laser modeling is also provided. The model that is presented here is the simple lumped model that is usually found in the available literature. This model is described here because in one way or the other it represents the starting point for more elaborate models such as the one that will be presented in Chapter 4. For the purpose of this thesis, the basic equations are modified to include longitudinal spatial variations and temperature dependence of most of the parameters as will be described in more detail in subsequent chapters.

The basic principles and construction of a laser are relatively straightforward; a rigorous analysis of the physics of the laser, on the other hand, is quite difficult and out of the scope of this work. The treatment that is given below is somewhat simplified and has the sole intention of providing the reader with the basic concepts necessary to follow up the material in the remaining chapters.

Three things are required to operate a laser: (i) a gain medium that can amplify the electromagnetic radiation propagating inside of it, (ii) spontaneous-emission of radiation and (iii) a confinement of the electromagnetic field through well defined optical modes that produces optical feedback of the gain. To provide the optical gain, a semiconductor laser needs to be externally pumped, and both electrical and optical techniques can be used for this purpose. A simple, practical, and most commonly used method employs current injection through the use of a forward bias p - n junction.

2.2 p - n Junction

The heart of a semiconductor laser is the p - n junction. The p - n junction provides the active gain medium, thus to obtain laser action we need only meet the other necessary requirements of radiation amplification and optical feedback. A p - n junction is formed by bringing a p -type and an n -type semiconductor into contact with each other. When they first come into contact, their quasi-Fermi levels do not match since the two are not in equilibrium. An equilibrium is, however, quickly established through diffusion of electrons from the n side to the p side, while the reverse occurs for holes. These diffusing electrons and holes recombine in the junction region. Eventually, a steady state is reached in such a way that further diffusion of electrons and holes is opposed by the built-in electric field across the p - n junction arising from the negatively charged acceptors on the p side and the positively charged donors on the n side. At this point, the Fermi level becomes continuous across the p - n junction as shown in Fig. 2.1, where the shaded areas represent the filled states.

When a p - n junction is forward-biased by applying an external voltage V_f , the built-in electric field is reduced, thereby making possible a further diffusion of electrons and holes across the junction. As Fig. 2.1b shows, in a narrow depletion region both electrons and holes are present simultaneously and can recombine either radiatively or non-radiatively. This narrow zone is called the *active region*.

Figure 2.1: Heavily doped p - n junction (a) in equilibrium and (b) with forward bias.

In the case of materials such as GaAs which have a direct bandgap, the electrons and holes have a high probability of recombining radiatively. Photons of energy $h\nu \equiv E_g$ are emitted during radiative recombination. The recombination radiation produced may interact with valence electrons and be absorbed, or interact with electrons in the conduction band, thereby stimulating the production of further photons of the same frequency. When the external current exceeds a critical value, the injected carrier concentration becomes large enough, and a condition known as *population inversion* is achieved ($n_c > n_v$); at this point, the rate of photon emission exceeds that of absorption. The p - n junction is then able to amplify the electromagnetic radiation and is said to exhibit *optical gain*. However, for a homojunction the thickness of the region where gain is sufficiently large, is very small ($\sim 0.01 \mu\text{m}$) and as there is no confinement mechanism, the charge carriers diffuse away from the active region and high threshold current densities result. The threshold current density, whose meaning will be explained in more detail later on this chapter, is an important parameter in semiconductor lasers and its minimization is usually sought during the designing process. The carrier confinement problem is solved through the use of a p - n heterojunction.

2.3 Heterojunction Lasers

As noted above, the resulting threshold current density for homojunction lasers is very large due to poor optical and carrier confinement. Dramatic reductions in the threshold current density at room temperature, coupled with higher efficiency, can be achieved using lasers containing heterojunctions. The properties of heterostructure lasers which

permit a low threshold current density and CW operation at room temperature can be illustrated with the double heterostructure (DH) laser shown in Fig. 2.2.

In this structure, a layer of GaAs, for example, is placed between two layers of the ternary compound $\text{Al}_x\text{Ga}_{1-x}\text{As}$ which has a wider energy gap than GaAs and a lower refractive index. Both $N-n-P$ and $N-p-P$ structures show the same behavior, where N and P represent the wider band gap semiconductor according to carrier type.

Fig. 2.2 also shows that carrier and optical confinement may be achieved simultaneously. The bandgap differences form potential barriers in both the conduction and valence bands which prevent electrons and holes injected into the GaAs layer from diffusing away.

(b)

Figure 2.2: Diagram illustrating the action of (a) single and (b) double heterojunction structures in confining the carriers and radiation to the gain region.

The GaAs layer then becomes the active region and it can be well defined, typically 0.1-0.3 μm in thickness. Similarly, the step change in refractive index provides a much more effective waveguide structure than was the case in homojunction lasers for the reasons explained in the next section. The radiation is, therefore, confined mainly to the active region. In addition, the fraction of the propagation mode which lies outside the active region is in a wider band gap semiconductor and is, therefore, not absorbed so that the losses are much smaller than in homojunction lasers.

2.3.1 Dielectric Waveguide

In double heterostructure semiconductor lasers, the optical confinement in the direction perpendicular to the junction occurs by virtue of a fortunate coincidence. The active layer with a smaller band gap also has a higher refractive index compared with that of the surrounding cladding layers (see Fig. 2.2). Because of the index difference, the active region in effect acts as a dielectric waveguide. The physical mechanism behind the confinement is total internal reflection, as illustrated in Fig. 2.3.

Figure 2.3: Dielectric waveguiding in a heterostructure semiconductor laser.

When a ray travelling at an angle θ (measured from the interface normal) hits the interface, it is reflected back if the angle θ exceeds the critical angle given by

$$\theta = \text{asin} \frac{\eta_1}{\eta_2} \quad (2.1)$$

where η_1 and η_2 are the refractive indices of the cladding and active layers respectively.

Thus, rays travelling nearly parallel to the interface are trapped and constitute the waveguide mode.

The confinement of the radiation field in the direction parallel to the junction may be achieved by different mechanisms. In the so called gain-guided (GG) semiconductor lasers, the waveguide is formed spontaneously due to the current spreading effect [5]. In this case, the waveguiding effect is dependent on the supply current, and a change of the excitation level of the laser results in unstable operation, however, the details are not important for the purpose of this thesis. Modern designs of stripe lasers usually use index guiding (IG). In these structures, the dominant factor affecting the waveguiding properties in the plane of the junction is the distribution of the real part of the complex refractive index, imposed by an appropriate internal waveguide built in the device by the designer. With such a waveguide, an increase in the supply current affects the waveguiding effect only slightly and the laser is more stable.

2.4 Recombination Mechanisms

When the current flowing through a semiconductor laser is increased, charge carriers (electrons and holes) are injected into the thin active region where they recombine through radiative or non-radiative mechanisms. As one may expect, non-radiative recombinations are not helpful for laser operation, and attempts are made to minimize their occurrence by controlling point defects and dislocations. However, a non-radiative recombination mechanism known as the Auger process, is intrinsic and becomes particularly important for long-wavelength semiconductor lasers operating at room temperature and above. Physically speaking, during an Auger recombination, the energy released by the electron-hole recombination is taken by a third charge carrier and is, eventually, lost to lat-

tice phonons. In the following sections, the various recombination mechanisms are briefly described.

2.4.1 Non-radiative Recombination

Non-radiative recombination mechanisms include (i) Auger recombination, (ii) recombination at defects, and (iii) surface recombination, among others. For long-wavelength semiconductor lasers, however, the Auger process is generally the predominant non-radiative recombination mechanism.

2.4.1.1 Auger Recombination

The Auger recombination process involves four particles states (three electrons and one hole, two electrons and two holes, and so forth). In this process, the energy released during the electron-hole recombination is transferred to another electron (or hole), which gets excited to a high energy state in the band. This electron or hole then relaxes back to achieve thermal equilibrium by losing its energy to lattice vibrations or phonons. The Auger recombination rate R_a may be approximately written as

$$R_a = C(T)n^3 \quad (2.2)$$

where n is the injected carrier density in an undoped semiconductor. The quantity $C(T)$ is called the Auger Coefficient and, in general, is temperature dependent [5].

It is generally accepted that Auger recombination can be a major non-radiative mechanism in narrow-gap semiconductors. The observed strong temperature dependence of threshold current in long-wavelength semiconductor is usually attributed to this effect.

The three major types are (i) band-to-band processes, (ii) phonon-assisted Auger processes, and (iii) trap-assisted processes. A detailed explanation of these processes is out of the scope of this thesis; however, as these processes play a very important role in semiconductor laser modeling, the band-to-band recombination process is briefly explained below to illustrate the nature of the process.

The basis of the Auger effect is the Coulomb interaction between two neighboring electrons in the conduction band. Electron 1 recombines with hole $1'$ (i.e. electron 1 makes a transition to empty state $1'$) and the excess energy is transferred to electron 2, which is excited to state $2'$. Electron $2'$ loses its energy to optical phonons when it relaxes back to thermal equilibrium. Fig. 2.4 shows schematically this recombination process.

Figure 2.4: Band-to-band Auger process shown schematically.

Band-to-band Auger processes are characterized by a strong temperature dependence and band gap dependence, i.e., the Auger rate decreases rapidly for low temperatures or for high band-gap materials.

2.4.1.2 Recombination at Defects

Defects in the active region of an injection laser can be formed in several different ways. In many cases they are grown-in during epitaxy. They can also occur, multiply, or propagate during a stress-aging test. In particular, defects can propagate along a specific crystal axis in a strained lattice.

Defects produce a continuum of states in a localized region. Electrons or holes that are within a diffusion length from the edge of the defect may recombine non-radiatively via the continuum of states. The rate of recombination at a defect or trap is usually written as

$$R_d = A_{nr}(T)n \quad (2.3)$$

where $A_{nr}(T)$ is given by

$$A_{nr}(T) = \sigma v N_t \quad (2.4)$$

where σ is the capture cross section of the trap, N_t is the density of traps, and v is the velocity of the electrons or holes. The quantity $A_{nr}(T)$ is the inverse of the non-radiative lifetime in the presence of a trap to account for non-radiative recombinations and, in general, is temperature dependent.

2.4.1.3 Surface Recombination

In an injection laser, the cleaved facets are surfaces exposed to the ambient. In addition, in many index-guided laser structures, the edges of the active region can be in contact with curved surfaces that may not form a perfect lattice. In general, an internal surface is a strong perturbation of the crystal lattice, creating many dangling bonds that can absorb impurities from the ambient. Hence, a high concentration of defects may occur on these surfaces, and the defects can act as centers of non-radiative recombination. When at the facets, such localized non-radiative centers, in addition to increasing the threshold current, may cause other performance problems in lasers such as thermal run-away. Another possible source for surface recombination is the heterostructure interface between the active and cladding layers.

The recombination rate of carriers at the surface is expressed in terms of a surface recombination velocity. If A is the surface area and n_{th} the threshold carrier density, then the surface recombination rate is given by

$$R_{sur} = V_{sur} A n_{th} \quad (2.5)$$

where V_{sur} is the surface recombination velocity. For an InP surface exposed to air V_{sur} is about two orders of magnitude smaller than that for GaAs. For the modeling of semiconductor lasers, it is customary to lump defect and surface recombination, and introduce a single parameter $A_{nr}(T)$.

2.4.2 Radiative Recombination

During a radiative recombination process, the energy E_g released by the electron-hole pair appears in the form of a photon whose frequency ν or wavelength λ satisfies the energy conservation relation $E_g = h\nu = hc/\lambda$. This can happen through two distinct ways: (a) the *spontaneous emission* process in which the electron drops to the lower level in an entirely random way and (b) the *stimulated emission* process in which the electron is triggered to undergo the transition by the presence of photons of the same energy. These processes are illustrated in Figs. 2.5 (a) and (b).

Figure 2.5: Schematic illustration of (a) spontaneous-emission and (b) stimulated-emission processes.

Under normal circumstances, the stimulated process is negligible because the probability of spontaneous process occurring is much higher. Because the spontaneous radiation from any atom is emitted at random, the radiation emitted by a large number of atoms clearly will be incoherent. In contrast to this, the stimulated emission process results in coherent radiation as the waves associated with the stimulating and stimulated photons have identi-

cal frequencies, are in phase, and have the same state of polarization. This means that with stimulated emission, the amplitude of an incident wave can grow as it passes through a collection of atoms in what is clearly an amplification process.

The total spontaneous radiative recombination rate R_{sp} can be approximated by

$$R_{sp} = B(T)np \quad (2.6)$$

where $B(T)$ is the radiative recombination coefficient and n and p are the electron density and the hole density respectively [5]. For undoped semiconductors, the previous equation becomes

$$R_{sp} = B(T)n^2. \quad (2.7)$$

2.5 Laser Operating Principles

Although stimulated emission can occur as soon as current is applied to the semiconductor laser, the laser does not emit coherent light until the current exceeds a critical value, known as the *threshold current* (I_{th}). This is so because stimulated emission has to compete against the absorption processes during which an electron-hole pair is generated at the expense of an absorbed photon. Since the electron population in the valence band generally far exceeds that of the conduction band, absorption dominates. At a certain value of the external current, a sufficient number of electrons become present in the conduction band in order to make the semiconductor optically transparent. With a further increase in current, a *population inversion* ($n_c > n_v$) occurs and the active region of the semiconductor

laser exhibits optical gain and amplifies the electromagnetic radiation passing through it. Spontaneously emitted photons serve as the noise input for the amplification process.

2.5.1 Population Inversion

The population inversion condition required for light amplification is a non-equilibrium distribution of atoms among the various energy levels of the atomic system. The Boltzmann distribution (appropriate for regions far from E_f), which applies to a system in thermal equilibrium, is given by:

$$n_j = \frac{g_j n_a \exp(-E_j/k_B T)}{\sum_i g_i \exp(-E_i/k_B T)} \quad (2.8)$$

where n_j , g_j and E_j are the population, degeneracy and energy of the j th level respectively, n_a is the total number of atoms, k_B is Boltzmann's constant and T is temperature. The population of a two level energy system is illustrated in Fig. 2.6. It is clear that as E_j increases, n_j decreases for a constant temperature. We note that if the energy difference between E_1 and E_2 was nearly equal to $k_B T$ (approximately 0.025 eV at room temperature), then the population of the upper level would be $1/e$ or 0.37 of that of the lower level. For an energy difference large enough to produce visible radiation (approximately 2.0 eV), however, the population of the upper level is almost negligible. Clearly then, if we are to create a population inversion, we must supply a large amount of energy to excite atoms into the upper level E_2 . This excitation process is called *pumping* and produces the required non-thermal equilibrium situation.

Figure 2.6: Population of a two level energy system: (a) in thermal equilibrium and (b) after a population inversion has been produced.

However, optical gain alone is not enough to operate a laser. The other necessary ingredient is optical feedback.

2.5.2 Optical Feedback

In the simplest semiconductor lasers, optical feedback is provided by the cleaved facets that form a Fabry Perot (FP) cavity [6]. The role of the FP cavity is twofold. First, it provides a direction selectivity for the process of stimulated emission, since only photons travelling along its axis are reflected back and forth. Second, it provides a wavelength selectivity, since the feedback is strongest for wavelengths corresponding to the longitudinal modes of the FP cavity.

One source of radiation loss is absorption in the facets. To reduce such a loss, high reflectivity, multilayer dielectric coatings are used rather than metallic coatings. In the so-

called multilayer stacks, there is a sequence of quarter-wave (i.e. $\lambda/4$) layers of alternate high and low refractive index dielectric materials. Because of the phase changes occurring at alternate interfaces, all the reflected waves are in phase and add constructively. More than twenty such layers may be needed to give reflectivities in excess of 99.9%. Clearly, this feedback mechanism will only be effective over a narrow wavelength range.

Because of the optical feedback, the number of photons travelling perpendicular to the facets increases when the current is large enough to satisfy the condition of net stimulated emission. However, some photons are lost through the partially transmitted facets, and some are scattered or absorbed inside the cavity. If the loss exceeds the gain, stimulated emission cannot sustain a steady supply of photons. This is precisely what happens below threshold, when the laser output consists of mainly spontaneously emitted photons.

2.5.3 Optical Gain

The gain is perhaps one of the most crucial parameters of interest in semiconductor lasers. A rigorous expression for gain in a semiconductor laser can be developed from quantum-mechanical theory considerations. A realistic analysis is extremely complex for several reasons. The inclusion of all the decay processes is usually difficult. Further, the description of the system requires knowledge of the band structure and density of states in the conduction and valence bands. A further complication arises in a semiconductor laser because the active region is often heavily doped and the band-tail effects become important [7].

An alternative approach has been used to calculate the optical gain. This approach is based on a generalization of Einstein's A and B coefficients [8], which relate the rate of spontaneous and stimulated emissions to the net gain or absorption coefficient. A shortcoming of this approach is that it provides only the small signal gain, while the gain saturation effects cannot be treated. However, the approach is useful for studying the dependence of the gain on various parameters such as the doping levels and the carrier density.

In view of the above mentioned difficulties, a phenomenological approach is generally used for describing a semiconductor laser, and has proven to be extremely successful. It is based on the observation that the numerically calculated gain at the lasing frequency varies almost linearly with the injected carrier density n for all values of J . The gain g can, therefore, be approximated by

$$g(n, T) = a(n - n_o(T)) \quad (2.9)$$

where a is the gain coefficient and $n_o(T)$ is the carrier density required to achieve transparency (corresponding to the onset of population inversion). Both of the parameters can be estimated based on numerical calculations or determined experimentally. Note that the product $an_o(T)$ is just the absorption coefficient of the unpumped material. In general, $g(n, T)$ can have a wide range of values depending on the intrinsic properties of the active medium.

We can now derive a threshold condition in terms of the parameters of the entire system for laser oscillations to occur.

2.5.4 Threshold Conditions and Laser Losses

Let us consider a collimated beam of perfectly monochromatic radiation of unit cross-sectional area passing through an active medium as in Fig. 2.7. We assume for simplicity that there is only one relevant electron transition, which occurs between the energy levels E_1 and E_2 , as illustrated in Fig. 2.6.

The change in irradiance of the beam as a function of distance is given by

$$\Delta I(x) = I(x + \Delta x) - I(x) \quad (2.10)$$

For a homogeneous medium, $\Delta I(x)$ is proportional both to the distance travelled Δx and to $I(x)$. That is, $\Delta I(x) = -\gamma I(x) \Delta x$, where the constant of proportionality, γ , is the *absorption constant*.

Figure 2.7: Radiation passing through a volume element of length Δx and unit cross-sectional area.

The negative sign indicates the reduction in beam irradiance due to absorption if γ is a positive quantity, or an increase due to optical gain if γ is a negative number. Writing this expression as a differential equation results in the following:

$$\frac{dI(x)}{dx} = -\gamma I(x), \quad (2.11)$$

integration of this equation gives

$$I(x) = I_0 e^{-\gamma x} \quad (2.12)$$

where I_0 is the incident irradiance.

A steady state level of oscillation is reached when the rate of amplification is balanced by the rate of loss. Thus, while a population inversion is a necessary condition for laser action, it is not a sufficient one because the minimum or threshold value of gain must be large enough to overcome the losses and sustain oscillations. The threshold gain, in turn, specifies the minimum population inversion required.

The total loss of the system is due to a number of different processes. The most important ones include:

- (1) transmission at the mirrors. The transmission from one of the mirrors usually provides the useful output, while the other mirror is made as reflective as possible to minimize losses;
- (2) absorption and scattering at the mirrors;
- (3) absorption in the laser medium due to transitions other than the desired ones (most laser media have many energy levels not all of which will be involved in the laser action);
- (4) scattering at optical inhomogeneities in the laser medium;
- (5) diffraction losses at the mirrors.

To simplify matters, let us include all the losses except those attributed to transmission at the mirrors in a single effective loss coefficient α_{int} . The net absorption coefficient can then be expressed as the difference between the effective loss coefficient and the gain as follows:

$$\gamma = \alpha_{int} - \Gamma g(n, T), \quad (2.13)$$

The constant factor Γ is introduced here phenomenologically to account for the reduction in gain that occurs because of the spreading of the optical mode beyond the active region. It is known as the *confinement factor* or the filling factor, and represents the fraction of the mode energy contained in the active region.

The threshold gain can be determined by considering the change in irradiance of a beam of light undergoing a round trip within the laser cavity. We assume that the laser medium fills the space between the mirrors M_1 and M_2 which have reflectances R_1 and R_2 , and a separation L as in Fig. 2.8. The thin central region provides the optical gain. The cleaved facets form a FP cavity and provide the feedback for the onset of laser oscillations. Then, in travelling from M_1 to M_2 the beam irradiance increases from I_0 to $I(L)$ where, from Eq. 2.12

$$I(L) = I_0 e^{-(\alpha_{int} - \Gamma g)L} \quad (2.14)$$

After reflection at M_2 , the beam irradiance will be $R_2 I_0 \exp(\Gamma g - \alpha_{int})L$, and after a complete round trip the final irradiance will be such that the round trip gain G is

$$G = \frac{Final}{Initial} = R_1 R_2 e^{2(\Gamma g - \alpha_{int})L} \quad (2.15)$$

Figure 2.8: Schematic illustration of a semiconductor laser and its associated Fabry-Perot (FP) cavity.

If G is greater than unity, a disturbance at the laser resonant frequency will undergo a net amplification and the oscillations will grow; if G is less than unity, the oscillations will die out. Therefore, we can write the *threshold condition* as:

$$G = R_1 R_2 e^{2(\Gamma g_{th} - \alpha_{int})L} = 1 \quad (2.16)$$

where g_{th} is the threshold gain. It is important to realize that the threshold gain is equal to the steady state gain in continuous output lasers. Such equality is the result of a phenomenon known as *gain saturation* which can be explained as follows. Initially, when laser action commences, the gain may be well above the threshold value. The effect of stimulated emission, however, will be to reduce the population of the upper level of the laser transition so that the degree of population inversion and, consequently, the gain will decrease. Thus, the net round trip gain may vary and be greater than or less than unity, so that the cavity energy density will correspondingly increase or decrease. It is only when G has been equal to unity for a period of time that the cavity energy (and laser output power) settles down to a steady state value, that is when the gain just balances the losses in the medium.

In terms of the population inversion, there will be a threshold value n_{th} corresponding to g_{th} . In steady state situations, the population inversion must be equal to n_{th} regardless of the amount by which the threshold pumping rate is exceeded. Further increase of the population inversion with pumping is impossible in a steady state situation. This would result in a rate of induced energy emission which exceeds the losses, and the total energy stored in the cavity would increase with time in violation of the steady state assumption.

The small signal gain required to support steady state operation depends on the laser medium through $g(n, T)$ and α_{int} , and on the laser construction through R_1 , R_2 and L .

From Eq. 2.16 we can see that

$$\Gamma g_{th} = \alpha_{int} + \alpha_m \quad (2.17)$$

where the first term represents the volume losses and the second the mirror loss, given by

$$\alpha_m = \frac{1}{2L} \ln\left(\frac{1}{R_1 R_2}\right) \quad (2.18)$$

2.6 Light-Current Characteristic

The light emitted by one facet of a semiconductor laser is measured as a function of the device current I . The resulting curve is often referred to as the light-current (L-I) curve and is strongly temperature dependent. Fig. 2.9 shows schematically the P_{out} vs. I behavior of a semiconductor laser at a given temperature.

The current I is related to the current density J that is injected into the active layer by

$$I = I_a + I_L = wLJ + I_L \quad (2.19)$$

where I_a is the current passing through the active region, L is the cavity length, and w is the active region width. The leakage current I_L accounts for the fact that, depending on a specific structure, a part of the total current may either not pass through the active layer (lateral leakage) or pass through it and leak over the potential barriers without recombination of carriers (vertical leakage). Eq. 2.19 is applicable for a broad area laser or a strongly index guided laser, for which the injected current density is nearly uniform over the active region width. If the current density is laterally non-uniform, J should be interpreted as the average injected current density over the width w .

Figure 2.9: Typical light current (L-I) curve for a semiconductor laser.

The form of the L-I curve is typical for any laser. The turning point, at which the light output abruptly starts increasing, corresponds to the laser threshold. When the applied current is below threshold, the light output mainly consists of spontaneous emission and its magnitude is small. When the current exceeds the threshold value, stimulated emission begins to dominate and the output power increases. In the vicinity of the threshold, the power increases by several orders of magnitude.

Once threshold is reached, the carrier density n remains pinned to its threshold value n_{th} ; a further increase in J leads to light emission through the process of stimulated

emission and the output power varies linearly. The slope dP_{out}/dI is a measure of the device's optical efficiency. It is found that the slope efficiency does not remain constant and the output power saturates for large values of I . Three factors may contribute to power saturation: (i) The leakage current may increase with I so that a smaller fraction of the device current leads to current injection into the active layer; (ii) The value of I_{th} is current dependent (by means of temperature) and increases with I . This increase may be attributed to Auger recombination, which, as previously mentioned, is highly temperature dependent; (iii) The internal loss α_{int} increases with I so that a smaller fraction of generated photons is useful as the output power.

Power saturation can be explained in the following way. When the input current is increased, additional heat is produced by Joule heating and non-radiative recombination processes (which increase with carrier density). In turn, the additional amount of heat raises the temperature inside the laser and this results in an increase of the non-radiative recombination processes and more heat production. This loop continues until a steady state condition is reached. The increase in the non-radiative recombination rate implies that more carriers are lost to heat and less contribute to the generation of photons. The overall effect is a diminishment of the laser efficiency and a saturation of power. Due to the cubic dependence on the carrier density of the Auger recombination process, this mechanism plays a very important role in power saturation.

2.7 Principles of Laser Modeling

Semiconductor lasers are often used as sources in optoelectronic computing and communication systems. The electrical and optical properties of the active region of a laser can be described by the well known rate equations which describe the rate of change of the electron and photon densities inside the active region in terms of physical parameters. The first rate equation relates the carrier density n to the pumping parameter, the carrier density J . The relation is accomplished through a rate equation that incorporates all the mechanisms by which the carriers are generated or lost. In general, the continuity equations for both electrons and holes should be considered. The two are interrelated because of charge neutrality, and it suffices to consider one rate equation for electrons. In its general form, the carrier density rate equation is

$$\frac{\partial n}{\partial t} = D(\nabla^2 n) + \frac{J}{ed} - R(n) \quad (2.20)$$

The first term on the right hand side of Eq. 2.20 accounts for carrier diffusion, and D is the diffusion coefficient. The second term governs the rate at which the carriers, electrons or holes, are injected into the active layer because of the external pumping. The electron and hole populations are assumed to be the same to maintain charge neutrality. In the second term, e is the magnitude of the electron charge and d is the active layer thickness. Finally, the last term $R(n)$ takes into account the carrier loss owing to various recombination processes, both radiative and non-radiative. In general, carrier diffusion plays an important role in semiconductor lasers and complicates their analysis significantly. Device geometry is important, however, and diffusion effects are sometimes of minor nature. This is the

case for index guided semiconductor lasers where the active region dimensions (in the plane perpendicular to the cavity axis) are often small compared to the diffusion length. Since the carrier density does not vary significantly over the active-region dimensions, it can be assumed to be approximately constant, and the diffusion term in Eq. 2.20 can be neglected. In the steady state, $\partial n / \partial t = 0$ and Eq. 2.20 becomes

$$J = edR(n) \quad (2.21)$$

The situation is entirely different for gain-guided semiconductor lasers where the lateral variation (parallel to the heterojunction) of the carrier density arising from carrier diffusion makes the gain in Eq. 2.9 spatially inhomogeneous and helps to confine the optical mode. The inclusion of the diffusion term in those cases is then necessary.

The threshold condition (Eq. 2.17) together with Eqs. 2.9 and 2.21 can be used to model the light current characteristics of semiconductor lasers. To complete the description of the carrier rate equation, we need a suitable form of the carrier recombination rate $R(n)$. As previously explained, the charge carriers recombine through several radiative and non-radiative mechanisms. A suitable form for $R(n)$ is

$$R(n, T) = A_{nr}(T)n + B(T)n^2 + C(T)n^3 + R_{st}N_{ph} \quad (2.22)$$

where it is assumed that the doping level of the active layer is well below the injected carrier density. The linear term $A_{nr}(T)n$ is due to non-radiative recombination at defects; the quadratic term $B(T)n^2$ is due to spontaneous radiative recombination; the cubic term $C(T)n^3$ is due to Auger recombination. The last term, $R_{st}N_{ph}$, is due to stimulated

recombination that leads to coherent emission of light. This term is directly proportional to the intracavity photon density N_{ph} given by

$$N_{ph} = \frac{(J - J_{th})}{v_g(\alpha_m - \alpha_{int})ed} \quad (2.23)$$

where $v_g = c/\eta_g$ is the group velocity, η_g is the group refractive index, and c is the speed of light in a vacuum. R_{st} is the net rate of stimulated emission

$$R_{st} = (c/\eta_g)g(n, T) \quad (2.24)$$

where $g(n, T)$ is the optical gain given by Eq. 2.9.

An expression for the threshold current density can be obtained by using this phenomenological model. As J increases, the gain also increases owing to an increase in the carrier density n . The threshold is reached when n achieves the threshold value n_{th} for which the gain equals loss. Using Eqs. 2.17 and 2.9, the threshold carrier density is given by

$$n_{th} = n_o + (\alpha_m + \alpha_{int})/(a\Gamma) \quad (2.25)$$

where n_o is the transparency value at which population inversion occurs. The threshold current density J_{th} is obtained using Eqs. 2.21 and 2.22. Near threshold, stimulated recombination can be neglected and, therefore, obtain

$$J_{th} = \frac{edn_{th}}{\tau_e(n_{th})} \quad (2.26)$$

where

$$\tau_e(n, T) = (A_{nr}(T) + B(T)n + C(T)n^2)^{-1} \quad (2.27)$$

is the carrier recombination time that, in general, depends on n and T .

The second rate equation is the photon rate equation

$$\frac{\partial S}{\partial t} = \beta B n^2 + v_g \Gamma g(n) S - \frac{S}{\tau_{ph}} \quad (2.28)$$

The rate equation for the photon density describes the various mechanisms by means of which photons are added to the lasing mode. The spontaneously emitted photons ($B(T)n^2$) are emitted randomly into all 4π steradians of space. The first term in Eq. 2.28 represents the fraction β of spontaneously generated photons that are coupled into the lasing mode. The second term models the increase in photons due to stimulated emission. Finally, the last term represents photon losses due to internal mechanisms within the laser cavity. The photon lifetime τ_{ph} incorporates the effects of mirror losses due to the facets reflectivities, and losses due to intrinsic internal absorption mechanisms α_{int} . It is defined as

$$\tau_{ph} = \frac{1}{v_g(\alpha_m + \alpha_{int})} \quad (2.29)$$

where it can be seen that the facet losses are distributed over the cavity length L . Note also that the previously defined rate equations apply for a *single-mode* laser. If multimode oscillation is to be considered, then a separate set of rate equations is required for each mode [9].

To complete the model, the output power from one of the facets is given by the following relation

$$P_{out} = \frac{h\nu}{e} \frac{\alpha_m}{\alpha_m + \alpha_{int}} \frac{\sqrt{R_1}}{\sqrt{R_2} + \sqrt{R_1}} \frac{1 - R_2}{1 - \sqrt{R_1 R_2}} (I - I_{th}). \quad (2.30)$$

Finally, the threshold current I_{th} is obtained using Eqs. 2.19 and 2.26; it is given by

$$I_{th} = \frac{eVn_{th}}{\tau_e} + I_L \quad (2.31)$$

where $V=Lwd$ is the active volume.

Table 2.1 summarizes the laser parameters described in this section.

Table 2.1: Definitions of laser parameters.

Parameter	Definition	Units
n	Electron density	m^{-3}
n_o	Transparency carrier density	m^{-3}
S	Photon density	m^{-3}
t	Time	s
I, I_L	Injected and leakage currents	A
J	Injected current density	A/m^2
e	Electronic charge	C
d, w, L	Active region thickness, width and length	m
$R_{1,2}$	Mirror reflectivities	--
P_{out}	Output power	W
$A_{nr}(T)$	Non-radiative recombination coefficient	s^{-1}
$B(T)$	Radiative recombination coefficient	m^3/s
$C(T)$	Auger recombination coefficient	m^6/s
Γ	Optical confinement factor	--
$g(n, T)$	Optical gain	m^{-1}
c	Speed of light in vacuum	m/s
v_g	Speed of light in the lasing medium	m/s
β	Spontaneous emission coupling coefficient	--
τ_{ph}	Photon lifetime	s
τ_e	Carrier recombination time	s
α_{int}, α_m	Internal losses, mirror losses	m^{-1}
η, η_g	Refractive index, group refractive index	--

2.7.1 Basic 0-D (Lumped) Model

In order to illustrate the model's predictions, the previously described steady state equations were solved to investigate a typical buried-heterostructure InP/InGaAsP laser emitting at 1.3 μm . For simplicity, the temperature dependence of the parameters involved was neglected. The parameter values used for the simulation are shown in Table 2.2.

Table 2.2: Typical parameter values for a 1.3 μm buried-heterostructure laser.

Parameter	Definition	Value
L	Cavity length	250 μm
w	Active-region width	2 μm
d	Active-layer thickness	0.2 μm
Γ	Optical confinement factor	0.3
α_m	Facet loss	45 cm^{-1}
α_{int}	Internal loss	40 cm^{-1}
η_g	Group refractive index	4
$A_{nr}(300)$	Non-radiative recombination rate	1 x 10 ⁸ s ⁻¹
$B(300)$	Radiative recombination rate	1 x 10 ⁻¹⁰ cm ³ /s
$C(300)$	Auger recombination rate	3 x 10 ⁻²⁹ cm ⁶ /s
a	Gain constant	2.5 x 10 ⁻¹⁶ cm ²
$n_o(300)$	Transparency carrier density	1 x 10 ¹⁸ cm ⁻³
β	Spontaneous emission coupling coefficient	1 x 10 ⁻⁵
τ_{ph}	Photon lifetime	1.6 ps
$\tau_c(300)$	Carrier lifetime	2.2 ns

The plots of the carrier density and gain as functions of the applied current are shown in Figs. 2.10 and 2.11 respectively.

Figure 2.10: Variation of the carrier population with device current.

From these figures one can observe and confirm the previously described phenomena of carrier and gain saturation. As the current is injected into the device, the gain increases linearly with the carrier density according to Eq. 2.9 until threshold is reached. At this point, the gain becomes equal to the total losses and saturates. The carrier density reveals a similar behavior: it increases almost linearly with the applied current until threshold is reached; after this point, it remains constant at its threshold value. It is impor-

tant to note that the values shown in these figures represent the average values at each bias point.

Figure 2.11: Variation of the gain with device current.

Fig. 2.12 shows the predicted output power versus the injected current. As it was mentioned before, this lumped model predicts a linear relation and is only valid in low power operation where $T = T_{nom}$.

A real $L-I$ plot (see Fig. 2.9) shows that as the current is increased far above threshold, the output power starts to roll off as the temperature effects become more critical [10], [11]. This effect is not predicted by this model because the temperature dependence of the threshold current was not included. This is not necessarily a constrain if the use of this

model is limited to low power simulations (not far above threshold) where the thermal effects are not important and the output power has, indeed, a linear behavior.

Figure 2.12: Laser output power as a function of the applied current.

In the case that thermal effects require further detailed investigation for a wider range of applied current, temperature dependence of parameters has to be considered for a prediction of a more accurate behavior. The model needs then to be reformulated, as it will be explained in Chapter 4.

Chapter 3

Pure Thermal Laser Model

3.1 Introduction

Much has been accomplished in the fabrication of semiconductor lasers over the last several years. Unfortunately their continuous wave (CW) performance and integration are still seriously limited by their thermal behavior. Average heat-flux densities generated inside the active regions of semiconductor lasers are usually very high. This leads to a substantial increase in temperature, accompanied by a corresponding increase in the threshold current density, a shift in the emission wavelength, and a reduction of the optical output power. Efficient heat dissipation, along with low threshold, is therefore critical to improve the semiconductor lasers performance. In addition, since the operating lifetime of semiconductor lasers decreases exponentially with temperature, it is essential to design lasers with consistently low self-heating.

As it was shown in Chapter 2, the various physical phenomena that define the laser's behavior are quite complex and coupled with each other. In order to get a better prediction of the laser's behavior, all of those physical processes should be considered simultaneously in a self-consistent model. This approach will be taken in Chapter 4 and the different trade-offs that come into play under such scenario will be analyzed in some detail there. On the other hand, by studying the electrical, optical and thermal effects in isolation, although not completely accurate, still valuable information about the laser's

behavior can be obtained. The degree to which this approach is valid depends on the knowledge of the individual phenomena that is available since this determines how well they can be formulated within the framework of the phenomenon that wants to be studied in isolation. In other words, the thermal behavior of a semiconductor laser structure can be investigated from a strictly thermal point of view provided the heat power contributions from the electrical and optical phenomena are known. This comes down to saying that a pure thermal analysis can be performed if the location of the heat sources and the power dissipated by each of them are known a priori and the coupling with the optical and electrical effects is small.

This chapter introduces the general concepts related to thermal modeling of semiconductor lasers as well as the temperature dependent phenomena that determine their thermal behavior. The typical heat sources that appear in semiconductor lasers are described and emphasis is put on those which are included in the self-consistent model that is presented in the next chapter. A three dimensional thermal model is also introduced as a tool to investigate the thermal properties of these optical devices. This model is introduced here because it is used in Chapter 5 to investigate and optimize the thermal properties of a real laser structure from a purely thermal standpoint. Finally, by solving a simple laser structure, a comparison between the analytical and numerical approaches to solving the heat conduction equation is made.

3.2 Distribution of Heat Sources in Semiconductor Lasers

Diode lasers are supplied with the electric power which, in the processes of current spreading, carrier recombination as well as emission and absorption of radiation, is transformed into output power and heat in the way illustrated in Fig. 3.1.

Figure 3.1: Power balance in diode lasers.

The power balance in a semiconductor laser may be expressed in the following form:

$$P = V_f I = (V_{pn} + I R_s) I \quad (3.1)$$

or

$$P = I^2 R_s + I V_{pn} \quad (3.2)$$

where P is the electric power supplied to the laser, V_f is the total voltage drop inside the device, V_{pn} is the voltage drop at the p - n junction, I is the operating current and R_s is the device resistance. The first term on the right side of Eq. 3.2 represents Joule heating and the second term represents all other forms of heat generation and the output power.

To model the thermal properties of semiconductor lasers, the following heat generating mechanisms should be taken into consideration: Joule heating, non-radiative recombination of carriers, and absorption of spontaneous and laser radiation at defects and heterostructure interfaces (scattering). Distributions of these heat sources are directly associated with distributions of carrier concentration, current density, and spontaneous and stimulated radiation intensities, hence they are influenced by current spreading, carrier diffusion and waveguiding mechanisms. Their magnitudes are influenced in turn by such device and material parameters as internal quantum efficiency, threshold current density, threshold carrier concentration, absorption coefficients and electrical resistivities; all in general dependent on temperature.

The distribution of heat sources in a semiconductor laser depends upon the efficiency of transformation of the energies of the electric and electromagnetic fields into heat energy and upon the way in which these energies are transferred.

Non-radiative Recombination. One of the main heat sources within a semiconductor laser is non-radiative recombination, i.e. the process in which the energy of the recombining electron-hole pair is transferred to phonons. The intensity of this process is proportional to the difference $(1 - \eta_{sp})$, where η_{sp} is the *internal spontaneous quantum efficiency* which shows the fraction of the carriers that emit photons through spontaneous recombination, it is given by

$$\eta_{sp} = \frac{B(T)n}{A_{nr}(T) + B(T)n + C(T)n^2}. \quad (3.3)$$

The non-radiative recombination process occurs within the active region where it competes with radiative recombination. The power density g_{NR} of the heat thus generated is given by

$$g_{NR} = \frac{1}{d} V_{pn} J_{th} (1 - \eta_{sp}) \quad (3.4)$$

where the parameters are as described before. Note that this equation assumes that below the threshold, i.e., for $j < j_{th}$, most of the electric power supplied to the active region is transformed into heat because the internal quantum efficiency for the spontaneous emission η_{sp} is much less than unity. Above threshold the probability of stimulated recombination occurring is very high and less carriers contribute to heat generation. As a consequence, Eq. 3.4 remains fixed at its threshold value for a constant temperature.

Absorption of radiation. Considerable portion of the radiation generated in the active region is reabsorbed, chiefly in the resonator. The absorption occurs through *scattering* processes, where the energy of the photons is directly absorbed into phonons by means of

inelastic collisions at a rate given by $c\alpha_{int}/\eta_g$. In many junction lasers, the thickness of the resonator is approximated to the thickness of the active region. The total power density g_A of heat there generated may be expressed as [7]

$$g_A = \frac{V_{pn}}{d} [J_{th} + (J - J_{th})(1 - \eta_d)] \quad (3.5)$$

where it has been assumed that the average energy of absorbed photons is approximately equal to eV_{pn} , and η_d is the *external differential quantum efficiency* given by

$$\eta_d = \eta_i \frac{\alpha_m}{\alpha_m + \alpha_{int}}; \quad (3.6)$$

In Eq. 3.6 η_i is the *internal quantum efficiency* indicating what fraction of injected carriers is converted into photons; it is given by

$$\eta_i = \frac{B(T)n^2 + R_{st}N_{ph}}{A_{nr}(T)n + B(T)n^2 + C(T)n^3 + R_{st}N_{ph}}. \quad (3.7)$$

Radiative Transfer. In heterojunction devices, the spontaneous radiation generated isotropically within the active region may easily penetrate through the existing passive regions of low absorption coefficient. As a result of this effect, known as the *radiative transfer*, the radiation is absorbed into heat outside the active region. This part of the spontaneous radiation which leaves the active region, is almost immediately absorbed in the surrounding areas (usually the cladding regions). The random nature of this process makes very difficult to estimate the spatial distribution and the extent to which this heat source penetrates into the substrate and other layers. In general, the thickness of this heat source is very

small and it is often a good approximation to assume that these absorption events take place also inside the active region. By taking this heat source into consideration, Eq. 3.5 becomes

$$g_A = \frac{V}{d} \frac{pn}{d} [J_{th}(1 - f_T \eta_{sp}) + (J - J_{th})(1 - \eta_d)] \quad (3.8)$$

where f_T is the radiative transfer coefficient [12] given by

$$f_T = K_f \sin^2(\theta/2) \quad (3.9)$$

where K_f is a constant that depends on the laser structure and θ denotes the total internal reflection angle.

From Eqs. 3.3, 3.6 and 3.9 it is clear that in one way or the other, the various efficiency parameters represent a measure of how the different recombination mechanisms interplay. Although useful in a general sense, the heat source representation in terms of these parameters does not allow to look into the details of the laser's behavior because these parameters are usually lumped into a single number. This suggests that in investigating spatial variations of temperature, a more efficient way of representing heat power should be used. The spatial distribution of minority carriers within the active region can be used instead, and the heat generation mechanisms can be associated with *carrier concentration distributions* rather than with laser's bulk parameters. This can be done if the type of non-radiative recombination mechanisms taking place, and the rate at which they occur are known. Under this approximation, it is assumed that each act of non-radiative recombination is followed by heat generation of energy equal to *about* $h\nu$, where h is the Planck constant and ν is the laser radiation frequency.

By approximating $eV_{pn} = h\nu$ and substituting Eqs. 2.26, 2.27 and 3.3 into Eq. 3.4, the heat generated through non-radiative mechanisms in the active region can be expressed as

$$g_{NR} = (A_{nr}(T) + C(T)n^2)nh\nu \quad (3.10)$$

By carrying out similar manipulations, the heat generated due to absorption of radiation in the active region can be expressed as

$$g_A = v_g \alpha_{int} h\nu S + (1 - \beta) B(T) n^2 h\nu. \quad (3.11)$$

The first term on the right side of Eq. 3.11 represents the heat generated from the absorption of scattering loss. The rate of absorption is given by $v_g \alpha_{int}$ where v_g is the group velocity, α_{int} is the absorption constant and S is the total photon density. The second term represents the absorption of spontaneous radiation which deviates outside the active region. Note that in this term, all the radiation that did not couple into the lasing mode is assumed to turn into heat.

Joule Heating. According to Joule's law of heating, in all semiconductor layers, an electric current flow is followed by generation of volume Joule heating g_J :

$$g_J = J^2 \rho \quad (3.12)$$

where ρ stands for the electrical resistivity (in Ωcm).

All heat generation processes represent volume heat sources (W/cm^3). They are however, often treated as flat (planar) heat sources located in the middle of their generation regions. Even though this simplifies the modeling of the different heat sources, such

approximation can deteriorate the accuracy of the calculated temperature profiles, especially in the close vicinity of the heat source. It is a good approximation, however, when the heat sources are relatively thin with respect to the thermal variations.

The importance of the individual heat sources varies considerably from one laser structure to another. In some semiconductor lasers, the heat generated due to absorption of radiation is very small compared with that generated due to non-radiative processes. Their thermal behavior is then mainly determined by the non-radiative recombination rates and the heat sources represented by Eq. 3.11 can be neglected. On the other hand, only Joule heating is important for some lasers. In general, a careful analysis of the specific laser structure to be modeled is required (such as layer resistivities, electrical and optical confinements, etc.) to assess the importance of the different heat sources involved.

In our model, all the previous heat sources are accounted for; in addition, the heat source representation in terms of the carrier density distribution is chosen because this formalism is required for investigating the distribution of temperature which we are interested in (longitudinal variations). In this way, the heat power is determined by Eqs. 3.10, 3.11 and 3.12.

3.3 Heat Conduction Equation

The flow of heat in a solid body is described by the heat conduction equation

$$\nabla(\kappa(T)\nabla T) = \frac{\kappa(T)}{\chi} \left(\frac{\partial T}{\partial t} \right) - g_T \quad (3.13)$$

where T is temperature, t is time, g_T is the spatial power density of heat sources, κ is the thermal conductivity of the material, and χ is its thermal diffusivity; the latter two quantities are thus parameters characterizing the thermal properties of the material. Table 3.1 shows these parameters for some III-V binary compounds.

Table 3.1: Room-temperature thermal conductivities and thermal diffusivities of III-V binary compounds [7].

Material	κ , W/mK	χ , 10^{-6} m ² /s
GaAs	44	26
InAs	26.5	19
InP	68	46
GaP	77	43
AlAs	91	48
InSb	17	11
GaSb	33	25
AlSb	57	43
AlP	90	-
AlN	160	-
GaN	130	44

The heat conduction equation (Eq. 3.13) is non-linear, since the thermal conductivity κ , the thermal diffusivity χ , and sometimes also the power density g_T of the heat sources depend on the temperature T . For GaAs, the temperature dependence of κ may be written as

$$\kappa_{GaAs}(T) = 44 \left(\frac{300}{T} \right)^{1.2} [\text{W/mK}] \quad (3.14)$$

This dependence is especially important for relatively *high temperature* increases because of its strong non-linear behavior. The relationships $\kappa(T)$ for other binary compounds are given in Table 3.2.

Table 3.2: Variation of thermal conductivities of III-V binary compounds with temperature [7].

Material	Thermal Conductivity, W/mK	Range, K
GaP	$77 (T/300)^{-1.364}$	300 - 550
InAs	$26.5 (T/300)^{-1.234}$	300 - 650
InP	$100/[1.47 + (T-300)/111]$	300 - 750
GaAs	$44 (T/300)^{-1.25}$	300 - 900
AlSb	$57 (T/300)^{-1.330}$	300 - 800
GaSb	$33 (T/300)^{-1.774}$	300 - 900
InSb	$0.48 + (4806/T)$	270 - 500

Many different heat-sinking mechanisms are involved in the removal of heat from a semiconductor laser, however, direct conduction through the metal contact to the heat sink has proved to decisively dominate this process [7]. Therefore, in top mounted lasers, one may neglect the heat removal through the vertical walls (vw) and through the lower surface (ls) of these devices so that

$$\left. \frac{dT}{d\mathbf{n}} \right|_{vw} = \left. \frac{dT}{d\mathbf{n}} \right|_{ls} = 0 \quad (3.15)$$

where $\mathbf{n}$ is the unit vector perpendicular to the surface and directed outwards the device.

The heat sink installed in radiation sources is usually many times as great in size as the source. It is usually treated as a region of infinitely large thermal capacity, so that the boundary condition at this surface is taken in the form

$$T_{HS} = T_A, \quad (3.16)$$

where T_A is the ambient temperature, and HS denotes the outer surface of the heat sink.

3.4 Solution of the Heat Conduction Equation

The thermal properties of a CW operating semiconductor laser are usually expressed in terms of the thermal resistance R_{TH} , defined as the ratio of the average temperature T_{av} of the active region and the supply power P [14], so that

$$T_{av} = R_{TH}P \quad (3.17)$$

In order to find the stationary temperature distribution within a laser, the steady state form ($\partial T / \partial t = 0$) of Eq. 3.13 has to be solved for specified boundary conditions.

3.4.1 Analytical Solution

An analytical solution for a stripe geometry laser that describes the heat flow within the laser structure has been previously obtained [14], [15]. Fig. 3.2 shows the structure of such a laser. In this model heat is generated uniformly and steadily in a planar stripe source which is embedded in a rectangular layered laser. No heat escape through the

sides, ends, or top of the laser is considered; heat flow is two dimensional into a highly conductive sink at temperature T_{HS} . The thickness of the i th layer is t_i , and its conductivity, which is taken to be a constant, is κ_i .

Figure 3.2: Schematic structure of a stripe geometry laser.

A separation of variables solution of the steady state heat equation for the temperature T_i within the i th layer is

$$T_i(x, y) = T_{HS} + \beta_{i,0}(1 - \gamma_{i,0}y) + \quad (3.18)$$

$$\sum_{n=1}^{\infty} \beta_{i,n} [\cosh(k_n y) - \gamma_{i,n} \sinh(k_n y)] \cos(k_n x)$$

where the undetermined coefficients $\beta_{i,n}$ and $\gamma_{i,n}$ should be calculated from the continuity conditions for the temperature distribution and the heat flux distribution at the layer boundaries (see Appendix A).

The term $\sin(k_n x)$ does not appear because of the symmetric stripe position, and the separation constant k_n is given by

$$k_n = \frac{2n\pi}{W} \quad (3.19)$$

in order to satisfy the requirement of no heat escape from the lateral faces, i.e.,

$$\left[\frac{\partial T}{\partial x} \right]_{|x| = \frac{W}{2}} = 0. \quad (3.20)$$

Positive y is in the direction away from the source with the origin of y in each layer at that face of the layer nearer the source. The uniform heat generation rate per unit stripe area of the heat source is a constant denoted by g (W/cm²).

The determination of the temperature reduces then to the calculation of a sufficient number of terms in Eq. 3.18. By renumbering the layers of the laser structure (c.f. Fig. 3.2), the heat source can be placed at any interface and all the equations remain unchanged. This makes the model useful to obtain the temperature distribution at any point (x, y) within a simplified layered-structure laser and to look into the individual contri-

butions of the various heat sources. Although accurate, the model requires extensive calculations when several heat sources are accounted for. In that case, Eq. 3.18 needs to be solved for each heat source and the final solution obtained by superposition.

In reality, most of the laser structures are quite complex and to come up with analytical solutions of Eq. 3.13 for such structures is a very challenging task. Specifically, the complications are: variety of boundary conditions difficult to deal with, heat spreading occurs in 3 dimensions, complex geometries, wide variety of materials with greatly differing and often temperature dependent thermal conductivities, etc. Thus, analytical solutions of Eq. 3.13 are difficult to obtain and their use is limited to specific simplified laser structures which are consistent with the formulation of the solution method.

3.4.2 Numerical Solution

A purely numerical approach to heat conduction problems offers more flexibility in terms of defining the device structure, and is often suitable to overcome the above mentioned problems.

3.4.2.1 Atar - a 3D Thermal Simulator

A number of general purpose numerical tools exist to solve thermal problems of this kind. The tools are based on finite elements or finite differences solution of the heat equation (Eq. 3.13). Two significant drawbacks in the use of these tools are the time and expertise that are needed to build a 3D model and the difficulty in obtaining self-consistent solutions of non-linear problems. This would normally require an iterative approach and

an automatic way of reading and writing data into the numerical package. Since this feature is usually not available to the user, self-consistent problems can not be easily handled.

The solution to these problems is a simulator that given a technology description and layout information will automatically create a 3D model of the device, discretize the model, solve for the temperature distribution in the device and provide an easy input/output user interface. A simulation tool that fulfills these expectations is *Atar* [16], a 3D thermal simulator for semiconductor integrated devices. *Atar* was extensively used in the thermal simulation of semiconductor lasers described in this work. for this reason the next pages are devoted to providing an insight into the engineering principles of this simulation program.

To produce a numerical solution of Eq. 3.13 an appropriately discretized 3D model must be built, boundary conditions applied, a mathematical model extracted and then solved using a non-linear method.

3.4.2.1.1 3-D Model Creation

The basic idea behind the model building is to create a complex mesh geometry using rectangular blocks of varying sizes. The model is created such that a block and its neighbors form a relatively simple set of topologies where by each block can have a maximum of two blocks on any vertical side and a maximum of four blocks on the top or bottom of a block in what is called a *quad tree mesh*. From these block topologies a network of thermal resistances and capacitances is extracted representing the thermal equation to be solved. By using these rules, complex meshes can be built. A second key part to the

model building is to build from chip surface down into the substrate and then up into the backend.

To improve accuracy, the model to be created must have a high density of elements or nodes in the regions of large heat flow or temperature gradients. This essentially means that a large number of small elements must be created near the heat generation regions in the simulation domain. *Atar* produces an appropriately refined rectangular mesh from layout information and an input file indicating the location of the heat sources.

3.4.2.1.2 Extraction of the Thermal Network

After the creation of the 3D model a thermal network is extracted consisting of non-linear resistors and capacitors. Each block is visited and using the topology of the connections with adjacent blocks a network of resistors to the neighboring nodes is created. The value of each resistor is calculated using the common cross-sectional area, the size of the two blocks and the local thermal conductivity. The thermal capacitance of the material is used to calculate the value of a capacitor connected from the block node to ground. As the material properties are, in general, temperature dependent the resistor and the capacitor values are a function of the node temperature. Next, the appropriate boundary conditions are placed on the model. Regions of the model can be designated to be at fix temperature, as would be used for a fixed backside wafer temperature. Constant power generation can be specified in volumes such as the active region, high resistivity areas where Joule heating is present or wherever the heat sources for a specific structure are located. Once the thermal model has been extracted and the boundary conditions set, the network needs to be solved. For this purpose, a direct solver is used which first sets up a

global sparse array and then solves the set of equations representing the thermal resistor network.

As mentioned before, when self-consistent simulations are required, the proper links to exchange data among the different programs involved (electrical, optical or mechanical simulators, for example) must exist. To this end, *Atar* allows that small volumetric regions be defined as *independent devices*. Each device is characterized by an array of data containing operating parameters such as power, temperature, and physical parameters such as volume and position. Thanks to *Atar*'s flexible input/output interface, the information can be read and updated by the user quite easily and by defining small devices in the appropriate locations, non-uniform power generation regions and hence self-consistent simulations can be handled.

3.4.3 Comparison Between the Numerical and Analytical Solutions

The purpose of this section is twofold, firstly, the building capabilities of *Atar* are demonstrated by building a laser structure from layout and technology information and, secondly, its accuracy is demonstrated by comparing *Atar*'s solution with the analytical one (Eq. 3.18) for a simple laser structure.

Fig. 3.3 shows schematically the nominal conductivities and thicknesses of the layers of a typical heterostructure laser. Note that even though this structure might seem unrealistic for the present art, it is used here for illustrative purposes only. To keep things simple, only heat generated in the active region was considered and all the different non-radiative recombination mechanisms were lumped into a constant uniform power source.

Figure 3.3: Thermal parameterization of a typical GaAs/GaAlAs heterostructure laser.

The heat source is considered flat and is embedded in the vertical center of the active layer by regarding this layer mathematically as two equal layers. It is apparent from discussions in the previous sections that this heat source representation is a very rough approximation for a semiconductor laser. It is good enough however, to demonstrate the accuracy and the usefulness of *Atar* as a simulation tool. To this end, the laser structure shown in Fig. 3.3 was built and simulated using *Atar*. Since the structure is symmetrical about the y axis, half-symmetry could be applied. In addition, heat flow is two dimensional

for a structure like this and this allowed to reduce the simulation domain to a thin slice of the original structure.

After the proper technology and layout information files were created, the structure to be modeled was extracted. In the discretized domain the laser structure looked as follows (note that the structure was built substrate down):

Figure 3.4: Original and discretized domain of the laser structure using Atar.

Fig. 3.4 shows a front view of the laser structure and its automatic discretization. Since large temperature and flux gradients were expected in the active region and in layers adjacent to it, a denser mesh was created at those locations in order to detect those small spatial variations. Fig 3.5 shows a detailed view of the mesh in the active region and the

adjacent layers. Conversely, deep down in the substrate the temperature gradients were expected to be small and, therefore, a rough mesh was enough to properly detect the changes taking place there.

Figure 3.5: Detailed view of the mesh in the active region and adjacent layers.

Note that the aspect ratio of the vertical (y direction) and lateral (x direction) dimensions in typical semiconductor lasers makes the discretization process a cumbersome task. The discretization step in the vertical direction is usually required to be very small so that the extremely thin layers can be included. In order to keep a good block aspect ratio, a small discretization step in the lateral direction must also be used. In turn, this can lead to a large number of blocks in the discretized domain and consequently to slow simulations. Due to these problems, the discretization process needs to be iterated

until a good compromise between accuracy and speed is reached. This is done by varying parameters such as maximum block size and maximum number of sublayers for a specific layer of the structure of interest in the *Atar* input files.

Fig. 3.6 shows a comparison between the analytical and *Atar* solutions of the lateral temperature profiles along the active region for a top boundary condition fixed at 300 K and total heat power of 26 mW.

Figure 3.6: Comparison between the analytic and *Atar* solutions of the lateral temperature profiles along the active region.

To obtain the analytic solution, only one hundred terms were considered in Eq. 3.18. The plot shows that the result obtained by either method is practically the same. The small discrepancy in the close vicinity of the active region can be explained in the follow-

ing manner. The analytic model requires that the heat source be infinitesimally thin for it to be applicable. In the Atar model however, the heat source had a finite thickness of $0.25\text{ }\mu\text{m}$ and this turned the planar source into a volumetric one. As can be seen from the figure, this had the effect of slightly decreasing the temperature right at the middle of the active region, where the heat source was located. Similarly, the transverse (y direction) temperature profiles are compared in Fig. 3.7.

Figure 3.7: Comparison between the analytic and Atar solutions of the transverse temperature profiles.

Once more, the result is practically the same for both analytic and Atar solutions. Only a small difference is present in the substrate because of the large blocks that were used to discretize that region.

After a thermal problem is solved, the solution can be displayed in the form of isothermals (temperature contours) in the structure of interest and the temperature distribution can be stored in data files and plotted afterwards. To illustrate Atar's post-processing capability, Fig. 3.8 shows a detailed temperature contour plot in the vicinity of the active region.

Figure 3.8: Temperature contours in the vicinity of the active region.

The temperature contours are especially important when the parameters of the physical structure are analyzed for optimization. The usual goal is to improve the heat flow across the different layers and minimize the temperature in the hottest areas (usually

the active region). The temperature contours become important in this analysis because they provide useful information about the direction in which the heat flows (the heat flows perpendicular to the isothermals) inside the actual structure. This assists in understanding the importance that the different layers play in the thermal behavior of a specific laser structure and suggests how the structure should be modified to enhance heat spreading and minimize the temperature in regions where necessary.

From this discussion we can conclude that *Atar* is a powerful tool capable of solving thermal problems within good accuracy, it can be used to study the thermal behavior of complex laser structures in isolation or it can be linked to assist in self-consistent simulations. In addition, its post-processing capabilities allow for quick and easy analysis of results. All these characteristics make *Atar* a better tool for thermal laser modeling as compared to analytic models. To further illustrate this, the chapter is closed with the following table that summarizes a comparison between *Atar* and the analytic model.

Table 3.3: Summarized comparison between *Atar* and the Analytical model.

Feature	Atar	Analytical Models
Flexibility	Can handle complex/non-linear structures	Limited to homogenous structures with regular geometry
Accuracy	Good (limited by convergence)	Good (limited by n in Eq. 3.18)
Computation time	Short (< 15 min for Fig. 3.4) but affected by the mesh size.	Large (> 5 min/point for 100 terms in Eq. 3.18). Every point requires a complete summation.
Memory requirements	Medium/Large (1M - 1G)	Small (< 1M)
Discretization step	Limited by the size of the domain/memory	As small as required ($\sim 0.01 \text{ m}\mu$)
Post-processing information	Temperature and Isothermals can be shown in the structure	Temperature is known in a few points. Isothermals require extensive calculations

Chapter 4

Self-Consistent Laser Model

4.1 Introduction

The thermal behavior of a semiconductor laser structure can be investigated from a strictly thermal point of view provided that the location of the heat sources and the power dissipated by each of them are known a priori. This approach was taken in Chapter 3 where a pure 3D thermal model was proposed to thermally optimize these devices. Two important constraints of this approach were that the power had to be uniformly distributed among the different heat sources and that the amount of power attributed to each of them was not easily determined (the total dissipated power has to be distributed over the heat sources taken into consideration). As it will be demonstrated later, this approach is good enough for some laser structures for it predicts the right tendency in temperature changes (average) in the device due to material parameters or structural variations. In real semiconductor lasers, the distribution of power in the heat sources is not uniform due to the continuous interaction among the different non-uniform physical phenomena. This interaction arises from the fact that many material parameters (such as electrical resistivities, refractive indices, absorption coefficients) and device parameters (such as threshold current and quantum efficiencies), which influence the heat generation, are strongly dependent on temperature. This can be included in the model using a self-consistent approach, when in successive iteration loops of the calculation, values of the above parameters determined in the

previous loop are used. Self-consistency is assumed to be reached when differences between results of calculations obtained in two consecutive loops are below given limits.

Strictly speaking, not only material and device parameters but also distributions of current densities and carrier concentrations within the whole laser structure are dependent on current temperature profiles. This is so because the current spreading and the carrier diffusion phenomena are also temperature dependent processes. In order to get more accurate temperature profiles, a thermal-electrical-optical self-consistent approach needs to be used. In this way, the optical processes with their mutual interactions with both the thermal and electrical processes are taken into consideration. The full picture of mutual interactions among all these processes is shown in Fig. 4.1. In lasers with strain active regions, additionally, mechanical processes should be included.

Fig. 4.1 also shows the different interactions that were included in the present model. As it can be seen there, most of them were taken into consideration with the only exceptions being those processes shown inside the shaded rectangles. From this figure one can also anticipate that in order to implement a self-consistent model at least three equations have to be solved, one for each of the physical phenomena that is to be modeled. In Chapter 3 a 3D thermal solver was described; while the same program was adapted to solve the thermal equation in a self-consistent manner without difficulty, this still left the burden of implementing the models to account for the electrical and optical phenomena. This chapter is devoted to the description of the procedure and analysis carried out to develop the optical and electrical models. Many of the equations shown in this chapter were previously described in earlier chapters since they constitute (in one way or the

other) the basic mathematical tools for laser modeling. The equations are shown here where necessary to emphasize how they were modified in order to improve the basic lumped model described in Chapter 2 to fit the needs of the present model.

Figure 4.1: Mutual interactions among thermal, electrical and optical processes in semiconductor lasers (the shaded areas show the phenomena which are not included in the model).

4.2 Optical Model

A laser beam is coherent electromagnetic radiation. The propagation must therefore obey Maxwell's equations. If we restrict our attention to beams that have small angular divergence and to media whose refractive index does not vary appreciably in the transverse direction, the vector wave equations can be reduced to a scalar wave equation [17].

4.2.1 Scalar Modes

Scalar modes are the eigensolutions of the equation

$$\nabla_T^2 E_z + k^2 E_z = \beta_s^2 E_z \quad (4.1)$$

where

$$\nabla_T^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (4.2)$$

is the transverse Laplacian and

$$k(x, y) = (2\pi\eta(x, y))/\lambda \quad (4.3)$$

where λ is the free-space wavelength. Under the scalar approximation, E_z , $\partial E_z/\partial y$, and $\partial E_z/\partial x$ are continuous throughout the solution domain.

The usual five-point difference approximation for Eq. 4.1 would allow us to replace it for the following discrete form [18]

$$\frac{E_{r,s-1}}{h_x^2} + \frac{E_{r-1,s}}{h_y^2} + (k_{r,s}^2 - 2/h_x^2 - 2/h_y^2)E_{r,s} + \frac{E_{r+1,s}}{h_x^2} + \frac{E_{r,1+s}}{h_y^2} = \beta_s^2 E_{r,s} \quad (4.4)$$

at the grid point (x_r, y_s) . By applying this approximation at each internal grid point, with appropriate amendments when (x_r, y_s) is adjacent to the outer boundary of the solution domain, the following algebraic eigenvalue problem is obtained

$$A_s \overline{E}_s = \beta_s^2 \overline{E}_s \quad (4.5)$$

in Eq. 4.5 A_s is a real *symmetric band* matrix, β_s^2 is the scalar mode propagation eigenvalue, and $\overline{E}_s$ is the corresponding normalized eigenvector representing the field profile $E_z(x, y)$. Eq. 4.5 can be solved by various techniques [19].

4.2.2 Optical Waveguide Analysis

The structure whose physical properties wanted to be investigated corresponded to that of a ridge laser that will be described in more detail latter. In order to determine the appropriate optical model to use, the analysis of the waveguiding properties of the device by applying Eq. 4.4, was the starting point. The boundary condition for the outer boundary of the solution domain of the wave equation was set to

$$E(x, y)|_{\partial\Omega} = 0 \quad (4.6)$$

Since the modeled device is symmetric, only half of it needed to be simulated. The boundary condition along the plane of symmetry ($\partial\Omega_s$), was

$$\frac{\partial E(x, y)}{\partial \hat{n}} = 0 \quad (4.7)$$

for the even modes and

$$E(x, y)|_{\partial\Omega_i} = 0 \quad (4.8)$$

for the odd modes.

The simulated device structure looked as follows:

Figure 4.2: Simulated optical waveguide.

For simulating the structure, the refractive index of InP was set to 3.17 and that of InGaAsP to 3.38. Since this structure corresponds to a weakly index guided semiconductor laser, part of the guiding mechanism is due to the carrier density distribution (gain guiding). The index step between the active region and the claddings governs the extent of index guiding. An index step in the range of 0.005-0.01 is enough to achieve index-guided regime operation [5], for this reason only this mechanism was considered in these simulations.

The Inverse Power Iteration Method [18], [20] was used to solve Eq. 4.5. This method while easy to implement is computationally expensive and therefore a good compromise between accuracy and speed must be sought. Fig. 4.3 shows the electric field profile of the fundamental mode supported by the waveguide. This result was obtained by using half symmetry and gradually reducing the computational domain in regions where the optical field was previously zero.

Figure 4.3: Normalized intensity of the fundamental optical mode.

As expected, the optical mode is highly confined in the active region. This can be seen better in Fig. 4.4, where a contour plot of this optical mode is shown.

Figure 4.4: Contour plot of the fundamental mode.

From Fig. 4.4 one can see that since the lateral optical intensity outside the ridge is very small, the volume of the active region can be approximated by the dimensions of the quantum-well under the metal contact where the radiative recombination processes are more likely to occur. In addition, the lateral variations occur in a very short distance compared with the dimensions of the whole thermal domain (the laser is $250\text{ }\mu\text{m}$ wide) and are of minor importance from a thermal standpoint. The portion of the optical field outside the active region is accounted for by the confinement factor Γ .

4.2.3 1-D Optical Model

From the analysis performed in the previous section, it can be seen that the inclusion of the lateral variations of the optical field is not justified because they are computationally expensive to determine and occur over a very short distance.

The longitudinal optical intensity variations however, occur in the same dimensional scale as that of the thermal problem and must therefore be incorporated. To this end, the optical phenomenon was modeled by the one dimensional equation

$$\pm \frac{c}{\eta_g} \frac{\partial I^{f,b}}{\partial z} = \frac{c}{\eta_g} [\Gamma g(n, T) - \alpha_{int}] I^{f,b} + \beta B(T) n^2 \quad (4.9)$$

where z is the axial length, $I^f(z)$ is the light intensity of the wave travelling in the positive z direction and $I^b(z)$ is the light intensity of the wave travelling in the negative z direction. This equation is very similar to the steady state rate equation for the photon density (Eq. 2.28). The differences are the inclusion of a spatially dependent term for the light intensity, the temperature dependence of the gain $g(n, T)$ and stimulated emission coefficient $B(T)$, and the fact that the facet losses are no longer distributed over the whole cavity as it was the case in the lumped model.

The temperature dependence of the gain model arises from the fact that the transparency carrier density increases with temperature. This is so because the effective number of states, N_c and N_v , increases with temperature while the band gap, E_g , decreases. In a parabolic band approximation, the transparency carrier density can be obtained from

Fermi-Dirac statistics (see Appendix B). To take this into consideration, Eq. 2.9 was replaced by:

$$g(n, T) = a \left(n - n_o \times \exp \left(-\frac{E_{int}}{k_B T} \right) \right) \quad (4.10)$$

where k_B is Boltzmann's constant and E_{int} is the activation energy. The linear dependence of the optical gain on the carrier density and the Arrhenius temperature dependence of absorption are adopted on the basis of the experimental evidence [21], [22].

In the same way, the spontaneous emission coefficient was replaced by

$$B(T) = B \times \exp \left(-\frac{E_2}{k_B T} \right) \quad (4.11)$$

where E_2 represents the activation energy of this mechanism. The other terms in Eq. 4.11 are as previously defined in Table 2.1.

In general, the gain $g(n, T)$, the carrier density n and the photon density $I^{f,b}$ are functions of their lateral positions in the optical waveguide. In Eq. 4.9 these three quantities simply represent the corresponding average values and are only longitudinally distributed.

The boundary conditions for the intensities are given in terms of the reflectivities R_1 at $z = 0$ and R_2 at $z = L$ by the expressions

$$I^f(0) = R_1 I^b(0), \quad I^b(L) = R_2 I^f(L) \quad (4.12)$$

To solve Eq. 4.9 a Runge-Kutta second order integration method was used [23].

4.3 Electrical Model

After the current is injected into the laser through the top contact, diffusion of carriers starts to occur. The excellent carrier confinement provided by the ridge however [24], keeps the lateral diffusion of carriers outside the area underneath the contact very small. The relatively short spatial variations are again small when compared with the size of the whole thermal problem and can be neglected for the same reasons explained above.

In order to be consistent with the formulation of the optical model, only the longitudinal spatial variations of carriers needed to be included in the electrical model. The variation in carrier density along the longitudinal direction is of the order of $1 \times 10^3 \text{ cm}^{-4}$ (approximately one order of magnitude variation due to spatial hole burning over a cavity length of the order of 1 mm) [20]. As a result, diffusion of carriers in the longitudinal axis is very unlikely to occur and under normal operating conditions can be neglected. Under certain conditions however, diffusion becomes important and should not be neglected without knowing the consequences that this implies. The importance of carrier diffusion under both, normal and high-power operating conditions will be analyzed in more detail in Chapter 5.

4.3.1 Diffusion of Carriers

The present model is specially suited (due to its longitudinal nature) to study the so called thermal runaway or catastrophic optical damage that occurs due to the increased density of defects in the vicinity of the cleaved facets. In this case, diffusion of carriers needs to be included in the electrical rate equation since the high temperature gradients in

those regions are not only going to give rise to longitudinal diffusion of carriers but in some cases, diffusion itself will be regarded as the mechanism responsible for this problem. Thus, in order to be in a position to analyze these physical effects, diffusion of carriers must be included in the electrical model. When diffusion is included, the electrical phenomenon is adequately modeled by the carrier rate equation

$$0 = D \frac{d^2 n}{dz^2} + \frac{J}{ed} - \frac{c}{\eta_g} [\Gamma g(n, T)] [I^f + I^b] - \frac{n}{\tau_c(n, T)} \quad (4.13)$$

where D is the diffusion coefficient, τ_c is the carrier recombination time described below and the other parameters are as previously defined in Table 2.1. Eq. 4.13 requires that the current flow into the active region, minus the current lost to stimulated and spontaneous emissions and non-radiative recombination processes be equal. The carrier recombination time τ_c is given by the equation

$$\tau_c(n, T) = \left(A_{nr} \times \exp\left(-\frac{E_1}{k_B T}\right) + Bn \times \exp\left(-\frac{E_2}{k_B T}\right) + Cn^2 \times \exp\left(-\frac{E_3}{k_B T}\right) \right)^{-1}. \quad (4.14)$$

One can readily notice that Eq. 4.13 is very similar to Eq. 2.20. It has been modified however, to account for the temperature dependence of the gain and the recombination terms that comprise it. The first term on the right hand side of Eq. 4.14 represents surface and defect recombination with activation energy E_1 , the second radiative recombination with activation energy E_2 , and the third Auger recombination with activation energy E_3 . The radiative recombination can be split into spontaneous emission which enters the lasing mode and spontaneous emission which is absorbed in the surround as heat. The constant β appearing in Eq. 4.9 represents the fraction of spontaneous radiation which enters the las-

ing mode. The electrons lost to the valence band by non-radiative and Auger recombination are assumed to have converted their energy into heat in the active layer.

The diffusion equation must in general be solved subject to the boundary conditions of zero flow of carriers at the edges of the simulated domain (laser facets); that is,

$$-D\frac{d}{dz}n(0) = 0, \quad -D\frac{d}{dz}n(L) = 0, \quad (4.15)$$

if, however, extra heating at the facets is to be investigated, the facets are thought of as sinks for the carriers, a surface recombination velocity is present, and the following boundary conditions must be applied

$$-D\frac{d}{dz}n(0) = -V_{sur}n(0), \quad -D\frac{d}{dz}n(L) = V_{sur}n(L) \quad (4.16)$$

in the special case when $V_{sur} = 0$, Eq. 4.16 reduces to Eq. 4.15. The dissipated power due to surface recombination can then be obtained as the product of the surface recombination velocity V_{sur} , the carrier density at the facets $n(z=0, L)$, and the photon energy $h\nu$ so that the heat power at these locations is calculated from

$$P_{0,L} = wdV_{sur}h\nu n(0, L) \quad (4.17)$$

where w and d are the width and the thickness of the active region respectively.

To solve Eq. 4.13 a standard five point difference scheme was used to approximate the second order derivative in a Newton iteration loop [25] (see Appendix C).

4.3.2 Non-diffusion Equation

As it was mentioned above, diffusion of carriers becomes an issue when high gradients are formed due to highly irregular longitudinal temperature profiles. These in turn are caused by localized extra heating at the facets. Under normal operating conditions (nominal current) and low power operation however, the longitudinal temperature in the active region is relatively small (approx. 20 - 30 K above ambient temperature) and consequently the temperature gradients are almost negligible. Under these conditions, diffusion of carriers can be neglected and Eq. 4.13 becomes

$$0 = \frac{J}{ed} - \frac{c}{\eta_g} [\Gamma g(n, T)] [I^f + I^b] \quad (4.18)$$

$$- n \left[A_{nr} \times \exp\left(-\frac{E_1}{k_B T}\right) + Bn \times \exp\left(-\frac{E_2}{k_B T}\right) + Cn^2 \times \exp\left(-\frac{E_3}{k_B T}\right) \right]$$

By substituting Eq. 4.10 into Eq. 4.18 and grouping terms, Eq. 4.18 becomes a cubic polynomial in n whose analytical solution can readily be obtained (see Appendix D).

It should be noticed that Eq. 4.18 can not be solved subject to any explicit boundary condition, it corresponds however to an implicit boundary condition of zero flow at the facets.

4.4 The Overall Model

After the details of the equations accounting for each of the thermal, optical and electrical phenomena have thoroughly been described, we are finally in a position to summarize the overall model that is being presented here. The model consists of the solution of the 3-D thermal heat equation (Eq. 3.13) subject to the boundary conditions given by

Eqs. 3.15 and 3.16; the solution of the 1-D optical rate equation (Eq. 4.9) subject to Eq. 4.12, and the solution of the 1-D carrier rate equation (Eq. 4.13) subject to Eq. 4.16, all of them solved in a self-consistent way. For convenience, all the equations comprising the self-consistent model are reproduced below.

Thermal Equations:

$$\nabla(x, y, z)[\kappa(T)\nabla(x, y, z)T] = -g_T \quad (\text{EQ 3.13})$$

with

$$g_T = (A_{nr}(T) + C(T)n^2)nh\nu + \nu_g \alpha_{int} h\nu S + (1 - \beta)B(T)n^2 h\nu + J^2 \rho \quad (\text{EQ 3.10-3.12})$$

subject to the following boundary conditions

$$\left. \frac{dT}{dn} \right|_{vw} = \left. \frac{dT}{dn} \right|_{ls} = 0, \quad T_{HS} = T_A \quad (\text{EQ 3.15, 3.16})$$

Optical Equations:

$$\pm \frac{c}{\eta_g} \frac{\partial I^{f,b}}{\partial z} = \frac{c}{\eta_g} [\Gamma g(n, T) - \alpha_{int}] I^{f,b} + \beta B(T) n^2 \quad (\text{EQ 4.9})$$

with boundary conditions

$$I^f(0) = R_1 I^b(0), \quad I^b(L) = R_2 I^f(L) \quad (\text{EQ 4.12})$$

Electrical Equations:

$$0 = D \frac{d^2 n}{dz^2} + \frac{J}{ed} - \frac{c}{\eta_g} [\Gamma g(n, T)] [I^f + I^b] - \frac{n}{\tau_c(n, T)} \quad (\text{EQ 4.13})$$

subject to:

$$-D \frac{dn}{dz}(0) = -V_{sur} n(0), \quad -D \frac{dn}{dz}(L) = V_{sur} n(L) \quad (\text{EQ 4.16})$$

Threshold Condition:

To complete the description of the model, the previous equations must also be solved subject to the laser threshold condition given by Eq. 2.17. When this equation is modified to account for temperature and spatial dependencies, it becomes:

$$\frac{1}{2} \ln\left(\frac{1}{R_1 R_2}\right) = \int_{z=0}^L \left[\Gamma a\left(n_{th}(T) - n_o \times \exp\left(-\frac{E_{int}}{k_B T}\right)\right) - \alpha_{int} \right] dz \quad (4.19)$$

where the threshold carrier density n_{th} is found from

$$\begin{aligned} \frac{J_{th}(T)}{ed} = n_{th}(T) & \left[A_{nr} \times \exp\left(-\frac{E_1}{k_B T}\right) + B n_{th}(T) \times \exp\left(-\frac{E_2}{k_B T}\right) + \right. \\ & \left. C n_{th}(T)^2 \times \exp\left(-\frac{E_3}{k_B T}\right) \right] \end{aligned} \quad (4.20)$$

The coupling between the different physical phenomena occurring in semiconductor lasers can now be seen from the previous set of equations. The heat equation is coupled to the electrical equation through the carrier density n and to the optical equation through the photon density S . It is also implicitly temperature dependent (non-linear) through n and S and explicitly through the $A_{nr}(T)$, $B(T)$ and $C(T)$ coefficients. The optical equation is coupled to the heat equation through the gain $g(n, T)$ and the spontaneous emission coefficient $B(T)$, and to the electrical equation through n . Finally, the electrical equation is coupled to the heat equation through the gain $g(n, T)$ and the carrier recombination time τ_e and to the optical equation through the photon density S (note that $S = I^f + I^b$). The temperature is also bias dependent through Joule heating and the lasing condition (Eq. 4.20). This equation also determines the temperature dependence of the threshold carrier density.

Fig. 4.5 shows a graphical representation of the self-consistent model. As can be seen there, all three equations are coupled with one another.

Figure 4.5: Graphical representation of the self-consistent laser model.

The heat and electrical equations are discretized and solved by using finite differences techniques, while the optical equation is straightforward solved by a Runge-Kutta (R-K) second order integration method. The details and the necessary iteration loops to

obtain a complete self-consistent solution of the three equations is explained in more detail in the next section.

In order to illustrate how the longitudinal model compares to the lumped model presented in Chapter 2, we used this model to solve for the same typical parameter values shown in Table 2.2.

The carrier density profile that was obtained from the simulation is shown in Fig. 4.6.

Figure 4.6: Longitudinal carrier density profile.

It is important to note that in order to be consistent with the physical principles of laser operation (carrier density and gain saturation) explained in Chapter 2 and predicted

by the lumped model described there: the average value of the longitudinal carrier density distribution above threshold must be equal to the threshold value predicted by the lumped model. This condition also holds for the gain and any other longitudinally distributed parameter.

In the previous simulations, only the isothermal case was considered for two reasons: *i)* a full self-consistent calculation would require knowledge of the exact structure and material parameters of the laser such as dimensions, temperature dependence of all the recombination coefficients, gain and transparency carrier density, thermal conductivities, etc. which were not known, and *ii)* even if those parameters could be estimated from an arbitrary structure, a comparison between the isothermal and the non-isothermal solutions would require knowledge of the L-I curve for each case. To obtain the curve for the non-isothermal case would be computationally expensive since for each point in the L-I curve, a complete self-consistent calculation would be required. Nevertheless, Fig. 4.7 illustrates how the thermal effects come into play and deviate the laser behavior from a linear L-I relation (predicted by the lumped model). The L-I curve corresponds to a real laser structure that will be described in more detail in Chapter 5.

As expected, the thermal effects manifest themselves by deviating the L-I curve from a linear relation. As the injected current is increased, so is the temperature in the laser; the non-radiative recombination processes increase and this causes the threshold current to increase as well. The overall effect is a reduction in the useful optical power at the output.

Figure 4.7: Real light-current curve with thermal effects included.

4.4.1 Solution Algorithm

Fig. 4.8 shows the algorithm that was used to solve the coupled equations described above. The program is initiated by assuming that the whole laser structure is at ambient temperature (usually 300 K) and the temperature dependence of the different parameters involved is calculated. Next, the diffusion and optical equations are solved in an inner loop by propagating back and forth a trial beam and calculating the carrier density profile until convergence is achieved. This loop proceeds as follows: the average longitudinal temperature is used to find the carrier and current densities at threshold that would be predicted by a lumped model (solution of equations 2.17 and 4.20). The carrier density profile for the first iteration is set equal to the threshold value (constant) and used to inte-

grate the optical equation. From here the light intensity distribution of the forward and backward propagating beams is obtained (solution of Eq. 4.9).

Figure 4.8: Solution Algorithm.

As an initial condition for the optical equation in the first iteration, we assume that there are no photons in the resonator. In the next iterations the previous values of the light intensity at the facets are used.

The threshold carrier density is then used as an initial guess in the diffusion electrical equation and the longitudinal spatial distribution of n is found (solution of Eq. 4.13). This procedure is repeated until equations 4.19 and 4.20 are satisfied. From this procedure the carrier density and light intensity distributions which satisfy the threshold condition are obtained. The heat sources (*power distribution*) are calculated next and fed into ATAR which will provide a new temperature profile $T(x_o, y_o, z)$ to start a new iteration. The convergence criteria that was used to stop the simulation was that the temperature at any node in the whole thermal domain did not change more than 1×10^{-4} K in two consecutive iterations.

It should be noticed that Atar itself uses an iterative method to find the temperature distribution in the laser structure under simulation. Thus, it must converge within a given criteria before a new temperature profile is fed into LOES. During the first iterations of the outer loop, the heat power and temperature change considerably in two consecutive iterations. As the power settles down, so does the temperature and these changes become smaller. For this reason, the initial temperature profiles fed into LOES do not require a high degree of convergence since temperature changes of 1 K hardly change the value of the heat power. Therefore, Atar's convergence criteria was gradually reduced from an ini-

tial value of 1 K used during the first iterations, to a final value of 1×10^{-4} K used for the final iterations (close to total convergence).

4.4.2 Convergence Criteria and Stability

The previous algorithm was tested under different operating conditions and some of them resulted in highly irregular temperature and light intensity profiles (see Fig. 5.18). The first case arose as a consequence of considerably increasing the surface recombination velocity to instigate thermal runaway, the second one was a consequence of having a laser structure with very different facet reflectivities. The procedure described above proved to be highly stable since convergence was achieved in all cases.

As in any other numerical technique, the convergence speed and accuracy of the solution depend significantly on the discretization step. In our case, the laser structure had to be discretized in 3 dimensions in order to solve the heat equation. For asymmetric lasers, it was not possible to use a quarter symmetry and at least half of the device had to be simulated. Thus, if the discretization step had been chosen too small, an excessive number of nodes (unknowns) would have resulted and the computational burden would have become incredibly heavy. On the other hand, the 1-D optical and electrical equations had to be discretized in one dimension only. This resulted in a small number of nodes (compared with the thermal domain) and the discretization step of these equations was only limited by the size of the longitudinal discretization step of the active region in the thermal domain. For this reason, it is obvious that most of the computational effort lay on the solution of the thermal equation. It was seen that a good compromise between accuracy and

speed was obtained when a step size of $0.5\ \mu\text{m}$ was used to longitudinally discretize the active region.

The simultaneous solution of the optical and electrical equations for a given temperature profile (inner loop in Fig. 4.8) was usually reached within 500 to 1500 iterations depending on the injected current density. In the simulations that will be presented in the next chapter, the convergence condition (Eq. 4.19) was reached within 0.5%. The convergence speed varied from approximately 1 minute when diffusion of carriers was not included to approximately 2 minutes when it was. It should be noticed that due to the fast convergence speed of the inner loop, the solution time of the electrical and optical equations was almost insensitive to bias conditions and other parameter values such as surface recombination velocity, radiative and non-radiative recombination processes, dimensions of the laser structure, etc.

The time that was needed to complete a self-consistent opto-electrothermal simulation (outer loop in Fig. 4.8) however, varied in a wide range depending on several factors such as the dimensions of the laser structure and its symmetry since this determined the degree of mirroring that could be applied and the total number of nodes. The convergence speed was also influenced by the surface recombination velocity that was used (uniformity of power distribution), the injected current density (as this determined the total amount of heat power), etc. For this reason it is difficult to characterize the convergence speed of the whole model with a single number. To give an idea of how time came into play however, a discussion for a $400\ \mu\text{m}$ long longitudinally asymmetric laser operating at nominal current is provided below.

When thermal runaway effects were not the objective of the simulation ($V_{sur} = 0$), the complete solution was reached in 50 to 60 iterations in approximately 3 hours. When thermal runaway effects were included, high temperature gradients arose and convergence occurred much slower. For surface recombination velocities in the order of 1×10^6 , convergence was reached in 80 to 100 iterations in approximately 5 hours. When the current density was increased far above threshold, the simulation time increased up to 9 hours. A way to get around these long simulation times was to study the thermal runaway effects in a smaller symmetric laser. This allowed us to reduce the simulation domain by applying a quarter symmetry and did not affect the results as the excessive heating occurred within a short distance from the facets. The simulations were run on a Sun Ultra-10 workstation.

Chapter 5

Applications of the Laser Model

5.1 Introduction

After the purely thermal and self-consistent modeling capabilities have thoroughly been established, we are finally in a position to use these tools to investigate the thermal properties of real optical structures. The purpose of this chapter is several fold, first, a real laser structure is analyzed using the pure thermal model described in Chapter 3 with the objectives of initially understanding its thermal behavior and then the optimization of it by proposing physically realizable structural and material parameter modifications. Second, the self-consistent model presented in Chapter 4 is used to analyze the same laser structure for the cases of equal and different facet reflectivities (optically symmetric and asymmetric structures). The purpose of the analysis is to compare the thermal solutions predicted by both models and determine under which conditions the interaction among the electrical, optical and thermal phenomena do not significantly change the pure thermal solution and consequently can be ignored. The laser device that is investigated here corresponds to a ridge laser structure that is currently produced by Nortel Networks. Third, the self-consistent model is used to investigate the high power operation conditions of InGaAsP/InP and AlGaAs/GaAs laser structures. A comparison between these two devices is made and a possible explanation as for what the thermal runaway problem is more severe in GaAs

than in InP lasers is proposed. Finally, the model is used for investigating a basic laser structure in an integrated optical environment to emphasize its 3D thermal capabilities.

5.2 Analysis of a Stand-alone Symmetric Laser Structure

As it was mentioned before, the laser device that is analyzed here corresponds to a ridge laser structure which is currently fabricated by Nortel and whose schematical representation is shown in Fig. 5.1.

Figure 5.1: Schematical representation of the simulated laser structure.

The laser was designed to emit at $1.5\text{ }\mu\text{m}$ and given the materials and the arrangement of the layers, the laser is said to be of the InGaAsP/InP type. The laser is structurally symmetric about the lateral (x) and longitudinal (z) axes, it has different facet reflectivities, however, and this makes it *optically asymmetric* in the z -direction. The reflectivities are 0.75 and 0.03 for the right and left facets respectively. In the context of this thesis, the term

symmetric laser refers to those devices which are both structurally (about x and/or z) and optically symmetric. At first glance and ignoring longitudinal optical effects, a 2D thermal model seems to be sufficient to investigate the thermal properties of this laser. This indeed would be the case if the top metal was not recessed from the chip's edges as it actually occurs in the device.

Figure 5.2: Top view of the simulated structure showing the top metal recess.

Fig. 5.2 shows a top view of the simulated laser structure. As can be seen here, the thick metal layer does not extend completely over the surface of the thin metal layer and a 25 μm recess is created. The recess is a result of the packaging technique currently used for the device and its elimination would require extensive changes in the packaging process. This feature imposes the requirement that a 3D thermal model be used to investigate

the thermal behavior of this laser since the recess can not be dropped from the simulation domain without first assessing its importance.

Fig. 5.3 shows a schematical representation of the laser-mounting structure. The chip is bonded with the ridge up on a metallized CVD diamond, which in turn is bonded on an AlN carrier. Two 25 μm diameter gold wires are bonded from the chip top contact (approximately 75 μm from the lateral edge of the chip) to the AlN carrier.

Figure 5.3: Laser-mounting structure.

5.2.1 Pure Thermal Approach

As it was mentioned earlier in this work, a semiconductor laser can be viewed as a pure thermal structure which can be optimized if the total dissipated power and the location of the heat sources dissipating that power are known a priori. This approach is justified by the assumption that, for a given laser structure, the main thermal effects take place in well known regions. This approach is initially taken in this section and the different

activities that were realized to improve the thermal properties of the laser structure shown in Fig. 5.1 are described below. Unless otherwise stated, the thermal conductivities used for the simulations are those shown in Table 5.1. The fact that the laser is optically asymmetric can not be taken into consideration within the framework of a pure thermal model and consequently has to be neglected. Conversely, the structural symmetry of the laser plays a very important role here since this allows that a reflective boundary condition [26] be applied with the consequent reduction of the simulation domain. For the laser structure of Fig. 5.1 symmetry can be applied about the x and z axes and only a quarter of the original structure needs to be simulated.

Table 5.1: Thermal conductivities used for the laser structure of Fig. 5.1.

Material	κ , W/mK
InP	68
InGaAs	6.8
InGaAsP	6.8
Gold	315
SiO ₂	0.14

A) Heat Source Power Definition

The first step in the pure thermal analysis of the structure was to determine the location of the heat sources to be considered. The structural information of this laser suggested that two heat sources be considered, namely the ridge and the active region. This is

so because the ridge's resistivity is very high compared with that of the other layers thus creating a well localized heat source in this region due to Joule heating. Likewise, most of the recombination processes take place in the active region. some of the carriers recombine radiatively and some others recombine non-radiatively creating phonons and giving rise to heat generation. This argument is also supported by the fact that the current is very well confined inside these two regions. In [27] it was shown that thermally speaking, a small variation in the location of the heat sources does not affect significantly the resulting temperature distribution when the heat regions are small compared with the overall structure since the heat spreads in a point-source like manner. Following this argument, the heat source associated with the ridge was assumed to be uniformly distributed in the whole ridge's volume. Similarly, the heat source associated with the active region was approximated by the dimensions of the active region underneath the ridge.

The next step in the thermal analysis was to determine the amount of heat power that was going to be attributed to each of the heat sources at a given bias condition. Since the goal was to improve the thermal behavior of this device under normal operating conditions, the nominal current and voltage were chosen as a reference point. The laser was known to produce an optical output power of 40 mW at 170 mA and 1.7 V. From this information and by invoking the law of energy conservation, the total heat power was calculated. Once the total dissipated power was known, the last step was to assign the right percentage to each of the heat sources that were accounted for.

The total laser's resistance is approximately equal to 4-5 Ω . If the resistivity of the *n*-substrate is neglected and if it is assumed that most of the resistance is due to the *p*-type

ridge (owing to its high resistivity), approximately half of the total heat power would be dissipated in this area at the previously described bias point. For this reason, a ratio of 1-1 between the two heat sources was chosen in the pure thermal simulations.

To start with the thermal analysis, the laser structure of Fig. 5.1 was built and discretized by means of *Atar*. Since the thermal importance of the top wire is investigated in a separate section, it was not initially included in the structure. Fig. 5.4 shows the discretized domain of the laser after it was built by using *Atar*.

Figure 5.4: Discretized *Atar* model of the laser structure.

As it was previously mentioned, thanks to the structural symmetry of the laser, the computational domain was reduced to a quarter of the original structure. The ridge and the active region are located in the upper-left side. The dimensions of these regions are very small

compared with those of the whole structure and can not be visualized from the figure. To show how these regions were built in the model, Fig. 5.5 shows a detailed view of the ridge and active region.

Figure 5.5: Detailed view of the ridge and active region.

Fig. 5.6a shows the longitudinal temperature distribution in the active region for different percentages of power attributed to each heat source. It should be noticed here that regardless the location of the heat power generation, the longitudinal temperature profiles shown in this section, are always plotted *along the active region* since this is the temperature that matters to determine the laser's behavior. As can be seen from this figure, the maximum temperature is obtained when all the power is assumed to be in the active region (solid line) and the minimum when it is assumed to be in the ridge (dashed line).

Figure 5.6: Temperature profiles for different percentages of power in each heat source
(a) Longitudinal and (b) transverse.

This result is somewhat obvious since the maximum temperature is expected to be in the region where the heat source is located. The fact that the temperature difference from one case to the other is so small (approximately 2 K) indicates that for this particular laser structure the error that could be introduced by incorrectly assigning the adequate percentage of power associated with each heat source is small. This is so because the ridge and the active region are very close to each other and the layers in between are thermally negligible (the layers in between are very thin, c.f. Fig. 5.1). When all the power is placed in the ridge, the maximum temperature in the laser, which occurs in this area, is higher than that obtained when all the power is placed in the active region because the heat source is now farther from the heat sink. This can be better appreciated from Fig. 5.6b, where the transverse temperature profiles for the previous distributions of power in each region are plotted. The inset in this figure shows the plotting line (transverse direction) along which the temperature was plotted, for more structural details see Figs. 5.1 and 5.5. The fact that the temperature does not increase significantly as compared with the case when all the heat power is in the active region suggests that, even though the thermal conductivity of the active region is one tenth of that of the substrate, this layer is thin enough to significantly hamper the heat flow towards the heat sink. Another important result that can be seen from Fig. 5.6a is that the heat flux in the z -direction (proportional to the temperature gradient) is very small and consequently the metal recess plays a minor role in the structure's thermal behavior when the laser is mounted in the present configuration (substrate down).

After the location of the heat sources and the heat power associated with each of them have been established, the thermal problem can be schematically represented in the form shown in Fig. 5.7.

Figure 5.7: Schematic representation of the thermal problem showing the heat flow pattern.

Given the simplicity of the structure and the location of the heat sources, it is evident that most of the heat will flow downwards through the substrate into the heat sink. A part of it, however, will flow upwards to the gold layer where it will diffuse first and then flow into the heat sink following either path, it will flow downwards through the oxide, substrate and other layers or upwards through the bonding wires. In this scenario, it is clear that there is little to be done in the lower part of the structure to optimize its thermal behavior

since there are no layers between the active region and the substrate to maneuver with. On the other hand, the upper part of the structure is more flexible and suggests that two possible ways to proceed with the thermal optimization problem are to enhance the heat flow from the metal to the heat sink or, to remove efficiently the heat that makes it to the metal layer through the top of the device. In this section the two cases are considered and explained in some detail below.

B) Analysis of the Silicon Dioxide Layer

An obvious way to intensify the heat flow from the top metal to the heat sink is to reduce the thermal resistance of the heat path. A possible way to do this is either, reduce the thermal resistance of the Silicon Dioxide layer by making it thinner, or replace it by another insulator with similar electrical properties but better thermal conductivity such as Silicon Nitride. Before any attempts to pursue this structural modifications are made, it should be investigated whether they are physically possible and/or justifiable. Fig. 5.8 shows the longitudinal temperature profiles for different thermal conductivities of the Silicon Dioxide layer. In order to get a better picture of the importance of this layer, the bonding wires were excluded from the simulation domain to force the heat in the metal layer to flow across the oxide downwards into the heat sink. From this figure one can see that the temperature reduction in the active region for the case when the thermal conductivity of the SiO_2 layer is increased up to 10 times its original value is small. This indicates that the amount of heat that flows laterally through the vertical walls of the ridge and upwards to the metal layer and then down into the heat sink is small compared with the amount of heat that flows straight down into the heat sink.

Figure 5.8: Longitudinal temperature profile for different thermal conductivity of the Silicon Dioxide layer.

Fig. 5.9 shows a contour plot of the heat flow pattern in the vicinity of the ridge and active region. This figure corroborates the idea that even though the Silicon Dioxide layer is very thin, it plays a very important role in the thermal behavior of the structure owing to its low thermal conductivity. In order for this layer to be effective in the heat removal process, its thermal conductivity would have to be increased beyond what conventional materials can offer.

This result suggests that there is no point in pursuing this structural modification because, although increasing the thermal conductivity of the SiO_2 layer within reasonable limits may not be a difficult task, the predicted improvement does not justify it. It is time

then to consider the second possibility, namely that of removing the heat through the top of the chip.

Figure 5.9: Heat flow pattern in the vicinity of the ridge and active region.

C) Analysis of the Bonding Wires

As shown in Fig. 5.3, in the actual laser structure two gold wires are bonded from the chip top metal to the AlN carrier. The wires are intended to aid the backend contact in the heat removal process through the top of the chip. Due to their thickness however, their efficiency was questionable and needed to be investigated. To this end, the wires were included in the *Atar* model and approximated by a square bump with the same cross-section area. Since the wires are connected to the heat sink, a fixed temperature boundary

condition was applied. Fig. 5.10 shows the *Atar* model after the wire was added on top of the structure. The figure shows the case of the wire located on the edge of the chip.

Figure 5.10: *Atar* model of the laser structure with the bump representing the wire on the edge.

Fig. 5.11 shows the longitudinal temperature profiles in the active region for three different positions of the bump representing the wire. As can be seen there, a considerable temperature reduction is obtained when the bump is placed right on top the ridge (active region). This is so because the heat generated in the ridge flows upwards directly through the bump into the heat sink without inducing heating in the active region. As the bump is moved towards the edge of the chip, its efficiency decreases since the thermal path and

hence the thermal resistance between the bump and the heat sources (located in the active region and in the ridge), increases. As the thermal resistance increases, part of the heat will find its way down easier and will flow through the active region and substrate to the back-end heat sink. This results in higher temperatures in the active region as compared to the previous case when the bump is sitting on top of the ridge.

Figure 5.11: Longitudinal temperature profile for different locations of the bump representing the bonding wire.

From Fig. 5.11 it seems obvious that the next step in improving the thermal behavior of the structure should be that of increasing the size of the bonding wires, specifically for the case when they are located on top of the ridge. Although this structural modification appears appealing from a thermal standpoint, a trade-off comes into play and it is not

physically possible to implement because the presence of the extra metal on top of the ridge would increase the mechanical stress in the active region.

Proceeding with the thermal optimization process, in the limit, the protuberance representing the wire was replaced by a metal stripe that extended all the way along the laser structure to see how this could help to reduce the temperature in the active region. For the reasons explained above, only the case when the stripe is located off the ridge was considered. Fig. 5.12 shows the *Atar* model for this case.

Figure 5.12: *Atar* model of the laser structure with a stripe on the edge of the chip.

Fig. 5.13 shows the corresponding longitudinal temperature profiles for this structural modification. To see how the size of the stripe could improve the temperature reduction,

two different stripe widths were considered. This figure shows that indeed the temperature in the active region can further be reduced by replacing the existing bonding wire by a metal stripe which sits on the edge of the chip and extends all the way along the laser structure.

Figure 5.13: Longitudinal temperature profiles for different stripe dimensions.

D) Final Modification

The structural modification just presented predicted a reduction of the laser's temperature and also seemed to meet all the technological requirements for its physical implementation. In the simplest case, the stripe could be implemented with an array (strap) of bonding wires connected from the top of the device to the AlN carrier. As the size of the stripe continues to increase, however, its implementation is not quite feasible, because for

it to be effective a real good contact between the stripe and the heat sink must exist. This suggested that if the optimization process was to continue to move in this direction a flip chip configuration should be sought. By further changing the dimensions and structure of the stripe, a final possible modification to reduce the temperature in the active region was obtained. Fig. 5.14 shows a slice of the *Atar* model (a) of the original laser structure and (b) after it was modified. Unlike the case of having a regular top boundary condition (flip chip), in the proposed structure no physical contact between the ridge and the stripe metal exists, condition that, as was brought up before, is important to keep the active region free of extra mechanical stress. This fact is emphasized by the presence of the thin metal layer on top of the stripe shown in Fig. 5.14b.

(a)

(b)

Figure 5.14: Atar model of the (a) original and (b) modified laser structure.

Fig. 5.15 shows a comparison of the transverse (y -direction in Fig. 5.1) temperature profiles for the previous structures. This figure was obtained by plotting a temperature line from the top of the laser to the backend contact (heat sink) in a similar way to that shown in the inset of Fig. 5.6. As can be seen there, by extracting the heat through the top of the laser, the temperature is practically reduced everywhere inside of the device. The highest point in the plot corresponds to the temperature in the active region. A reduction of this value of approximately 17% is predicted by the modified structure with respect to the original one. To see how the proposed structure compares to the best case scenario for the

laser, namely that when the top and back boundary conditions are present simultaneously, the transverse temperature profile for this case is also shown in the figure.

Figure 5.15: Transverse temperature profiles of the original and modified laser structures.

As the figure shows, by having both boundary conditions applied simultaneously, a temperature reduction of approximately 25% with respect to the original structure is predicted by the model. This is only 7% better than the proposed structure and indicates that under these conditions, no considerable improvement can be obtained by removing heat from the back of the chip. On the other hand, the double boundary condition is not physically possible to implement because in this situation the chip would be loaded with con-

siderable mechanical stress and, several packaging problems would arise. The details are out of the scope of this work.

The pure thermal analysis of the laser structure performed in this section ignored the electro-opto-thermal interactions that occur mainly in the active region of the semiconductor laser. To account for these, the self-consistent model described in Chapter 4 needs to be used. Before starting with the self-consistent simulations, however, the importance that diffusion of carriers plays in the different scenarios imposed by the simulated operating conditions addressed in this thesis, needs to be assessed.

5.2.2 Assessment of the Importance of Diffusion of Carriers in Self-consistent Simulations

The self-consistent model was intended to be used for investigating the thermal behavior of lasers under both, normal operating conditions and those which bring it close to the breaking point. Since there is not complete agreement in the literature as whether or not longitudinal diffusion of carriers is important [28], a small analysis was performed to determine the importance that this effect plays for different laser structures (and materials) under the two different operating conditions. To this end, a comparison between the solutions of equations 4.13 and 4.17 for given temperature profiles was performed. The results obtained from this analysis provided guidance when determining which electrical equation to use during the self-consistent modeling to meet the objectives of the different simulations performed.

i) Low power operating conditions

We can regard the temperature profile shown in Fig. 5.16 as a typical one for moderate power operation conditions (it was obtained self-consistently by using this model and the details are not important at the moment) and a small amount of facet heating for a symmetric laser (same facet reflectivity).

Figure 5.16: Typical longitudinal temperature profile for moderate power laser operation.

The corresponding carrier density profiles for three different values of the carrier mobility are shown in Fig. 5.17. The dot-dashed and dashed lines are the solutions of the diffusion equation (4.13) and correspond to mobility values of GaAs and Si respectively. The solid line shows the solution of the equation with no diffusion (4.17) and corresponds

to a material with zero mobility of carriers. From this figure one can see that for low power laser operation the carrier density profiles do not differ significantly from one another.

Figure 5.17: Corresponding carrier density distributions for the temperature profile of Fig. 5.16 and different carrier mobility.

This result could have been anticipated from the fact that when the temperature distribution is almost uniform (± 2 K), the carrier density profile is mainly shaped by spatial hole burning [6]; and as the longitudinal light intensity distribution is quite uniform for a symmetric laser (as will be demonstrated latter) this results in small carrier gradients. The solution differs a little at the edges because the small temperature increase in those regions gives rise to a small gradient of carriers which is smoothed out by diffusion. We can also see that even though the degree to which the solutions differ from one another depends on

the value of the carrier mobility that is used, the difference is practically negligible in all cases.

From this comparison we can conclude that diffusion of carriers can be neglected under low power operating conditions regardless of the material the active region is made of. In this case, the equation modeling the electrical phenomenon reduces to a cubic polynomial whose solution can be analytically obtained in a very short time. A trade off comes into play however, because the implicit boundary condition of zero flow of carriers at the facets can not be changed and the thermal runaway phenomenon can not be modeled as a function of carrier diffusion towards the facets.

ii) High power operating conditions

Similarly, we can regard the longitudinal temperature profile shown in Fig. 5.18 as a typical one for high power laser operation (close to run-away) and high surface recombination velocity and realize the same comparison as above. The corresponding solutions for this temperature profile are shown in Fig. 5.19. As can be seen there, the highly non-uniform temperature distribution generates considerable gradients of carriers and these in turn cause diffusion. If these carrier density profiles were to be used in successive iterations of a self-consistent calculation, the final result that would be obtained for each case would greatly differ from one another. This difference would be gradually built up by the fact that the optical and thermal solutions obtained from the different carrier density profiles at each iteration would introduce accumulative error in the final self-consistent solution. It should be noticed that the carrier density profiles displayed here do not correspond to real self-consistent steady state solutions and are rather arbitrary. This is due to the fact

that the diffusion equation was solved subject to a boundary condition of zero flow at the facets.

Figure 5.18: Typical temperature profile for high power laser operation.

In reality, the facets act as sinks for the carriers (which in principle raises the temperature there) and a different boundary condition should have been applied if a complete self-consistent solution wanted to be obtained (see Eq. 4.16). Unfortunately, those boundary conditions can not be straightforward implemented in the non-diffusion equation, and a comparison between the two equations would not be possible. The results are still valid though, because Fig. 5.19 shows the difference among the predicted carrier densities when equations 4.13 and 4.17 are solved for a prescribed (physically possible) temperature profile.

Figure 5.19: Carrier density distribution (a) longitudinal and (b) detailed view.

From this analysis we can conclude that diffusion of carriers plays an important role in high power laser operation and it should be included in the model if more accurate results are desired. Its importance is emphasized by the fact that different materials have different carrier mobility and this can not be taken into consideration if the diffusion term in the electrical equation is not included. The inclusion of this term brings along an increase in the computational effort since the electrical equation must now be solved iteratively. Nevertheless, convergence is achieved relatively fast (10-15 iterations) and the overall increase in time is very small.

5.2.3 Self-consistent Approach

After the importance of diffusion of carriers under different scenarios has been demonstrated, we are finally in a position to use the self-consistent model to study the laser structure shown in Fig. 5.1 for the case of equal facet reflectivity. Since the simulations are performed at nominal operating conditions, it follows from the results obtained in the previous section that the diffusion term in the electrical equation can be dropped out. To focus on the effect that the longitudinal optical power distribution has on the temperature profile and, supported by the results obtained in Section 5.2.1, the top metal recess was not included in the self-consistent analysis presented in this section. Its importance is investigated, however, in the self-consistent analysis of the optically asymmetric structure presented in Section 5.3.2. The *Atar* structure used for the simulations of this section is the same as that shown in Fig. 5.4 but with the top metal extending over the entire surface.

Unless otherwise stated, the parameters used for the simulations are those shown in Table 5.2 where some of them are reliable and others best estimates.

Table 5.2: Parameters used for the self-consistent simulations.

Parameter	Symbol	Value
Non-radiative Coefficient.	A_{nr}	$1.4 \times 10^8 \text{ s}^{-1}$
Activation Energy	E_1	$8.61 \times 10^{-3} \text{ eV}$
Radiative Coefficient	B	$1.4 \times 10^{-10} \text{ cm}^3/\text{s}$
Activation Energy	E_2	$8.61 \times 10^{-3} \text{ eV}$
Auger Coefficient	C	$1.3 \times 10^{-28} \text{ cm}^6/\text{s}$
Activation Energy	E_3	$5.16 \times 10^{-2} \text{ eV}$
Wavelength	λ	$1.5 \text{ } \mu\text{m}$
Spontaneous Emission Coupling Coefficient	β	8×10^{-5}
Optical Confinement Fac- tor	Γ	0.3
Transparency Carrier Den- sity	n_o	$3.0 \times 10^{19} \text{ cm}^{-3}$
Activation Energy	E_{int}	$8.61 \times 10^{-2} \text{ eV}$
Internal Losses	α_{int}	20 cm^{-1}
Group Refractive Index	η_g	4.0
Gain Coefficient	a	$2.5 \times 10^{-16} \text{ cm}^2$

Fig. 5.20 shows the longitudinal carrier density distribution for the laser structure shown in Fig. 5.1 when a reflectivity of 0.32 is used for both facets. Since the temperature is rather uniform (see Fig. 5.23) and diffusion effects are neglected, the shape of the carrier density profile is mainly determined by spatial hole burning, that is, the carrier density is smaller in regions where the optical intensity is higher and vice-versa. This is so because

as the light intensity increases, more electrons recombine through stimulated emission and a diminishment of the electron population in those regions results.

Figure 5.20: Longitudinal carrier density profile.

Fig. 5.21 shows the light intensity profiles of the forward $I^f(z)$ and backward $I^b(z)$ running waves. Since the facet reflectivities are equal, the laser is optically symmetric and so are the waves. This figure also shows that the total photon density for the symmetric laser is also rather uniform. The minimum intensity occurs right in the middle of the laser and the maximum at the facets. The difference between these two extremes is approximately 15% and this results in a weak shaping of the carrier density profile due to spatial hole burning (see Fig. 5.20).

Figure 5.21: Light intensity distribution of the forward $I^f(z)$ and backward $I^b(z)$ running waves.

In turn, the longitudinal power distribution in the active region is determined by the solution of the electrical $n(z)$ and optical equations (I^f and I^b) through Eqs. 3.10 and 3.11. Fig. 5.22 shows the heat power contribution of the different heat sources in the active region. As can be seen from Eq. 3.10, the non-radiative recombination processes and the absorption of spontaneous radiation are proportional to the carrier density and as a consequence, these heat sources have a distribution profile similar to that of n . On the other hand, the heat power due to absorption of stimulated radiation is proportional to the photon density S and this results in the spatial distribution of this heat source being similar to

that of S . It should be emphasized at this point that the parameters of Table 5.2 were adjusted so that the heat power contribution of the heat sources located in the active region would account for 50% of the total heat power dissipated at the simulated bias point. The remaining 50% was accounted for by Joule heating and was uniformly distributed in the volume of the ridge.

Figure 5.22: Heat power contribution of the different sources in the active region (account only for 50% of the total heat power).

As can be seen from Fig. 5.22, at nominal operating conditions the most important contribution to the heat power in the active region comes from the non-radiative processes followed by absorption of stimulated emission. The heat generated by absorption of spontaneous radiation is small compared with that generated by the other two mechanisms and

it is usually neglected in most semiconductor lasers. It should be noticed that the magnitude of the heat sources shown in Fig. 5.22 is bias dependent and does not remain constant. In general, the heat power contribution from non-radiative processes increases faster with current and temperature than the contribution from absorption processes. This is so because of the cubic dependence on the carrier density of Auger recombination. It should also be noticed that the spatial distribution of the two main heat sources have opposite profiles and thus partially cancel each other thereby helping to flatten the profile of the total heat power.

5.2.4 Self-consistent vs. Pure Thermal

After the laser structure of Fig. 5.1 has been analyzed both purely thermal and self-consistently, a comparison between the temperature profiles obtained by either approach can be made. Fig. 5.23 shows the longitudinal temperature profiles in the active region that were obtained in each case. From this figure it is obvious that the spatial temperature variation induced by the non-uniform power distribution shown in Fig. 5.22 is very small. The difference between the extremes is less than 0.5 K and this makes it negligible for most practical purposes. For this reason, we can conclude that for a structurally and optically symmetric laser a pure thermal approach is sufficient to determine its thermal behavior. The location of the existing heat sources and their corresponding contribution to the total heat power should properly be assigned as explained in Section 5.2.1. This result is really important since it directly translates into a considerable reduction of the computational effort and in a simplification of the modeled problem.

Figure 5.23: Comparison between the self-consistent and pure thermal solutions.

5.3 Analysis of a Stand-alone Asymmetric Laser Structure

As it was mentioned earlier in this chapter, the real laser structure shown in Fig. 5.1 is optically asymmetric due to its having different facet reflectivities. In this section, the degree to which the results obtained in Sections 5.2.3 and 5.2.4 for a completely symmetric laser change, owing to the optical asymmetry, is investigated. For an asymmetric laser, it is not possible to perform a pure thermal analysis which accounts for unequal facet reflectivity because the assumption of uniform heat power distribution leads to the previously described symmetric case. Since there is no a straightforward method to recalculate

the heat power distribution in the active region to include the effect of the optical asymmetry, a complete self-consistent analysis must be performed to determine how the heat power profile changes for this case.

5.3.1 Self-consistent Approach

Fig. 5.24 shows the longitudinal carrier density profile for the laser structure of Fig. 5.1 for right and left facet reflectivities of 0.03 and 0.75 respectively.

Figure 5.24: Longitudinal carrier density profile.

Once more, the simulation is performed at normal operating conditions and for these, the expected gradients and temperature rise are small (see Fig. 5.28). This results in the carrier distribution being mainly shaped by spatial hole burning similarly to the symmetric case.

By comparing Figs. 5.24 and 5.20 one can see that the carrier density variation due to spatial hole burning is more severe for the asymmetric laser than for the symmetric one. This is of course very closely related to the optical intensity distribution which as seen before, turned out to be quite uniform for the symmetric case.

Fig. 5.25 shows the light intensity distributions of the forward $I^f(z)$ and backward $I^b(z)$ running waves for the asymmetric laser.

Figure 5.25: Light intensity distribution of the forward $I^f(z)$ and backward $I^b(z)$ running waves for the asymmetric laser.

As can be seen here, the different reflectivity of the facets results in the forward and backward travelling waves having very different spatial distributions. The right facet

of low reflectivity (0.03) allows that most of the optical radiation of the forward travelling wave leave the laser cavity to produce the useful output light beam. The small portion that is reflected becomes the backward travelling wave and is slowly amplified (the amplification process being proportional to the irradiance of the beam and the travelled distance according to Chapter 2) as it passes across the active medium until it reaches the surface of the left facet. At this point, 75% of the wave intensity is reflected and 25% is transmitted through the facet. The portion of radiation which is reflected becomes the forward travelling wave and is amplified at a faster rate (due to its higher irradiance) as it travels across the active medium towards the right facet. This process continues until the light intensity distribution shown in Fig. 5.25 results. As can be seen there, the light intensity has its minimum value on the side of the laser where the high reflectivity facet is located and vice versa. For this case the difference between the two extremes is approximately 35% and this results in a stronger shaping of the carrier density profile due to spatial hole burning (see Fig. 5.24).

Fig. 5.26 shows the heat power contribution of the different heat sources in the active region for the asymmetric case. As can be seen there, once again the heat power generated by non-radiative recombination processes and by absorption of spontaneous radiation has a distribution profile similar to that of n while the heat power due to absorption of stimulated radiation has a spatial distribution similar to that of S . Once more, the smallest contribution to the total heat power comes from absorption of spontaneous radiation but, unlike the symmetric case, this mechanism can not be readily neglected considering that in the vicinity of the high reflectivity facet the heat contribution due to absorption

of stimulated radiation is even smaller. Similarly to the symmetric case, the main heat sources are spatially distributed in such a way that partially cancel one another. This results in a relatively flat heat power profile but, contrary to the symmetric case, the difference between the maximum and minimum intensity is bigger than 20%. Finally, it should be noticed that the fact that non-radiative recombination of carriers dominates the heat generation process results in the maximum heat power being generated in regions where the optical intensity is the smallest.

Figure 5.26: Heat power contribution of the different sources for the optically asymmetric laser.

In general, Fig. 5.26 shows that the heat power distribution in the active region is relatively uniform for the laser structure of interest. As it was emphasized before, this dis-

tribution is mainly the result of having different facet reflectivities and these are well known laser parameters. On the other hand, its shape also depends on the selection of parameters used for the simulations (Table 5.2) and given the wide range of values found in the literature, a certain amount of uncertainty is introduced thereby.

To illustrate how the optical asymmetry introduced by the different facet reflectivities affects the longitudinal temperature profile, Fig. 5.27 shows the temperature distribution in the active region.

Figure 5.27: Longitudinal temperature distribution for the asymmetric laser.

The dashed line in Fig. 5.27 represents the average temperature that would be obtained if the spatial variations of the heat power were neglected. By comparing Figs. 5.27 and 5.23 we can see that the spatial variation of the longitudinal temperature is more

severe for the asymmetric laser than for the symmetric one, and can not be readily neglected. In the following section a more detailed analysis and discussion of the resulting longitudinal temperature profiles under different scenarios is presented.

5.3.2 Self-consistent Analysis of the Top Metal

Structurally speaking, the proximity of the top metal to the heat sources offers the possibility of decreasing the effect of the non-uniform power distribution on the resulting temperature profile by means of an efficient redistribution of the heat power (shortening effect). In order to investigate this, several self-consistent simulations were performed for different conditions of the top-metal and the results are shown in Fig. 5.28. As can be seen from this figure, the temperature difference (ΔT) from one end of the laser to the other does not depend considerably on the top metal conditions. This is so because as it was appointed during the pure thermal analysis of this laser, the effect of the top metal in the temperature distribution is moderate when the device is mounted substrate down. This result still holds here and the role that the metal plays in this scenario does not help to considerably flatten the temperature profile in the active region. Yet, the maximum ΔT is obtained when the top metal is completely removed (thin and thick Gold layers in Fig. 5.1) since in this case the heat redistribution process is not present at all. The solid line shows the temperature profile of the actual laser structure (both metal layers present and a recess of 25 μm). In this case, a ΔT of around 2.6 K is obtained and this corresponds to approximately 17% of the maximum temperature rise. Fig. 5.28 also confirms that the presence of the metal recess plays a small effect in the laser's thermal behavior even for the self-consistent case since the local temperature rise associated with it, is approximately 0.5 K.

Figure 5.28: Self-consistent temperature profile for different top-metal conditions.

It should be emphasized that these observations can not straightforward be generalized to other semiconductor lasers with similar constructions since the temperature difference obtained here does not only depend on bias conditions, but also on various factors such as the specific laser structure that is being analyzed and the materials the laser is constructed of. This is so because in general, the parameters shown in Table 5.2 greatly differ from one laser to the other. For this reason, it is quite possible that under certain circumstances one of the heat sources may completely dominate the others with a resulting increase in the heat power difference (and temperature) from one end of the laser to the other. Conversely, for a given laser structure the different heat sources involved might display a better cancellation of one another with a corresponding highly uniform heat power

distribution. To illustrate the sensitivity of ΔT to the laser parameters. Fig. 5.29 shows the temperature profiles for the laser structure of Fig. 5.1 and the values of Table 5.2 for two different substrate materials.

Figure 5.29: Temperature profile for the laser structure of interest for different substrate material.

Fig. 5.29 shows that by simply changing the substrate thermal conductivity of the laser, the ΔT increases from 2.6 K to approximately 3.6 K. While the percentage of this difference with respect to the highest temperature remains almost the same (approx. 17%), the absolute temperature difference from one end of the laser to the other increases by 1 K. This comparison is intended to show that in order to determine the importance that longi-

tudinal thermal effects play in optically asymmetric lasers, each laser structure should be studied separately subject to the appropriate modeling parameters for each case.

It should be noticed that the results obtained in this section show that the longitudinal temperature profile for the laser structure of Fig. 5.1 is relatively uniform even for the asymmetric case. After the predicted temperature difference (17%) between facets is known, it is up to the designer to determine whether this difference is important or not (depending on the application the laser is aimed to) and to decide if further effort should be put into its minimization. Should this be the case, the longitudinal self-consistent model becomes indispensable and the complications of selection of parameters and their associated temperature dependence, as well as the long simulation times involved, must be dealt with. On the other hand, if the predicted temperature difference can be neglected, the thermal behavior of the laser can be investigated (at least qualitatively) by using the pure thermal model that was described before.

5.4 Thermal Runaway Analysis

5.4.1 Overview

Thermal runaway has been observed in many physical systems where a temperature dependent heat source is present. The origin of this problem is positive feedback when the heat power generated increases with rising temperature at a faster rate than that at which the heat is removed from the system. Although this effect has been found useful for some applications [29], it is mostly undesirable since the heat power generation due to

the positive feedback mechanism continues to increase in an uncontrolled way until, for a certain critical temperature value, catastrophic damage of the device results.

Thermal runaway has widely been observed in events ranging from chemical reactions, to electronic components such as Bipolar Junction Transistors (BJT) and recently in Power MOSFET's [30]. Semiconductor lasers also exhibit thermal runaway leading to catastrophic damage of the laser facets thereby limiting the high power operation of these devices. This problem is believed to be caused by the intensive heating in the vicinity of the facets originated by the non-radiative recombination of charge carriers and other heat sources to be discussed latter. Therefore, the facet heating process must be carefully investigated in order to understand and minimize this problem.

The feedback mechanism in semiconductor lasers arises from the strong temperature dependence of one of the physical processes supplying carriers into the active region, namely reabsorption of radiation. Although non-radiative recombination processes are present everywhere inside the active region, these processes take particular importance in regions close to the cleaved facets where they are enhanced by the high concentration of mechanical defects which are formed during cleavage or are thermally induced during laser operation [31].

The issue of modeling facet heating in semiconductor lasers has been addressed by several authors and both, analytical and numerical approaches have been taken. In [3], [32] a one-dimensional fully self-consistent model to investigate facet heating was presented. A modified heat conduction equation was used to obtain the temperature distribution in the active region and a parameter to phenomenologically describe the heat flow to

the heat sink was introduced. In [28], a similar approach was taken but this time the laser structure was thermally analyzed by partitioning it into two areas, each described by a one-dimensional heat equation. Unfortunately, it will be demonstrated later that one-dimensional models tend to greatly overestimate facet heating and thus, a better thermal model should be used to study this effect under more realistic bases. On the other hand, the existing three-dimensional thermal models [4],[33] do not handle longitudinal variations of photons and carriers self-consistently and reduce the simulation domain to a small area in the vicinity of the facets called the catastrophic degradation (CD) area, thereby limiting the investigation of this effect from a more general perspective that includes the scale of the real thermal domain. An immediate limitation of these models is that the heat generated far from the CD area can not be straightforward taken into consideration nor can the thermal importance of layers non-adjacent to the active region be clearly assessed. Correct thermal modeling of facet heating requires then a 3D model and a large simulation region that includes the structural details of the laser.

In this section, the numerical self-consistent thermal/optical/electrical model presented in Chapter 4 is used to investigate facet heating in the ridge semiconductor laser structure shown in Fig. 5.1. Firstly, a comparison is made between the predictions of the present model and those predicted by similar ones using lower dimensional thermal equations. Next, the importance of a non-linear thermal conductivity and Joule heating in the facet heating problem is investigated. Finally, a comparison of the facet heating effect between InGaAsP/InP and AlGaAs/GaAs lasers is presented.

Due to the fact that thermal runaway is more likely to occur in GaAs than in InP lasers, the laser structure of Fig. 5.1 was replaced by its GaAs counterpart as shown in Fig. 5.30.

Figure 5.30: AlGaAs/GaAs laser structure used to study facet heating.

In addition, the laser structure was simulated with an equal facet reflectivity of 0.32. This allowed the reduction of the simulation domain to a quarter of the original structure. The reduction was particularly important for the solution of the 3D thermal equation because most of the computational effort lay with it.

Although *Atar* automatically produces a finer mesh in regions where it is needed such as in the vicinity of the heat sources, an additional refinement was made near the facets to correctly model the rapid temperature changes occurring there. Fig. 5.31 shows a lateral view of the discretized structure in the facet region. As the figure shows, the mesh is gradually coarsened as one moves longitudinally away from the facets.

Figure 5.31: Discretization in the facet region.

The temperature rise at the facets of a laser greatly depends on the values of the laser parameters used for the simulations and on their temperature dependence as well. Thus, they must be carefully chosen to avoid unrealistic estimations. Unless otherwise stated, the parameters used for the simulations are the same as those shown in Table 5.2, except for the following modifications (Table 5.3).

Table 5.3: Correction values for the AlGaAs/GaAs laser.

Parameter	Symbol	Value
Auger Coefficient	C_0	$3.25 \times 10^{-29} \text{ cm}^6/\text{s}$
Wavelength	λ	$0.88 \text{ } \mu\text{m}$

Similarly, the thermal conductivities of the materials were changed according to Table 5.4.

Table 5.4: Thermal conductivities used for the simulations.

Material	κ_0 , W/mK
GaAs	44
AlGaAs	11.4

5.4.2 Modeling the heat flow - dimensionality

The predicted temperature rise at areas of local heating such as the laser facets depends on an accurate model of the heat flow from this region and unfortunately, this issue is rarely addressed in the literature. This subsection investigates how the thermal approach that is taken to investigate facet heating in a laser, greatly influences the overall result. Specifically, we focus on the dimensionality of the heat equation comprising the self-consistent laser model. To this end, a similar approach to that shown in Ref. 3 is ini-

tially taken and Eq. 3.13 is replaced by a one-dimensional heat equation used in this reference. To be consistent with the formulation of that model, only the heat power generated within the active region is initially considered. Next, a two-dimensional thermal approach is taken by reducing the simulation domain of the laser structure of Fig. 5.30 to a slice with the same width as that of the ridge. To model the lateral spreading of the heat in the x direction as it flows from the active region to the bottom of the die, the thermal conductivity of the substrate was replaced by an effective thermal conductivity κ_{eff} . The value of this conductivity will be substantially lower than that of the bulk as it represents the increasing cross-sectional area of the heat flow as it moves towards the heat sink. Finally, the complete laser structure shown in Fig. 5.30 is simulated with the laser model comprising the full heat equation without any structural simplifications. Fig. 5.32 shows a schematical representation of the three different approaches taken to simulate facet heating.

Figure 5.32: Schematical representation of the three different thermal approaches taken to simulate facet heating, (a) 1-D, (b) 2-D and (c) 3-D.

In Fig. 5.32a the 1D approach is presented. The heat flows longitudinally along the active region (the laser is represented by a line of nodes). Each node is connected to its neighbors allowing for a longitudinal heat flow along the laser from the hot facet. Between each node, a thermal resistance R_b given by $R_b = L_b / \kappa(T)A_b$ is connected, where L_b and A_b are the block's length and cross-section area respectively. The heat flow down to the fixed temperature at the bottom of the die is represented by a thermal resistor R_F from each node to the fixed boundary condition. The chief limitation of the 1D model is the lack of heat spreading for the heat flowing from the facet. The 2D model is presented in Fig. 5.32b. The model is a slice of the laser with a thickness equal to the ridge. The metal contact at the top is included in the model. To model the spreading of the heat flow in the x direction as it flows downwards towards the heat sink, an effective conductivity $\kappa_{eff} = W_{eff} \times \kappa$ is used to characterize the region under the ridge. The heat flow from the facet in this model was determined using a 2D version of *Atar* and would be purely thermal with spreading present in only the y and z directions. A full 3D model is presented pictorially in Fig. 5.32c. For this case, a complete *Atar* model was built and the heat flow from the hot facet would spread in all three directions.

Fig. 5.33 shows the temperature distribution along the active region predicted by the laser model for each of the different thermal approaches described above. To instigate high facet heating, a surface recombination velocity V_{sur} of 3×10^6 cm/s and an input current of 0.5 A were used in the calculations. In addition, to visualize better the very different facet temperatures predicted by each model, the phenomenological parameter R_F of

the 1D thermal equation of the fully 1D laser model and the effective thermal conductivity of the substrate for the 2D model κ_{eff} , were adjusted so that the temperature in the bulk of the laser matched that predicted by the laser model using the 3D thermal equation.

Figure 5.33: Temperature profiles for different dimensionality of the heat equation.

From Fig. 5.33 it is apparent that the facet temperature predicted by the 1D thermal model is greatly overestimated as compared with that predicted by the 3D thermal model. This suggests that the phenomenological parameter that accounts for heat flow towards the heat sink fails to capture the three-dimensional nature of the heat removal process in the vicinity of very hot spots such as the facets. Conversely, the solution of the 2D thermal model turned out to be very similar to that of the 3D model for this specific laser structure. Contrary to the 1D case, the temperature at the facets is slightly underestimated

as compared with that predicted by the 3D model and this can be explained as follows. The 3D heat flow from the active region to the heat sink is unconstrained in the x direction and will spread out as it moves down through the wafer. The effective width W_{eff} of this flow can be quite large. This lateral spreading is represented in the 2D model by κ_{eff} and to match the bulk laser temperature to that predicted by the 3D model, a relatively large value of approximately 6 times that of gold is needed. This large value of κ_{eff} results in a lower facet temperature due to an exaggerated heat flow from the facet towards the heat sink and a lower thermal resistance for the facet region.

The complexity of the three dimensional heat flow can be clearly seen in the two plots of Fig. 5.34. The first plot shows a contour plot of the temperature distribution in a 2D Atar simulation. The contours show the 2D heat spreading from the hot facet region (the flow is, of course, perpendicular to the contours) with both longitudinal and vertical flow being present. The distortion of the upper most contour is due to the presence of the AlGaAs layer. Additionally, the figure corroborates the previous discussion about the lower thermal resistance for the facet region since the large value of κ_{eff} results in a pseudo-fixed temperature boundary condition at the top of the substrate.

The second plot is a contour plot of the 3D model in the vicinity of the facet. For this case the contours show 3D heat spreading from the facet in the lateral, longitudinal, and vertical heat flow being exhibited.

Figure 5.34: Atar temperature contour plots of the facet region of the laser,
a) 2D model b) 3D model.

Fig. 5.35 shows the corresponding longitudinal carrier density distribution in the active region for each of the temperature profiles shown in Fig. 5.33.

Figure 5.35: Longitudinal carrier density distribution.

In the vicinity of the facets, the non-radiative recombination processes are enhanced by the increased density of defects and this results in higher local temperature. As the temperature at the facets increases, so do the transparency carrier density and the carrier generation by reabsorption of radiation. In turn, the carriers diffuse towards the facets thereby increasing the temperature according to Eq. 4.17. This process continues until either, a steady state situation is reached as actually occurred in this case, or the temperature continues to increase in an uncontrolled way such that a steady state solution ceases to

exist and a runaway situation arises. By comparing the carrier density profiles of Fig. 5.35 one can see that as the temperature continues to increase, the facets can no longer sink all of the carriers generated by reabsorption of radiation. The carriers start to pile up near the facets and a bump on the carrier density profile is formed.

As the figure shows, the bump is far more severe for the 1D thermal model and originates from the poor heat removal from the facets. The presence of this bump in the carrier density profile increases the heat generated by non-radiative recombination in these regions (specially Auger recombination through its cubic dependence on carrier density) and this further contributes to the overestimation of the facet heating.

Finally, Fig. 5.36 shows a comparison among the photon density distributions in the active region corresponding to each of the temperature profiles shown in Fig. 5.33. The variations near the facets show how laser radiation is reabsorbed and this is consistent with the carrier density increase in those regions. This figure indicates that as the temperature increases, the photon density distribution shifts downwards due to the fact that the non-radiative recombination processes increase with temperature and for a constant current injection, the carriers that contribute to the useful laser beam are decreased.

From the previous figures we can see that for the case of only heat generation in the active region, the 2D approach taken here matches relatively well the results predicted by the 3D model and is a much better approximation than the 1D model. However, Fig. 5.33 shows that when only the heat generating mechanisms occurring in the active region are considered, the temperature rise at the facets predicted by the 3D model, is moderate even for a relatively large value of V_{sur} . This low temperature is far below the value that is

presumably needed to induce thermal runaway and cause catastrophic damage of the facets. Under this scenario, other heat sources become critical to explain this damage that other authors have estimated to occur in GaAs lasers at approximately the value of surface recombination velocity shown above [33].

Figure 5.36: Longitudinal photon density distribution.

5.4.3 Non-linear thermal conductivity and Joule Heating

An accurate model of the laser needs to include not only the heating in the active region, but also heat generation due to ohmic losses. Both the 2D and 3D models can be extended to include Joule heating in the ridge of the laser. However, it was found that for the 2D model the effect of Joule heating is greatly overestimated. This is so because, by

reducing the structure to a single slice, the lateral heat flow from the ridge is completely eliminated: all the heat is forced to flow downwards across the active region into the heat sink, thus increasing the temperature there. In addition, the heat spreading effect of the top metal in the x direction is also completely eliminated and this contributes to the overestimation of the temperature in the active region and facets as well. The combination of an artificial constraint of 2D heat flow; the temperature dependence of the thermal conductivity and heat generation terms; and Joule heating produces a very strongly coupled thermal system. In our case, when Joule heating was included in the 2D model no steady state solution was reached. The temperature increased far beyond the melting point of the facet materials in what could be thought of as a runaway situation.

This situation can be contrasted with the more realistic situation of the 3D model where a numerically stable solution was found. To illustrate the importance that Joule heating and the temperature dependence of the thermal conductivity play in the overall heating process, Fig. 5.37 shows the longitudinal temperature distribution along the active region for three different scenarios. When Joule heating is included in the model and a temperature dependent thermal conductivity is used, the facet temperature increases by approximately 300 K whereas the temperature in the bulk of the laser is only increased by approximately 130 K. A larger temperature increase at the facets occurs because the temperature rise induced by this new heat source and by those located in the active region are added non-linearly owing to the temperature dependence of the thermal conductivity and the heat sources. On the other hand, when only Joule heating generated in the ridge is considered and the thermal conductivity is kept constant, the temperature in the active region

and the facets is essentially increased linearly because the temperature dependence of the heat sources (mainly through absorption of radiation at the facets) is still weak at these temperatures.

Figure 5.37: Effect of Joule heating and non-linear thermal conductivity on the temperature distribution.

Fig. 5.37 shows that the inclusion of Joule heating by itself does not considerably raise the temperature at the facets if the thermal conductivity of the laser materials is assumed to be constant. However, when the non-linear temperature dependence of the thermal conductivity is included, the result is high facet temperatures leading to a possible increase of the density of defects producing facet degradation.

5.4.4 Comparison between an InP and a GaAs laser

To further demonstrate the importance of studying facet heating with a 3D thermal model that includes the entire thermal domain, a comparison between an InGaAsP/InP and an AlGaAs/GaAs laser is made next. The laser structure of Fig. 5.30 is used here and in order to minimize the uncertainties introduced by the selection of parameters for each laser, the comparison is only focused on those basic properties of the lasers which influence their thermal behavior the most. To start with, only the thermal conductivity of the substrate was changed to that of InP and the laser parameters were kept the same as those shown in Tables 5.3 and 5.4. Fig. 5.38 shows the longitudinal temperature profile in the active region for the two structures (denoted as InP and GaAs in the figure).

Figure 5.38: Longitudinal temperature distribution for different laser structures.

As the figure shows, the higher thermal conductivity of InP plays a very important role in the laser heating process. The predicted facet temperature for the InP laser is considerably lower (approximately 200 K) than that predicted for the GaAs laser and this difference is due only to the different thermal conductivity of the substrate. Although the surface recombination velocity used for the calculations can be considered high compared with typical values used for InP lasers, the facet temperature is still relatively low. Initially, the heat power generated at the facets is the same for both lasers, as the heat starts to flow away from the facet region, the heat removal process is more efficient in the InP laser and a lower facet temperature results. At steady state, the non-linear temperature dependence of the thermal conductivity is crucial causing the GaAs laser to have a much higher facet temperature.

Fig. 5.38 also shows the longitudinal temperature profile for a more realistic InP laser. In addition to changing the thermal conductivity of the substrate, the surface recombination velocity was decreased to 3×10^5 cm/s and the energy gap and Auger recombination coefficient were changed according to the values of InP shown in Table 5.2. As can be seen from the figure, the combination of better thermal properties and lower V_{sur} for the InP laser results in very low temperature at the facets. The first factor appears to be dominant, however, because after the surface recombination velocity is set again to 3×10^6 cm/s, only a moderate temperature increase at the facets is observed. The lower temperature in the bulk of the laser is due to the fact that even though the non-radiative recombination rate is increased by Auger recombination, it does not compensate for the power reduction

introduced by the smaller energy $h\nu$ associated with each act of non-radiative recombination. This result would imply that the resistance of InP lasers to facet degradation is not a justification in itself for the assumption of a low surface recombination velocity.

5.5 Analysis of Semiconductor Lasers in Integrated Optical Environments (IOE)

From the pure thermal and self-consistent analyses of the laser structure shown in Fig. 5.1 that were presented in Sections 5.2 and 5.3 it became apparent that in order to investigate the thermal importance of some structural issues such as the fact of having different facet reflectivities and the presence of the bonding wires and the top metal recess, a 3D thermal model must be used. While a 3D thermal model is capable of modeling the real laser structure without resorting to excessive simplifications, it is also difficult and computationally expensive to implement. For this reason, most authors avoid dealing with such complications [1], [2] by neglecting those structural variations and by reducing the simulation domain in one or two dimensions. As it was also shown from the results obtained in those sections, the thermal importance of some of the structural issues demanding for the 3D thermal model, such as the metal recess, proved indeed to be almost negligible. Furthermore, if the analysis of the semiconductor laser is performed at the qualitative level, the structural details of the structure play a minor role in the overall laser's behavior and can be dropped out from the simulation domain with the consequent result that a 3D thermal model is no longer necessary. On the other hand, some of those structural issues such as the differing facets and the presence of the bonding wires on the

top contact proved to be relatively important, the former by determining the maximum ΔT from one end of the laser to the other and the latter by helping in the heat removal process.

A more suitable scenario to illustrate the importance of a 3D thermal model arises in integrated optical devices. In this case, heat flow actually occurs in 3 dimensions and the thermal optimization problem becomes more demanding. In this section, a simple semiconductor laser structure in an integrated optical environment is analyzed to emphasize the importance that a 3D thermal model plays under these conditions. Fig. 5.39 shows the structure that is considered here and corresponds to a simple semiconductor laser built on top of a die.

Figure 5.39: Basic semiconductor laser on an OIE.

In order to improve the heat removal process, the laser has a heat metal spreader in a set up which imitates a commonly used technique in BJT devices. The spreader can be seen better in the detailed view of the structure shown in Fig. 5.40. It should be emphasized that the arrangement shown in this section is somewhat arbitrary and is only considered here for illustrative purposes.

Figure 5.40: Detailed view of a laser in an OIE with heat spreader.

In order to show the importance that the location of the heat sources (such as a laser) plays in these scenarios, the previous structure was simulated with the laser placed on different positions of the die. Three different cases were considered; firstly the laser was located in the center of the die and afterwards it was offset in the vertical (y) direction and in the vertical and lateral (xy) directions. Fig. 5.41 shows the resulting temperature

contours for each case. As the figure shows, the heat flow pattern considerably differs from one structure to the other and this will obviously result in different temperature distributions in the die. The distortion of the elliptic contours is due to the presence of the heat spreader and takes particular importance in the vicinity of the heat source where the structural details should not be neglected.

Figure 5.41: Contour plot of the (a) the centered, (b) y -offset and (c) xy -offset structure.

Fig. 5.42 shows the normalized lateral temperature profile for the three cases. As can be seen there, the resulting temperature distribution considerably differs from one case to the other, especially when the laser is offset in both directions. A typical integrated optical device is comprised of several optical components such as modulators, amplifiers, detectors, etc., since some of them are very sensitive to small temperature variations (through the temperature dependence of the various physical parameters, e.g. refractive index), their layout distribution with respect to the optical devices generating heat power

must carefully be assigned. In addition, some applications require that an array of lasers be used. In such cases, it is not unusual to have the various lasers operate in different ON/OFF states and this creates several combinations of heat power distribution in the overall optical device. In these scenarios, a good knowledge of the three dimensional temperature distribution for different combinations of the heat sources is essential to investigate the effects of thermal coupling on the operating conditions of the neighboring devices.

Figure 5.42: Lateral temperature profile for different positions of the laser structure.

Fig. 5.43 shows the normalized temperature distribution for the same three different locations of the laser on the die. From this figure it is apparent that not only the tem-

perature distribution in the close vicinity of the device but also inside the device itself is a function of its position on the die.

Figure 5.43: Longitudinal temperature profile for different positions of the laser structure.

Finally, it should be said here that the temperature rise at a given location does not necessarily constitute a drawback in a whole sense. In some cases, if the temperature dependence of a specific device and its environmental thermal operating conditions are known, thermal effects can be used to the designer's advantage.

Chapter 6

Conclusions

In this thesis a comprehensive thermal analysis of a ridge semiconductor laser structure was performed. The goal was to find a possible solution for the problems stated in Section 1.2. To this end, two approaches were taken; the structure was first analyzed from a strictly thermal standpoint with the aid of *Atar*: a 3D thermal simulator. Next, a thermal/optical/electrical self-consistent model was developed and used to study the same laser structure. The two different approaches were compared to each other, and the necessary conditions for each approximation (model) to be valid, were discussed. Generally speaking, the objectives which gave rise to this thesis, were accomplished. However, this is only true from a theoretical point of view and the results here presented are still subject to experimental validation. With this in mind, the most important results can be summarized as follows.

The use of the pure thermal model led to the following concluding results:

- 1.- Although a pure thermal analysis of a semiconductor laser neglects the interactions among the various optical, electrical and thermal phenomena; this approach is still capable of identifying the right trends toward the laser's thermal optimization.
- 2.- A 17% reduction of the laser's average temperature in the active region was predicted by this model. This reduction would require that the current technique used to mount the laser, be changed from a wire bonded configuration to a flip chip configuration. The modification appears to meet both structural and technological requirements.

- 3.- Although a 3D thermal model is suitable for studying three dimensional structural details in quasi-2D structures such as a stand-alone laser, it finds its major application in optical integrated environments where the three dimensional heat spreading is strong and the thermal optimization problem more demanding.

Similarly, the use of the self-consistent model in analyzing the semiconductor laser led to the following conclusions:

- 4.- When the entire laser structure is studied for optimization, a large scale model (3D) is required and the *longitudinal* variations of the optical and electrical parameters need only be modeled as they occur in the same scale as that of the thermal problem.
- 5.- For a structurally and optically symmetric laser (equal facet reflectivity), a pure thermal approach is sufficient to determine its thermal behavior as the longitudinal spatial temperature variation induced by the non-uniform power distribution is very small, and consequently negligible for most practical purposes.
- 6.- The spatial variation of the longitudinal temperature is more severe in asymmetric lasers than in symmetric ones. For the present structure, a temperature difference of 2.6 K (approx. 17%) from one end of the laser to the other was predicted by the model. The assessment of the importance of this ΔT requires further investigation.
- 7.- Through the investigation of facet heating in a ridge laser, it was shown that when a 3D thermal model is used, the inclusion of Joule heating and non-linear thermal conductivity is indispensable to obtain the high facet temperatures that are required to expect facet degradation.

8.- By comparing an InGaAsP/InP and an AlGaAs/GaAs laser, it was shown that InP lasers exhibit a better thermal behavior at the facets. The combination of lower surface recombination velocity with enhanced heat flow from the facets provided by the higher thermal conductivity of the substrate, and the lower heat power generated ($h\nu$) with each act of non-radiative recombination, results in lower facet temperature as compared with GaAs lasers.

5.1 Future work

The pure thermal and self-consistent approaches described in this work were selected because they offered a possible solution to the original thermal problems to be looked at. Although the models were mainly used to study a specific structure (Ridge Laser), they can be extended to study other semiconductor lasers. Given the wide variety of these optical devices, each with its own optimization problems, the present models could be extended in many different directions. The specific problem to solve, the laser parameters aimed for optimization and/or the required computational effort, would determine the way to proceed in the overall improvement process. As examples, some possible ways to further develop the present model are mentioned below in order of increasing difficulty.

The fact that *Atar* allows the building of any semiconductor structure with relative ease, implies that not only the pure thermal, but also the self-consistent model could straightforward be extended to analyze other semiconductor lasers. This is true because the *Atar* model to use in the pure thermal model, could easily be adapted to fit within the

self-consistent environment. In addition, the model could also be improved by taking into consideration leakage current and other loss mechanisms such as free carrier absorption.

Another obvious way to proceed would be to increase the dimensionality of the electrical and optical equations with the inclusion of lateral effects. This could be done by neglecting longitudinal variations and approximating the laser structure as a two-dimensional device. Should this be the case, the reduction in the thermal domain would compensate for the increase in the computational effort associated with the solution of the 2D optical (see Section 4.2.2) and electrical (Poisson) equations. Under this approximation, longitudinal variations could also be taken into consideration at the expense of a large computational effort [34].

As described in Section 5.5, both the pure thermal and self-consistent models could be extended to handle integrated optical devices. However, achieving a self-consistent solution of the various electro/optical devices involved would be a very challenging task. The resulting computational effort and long simulation times that would be required, would probably make unfeasible the use of a self-consistent approach under this scenario.

Finally, regardless of the approach that is taken to theoretically investigate any semiconductor laser, experimental results are always desired to validate and back up the model's predictions.

References

- [1] Gen-Li Tan, N. Bewtra, K. Lee, J. M. Xu, "A Two-Dimensional Nonisothermal Finite Element Simulation of Laser Diodes", *IEEE Journal of Quantum Electronics*, Vol. 29, No. 3, pp. 822-835, 1993.
- [2] M. Gault, P. Mawby, A. R. Adams and M. Towers, "Two-Dimensional Simulation of Constricted-Mesa InGaAsP/InP Buried-Heterostructure Lasers", *IEEE Journal of Quantum Electronics*, Vol. 30, No. 8, pp. 1691-1700, 1994.
- [3] U. Menzel, R. Puchert, A. Barwolff, A. Lau, "Facet Heating and Axial Temperature Profiles in High-Power GaAlAs/GaAs Laser Diodes", *Microelectronics Reliability*, 38, pp. 821-825, 1998.
- [4] W. Nakwaski, "Thermal model of the catastrophic degradation of high-power stripe-geometry GaAs/(AlGa)As double-heterostructure diode lasers". *Journal of Applied Physics*, Vol. 67, No. 4, pp. 1659-1668, 1990.
- [5] Govind P. Agrawal, Niloy K. Dutta, *Semiconductor Lasers*, Second Ed. New York: Van Nostrand Reinhold, 1993.
- [6] A. Yariv, *Quantum Electronics*. New York: John Wiley & Sons, Inc., 1989
- [7] B. Mroziwicz, M. Bugajski, W. Nakwaski, *Physics of Semiconductor Lasers*. Warszawa, Polish Scientific Publishers, 1991.
- [8] J. Wilson, J. F. B. Hawkes, *Optoelectronics: An Introduction*. Prentice-Hall International, Inc., 1983.
- [9] James A. Morikuni, Sung-Mo Kang, *Computer-Aided Design of Optoelectronic Integrated Circuits and Systems*, Prentice Hall, Inc., 1997.
- [10] L. V. Asryan, N. A. Gun'ko, A. S. Polkovnikov, R. A. Suris, G. G. Zegrya, B. B. Elenkrig, S. Smetona, J. G. Simmons, P. Lau, T. Makino, "High-Power and High-Temperature Operation of InGaAsP/InP Multiple Quantum Well Lasers", *Semiconductor Science and Technology*, 14, pp. 1069-1075, 1999.
- [11] T. Makino, J. D. Evans, G. Mak, "Maximum Output Power and Maximum Operating Temperature of Quantum Well Lasers, *Applied Physics Letters*, Vol. 71, No. 20, pp. 2871-2873, 1997.

- [12] W. Nakwaski, "Thermal Properties of Vertical-Cavity Surface-Emitting Semiconductor Lasers", E. Wolf, Progress in Optics XXXVIII, Elsevier Science B. V., pp.165-262, 1998.
- [13] Carslaw H. S., Jaeger J. C., *Conduction of Heat in Solids*, Clarendon Press, Oxford, 1959.
- [14] W. B. Joyce, R. W. Dixon, "Thermal Resistance of Heterostructure Lasers", Journal of Applied Physics, Vol. 46, No. 2, pp. 855-862, 1975.
- [15] M. Ito, T. Kimura, "Stationary and Transient Thermal Properties of Semiconductor Laser Diodes", IEEE Journal of Quantum Electronics, Vol. QE-17, No. 5, pp. 787-795, 1981.
- [16] T. Smy, D. Walkey, S. K. Dew, "A 3D Thermal Simulation Tool for Integrated Devices - *Atar*", IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems", Vol. 20, No. 1, pp.1-11, 2001.
- [17] M. Reinhard, *Integrated Optics: Desing and Modeling*, Artech House, Inc., 1995.
- [18] M. S. Stern, "Finite Difference Analysis of Planar Optical Waveguides", Progress in Electromagnetics Research, PIER 10, pp. 123-186, 1995.
- [19] J. Buus, "Principles of Semiconductor Laser Modeling", IEE Proceedings, Vol. 132, Pt. J, No. 1, pp. 42-51, 1985.
- [20] A. M. Sarangan, W. P. Huang, G. P. Li, T. Makino, "A Ridge Waveguide DFB Laser Model Including Transverse Carrier and Optical Effects", IEEE Journal of Quantum Electronics, Vol. 32, No. 3, pp. 408-416, 1996.
- [21] P. M. Boers, M. T. Vlaardingerbroek, "Dynamic Behavior of Semiconductor Lasers", Electronics Letters, Vol. 11, pp. 206-208, 1975.
- [22] A. Champagne, R. Maciejko, Technical Report, "A Comparison Between LAS2D Simulations and Experimental Results", The Optoelectronics Laboratory, Ecole Polytechnique de Montreal, 1999.
- [23] S. Nakamura, *Numerical Analysis and Graphic Visualization With Matlab*, Prentice Hall Inc, 1997.
- [24] A. D. Sadovnikov, X. Li, W. P. Huang, "A two-dimensional DFB Laser Model Accounting for Carrier Transport Effects", IEEE Journal of Quantum Electronics, Vol. 31, No. 10, pp. 1856-1862, 1995.

- [25] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Ed. Cambridge University Press, 1999.
- [26] M. Necati Ozisik, *Heat Conduction*, 2nd Ed. John Wiley & Sons, Inc., 1993.
- [27] C. J. Reimer, T. Smy, D. J. Walkey, B. C. Beggs, R. Surridge, "A Simulation Study of IC Layout Effects on Thermal Managment of Die Attached GaAs ICs", IEEE Trans on CMPT, Vol. 23, No. 2, pp. 341-351, 2000.
- [28] W. R. Smith, "Mathematical Modeling of Thermal Runaway in Semiconductor Laser Operation", Journal of Applied Physics, Vol. 87, No. 12, pp. 8276-8285, 2000.
- [29] G. A. Kriegsmann, "Thermal runaway in microwave heated ceramics: A one-dimensional model", Journal of Applied Physics, Vol. 71, No. 4, pp. 1960- 1966, 1992.
- [30] <http://www.smpstech.com/tmos0000.htm>
- [31] H. Temkin, S. Mahajan and M. A. DiGiuseppe, "Optically induced catastrophic degradation in InGaAsP/InP layers", Applied Physics Letters, Vol. 40, No. 7, pp. 562-565, 1982.
- [32] U. Menzel, A. Barwolff, P. Enders, D. Ackermann, R. Puchert, M. Voss, "Modeling the Temperature Dependence of Threshold Current, External Differential Efficiency and Lasing Wavelength in QW Laser Diodes", Semiconductor Science and Technology, 10, pp. 1382-1392, 1995.
- [33] R. Schatz, C. G. Bethea, "Steady State Model for Facet Heating Leading to Thermal Runaway in Semiconductor Lasers", Journal of Applied Physics, Vol. 76, No. 4, pp. 2509-2521, 1994.
- [34] J. Piprek, P. Abraham and J. E. Bowers, "Self-Consistent Analysis of High-Temperature Effects on Strained-Layer Multiquantum-Well InGaAsP-InP Lasers", IEEE Journal of Quantum Electronics, Vol. 36, No.3, pp. 366-374, 2000.

Appendix A

Calculation of the Coefficients of the Analytical Solution of the Heat Equation

The condition of zero heat flux from the top of the device and Eq. 3.18 imply that in the top layer

$$\gamma_{2a,n} = \tanh(k_n t_{2a}), \quad (\text{A.1})$$

similarly the condition of fixed temperature at the heat sink yields in the bottom layer

$$\gamma_{2b-1,n} = \coth(k_n t_{2b-1}). \quad (\text{A.2})$$

At a source-free interface between any two layers the continuity of temperature and normal heat flow relates the coefficient $\gamma_{i,n}$ as follows:

$$\gamma_{i,n} = \frac{\tanh(k_n t_i) + (\kappa_{i-2}/\kappa_i)\gamma_{i-2,n}}{1 + (\kappa_{i-2}/\kappa_i)\tanh(k_n t_i)\gamma_{i-2,n}}. \quad (\text{A.3})$$

Starting with Eq. A.1 in the top layer and applying Eq. A.3 repeatedly yields $\gamma_{2,n}$ in the layer just above the source. Similarly Eqs. A.2 and A.3 yield $\gamma_{1,n}$ below the source.

The remaining γ coefficients are given by

$$\gamma_{2b-1,0} = \frac{1}{t_{2b-1}} \quad (\text{A.4})$$

and

$$\gamma_{i,0} = \frac{\gamma_{i+2,0}}{\frac{\kappa_{i+2}}{\kappa_i} + \gamma_{i+2,0} t_i}. \quad (\text{A.5})$$

With $\gamma_{2,n}$ and $\gamma_{1,n}$ known in the layers adjacent to the source, $\beta_{1,n}$ is found from the fact that the combined flow per unit area from both sides of the source stripe is by assumption constant at each point. From the equality $\beta_{1,n} = \beta_{2,n}$ (that represents the continuity of T through the source plane) and Eq. 3.18 the following equation can be obtained

$$\beta_{1,n} = \frac{4g}{Wk_n^2} \frac{\sin((k_n A)/2)}{\kappa_1 \gamma_{1,n} + \kappa_2 \gamma_{2,n}}, \quad (\text{A.6})$$

the other n coefficients are obtained from

$$\beta_{i-2,n} = \beta_{i,n} (\cosh k_n t_i - \gamma_{i,n} \sinh k_n t_i). \quad (\text{A.7})$$

Finally, the $n = 0$ coefficients are given by

$$\beta_{1,0} = \beta_{2,0} = \frac{Ag}{W(\kappa_1 \gamma_{1,0} + \kappa_2 \gamma_{2,0})} \quad (\text{A.8})$$

and

$$\beta_{i-2,0} = \beta_{i,0} (1 - \gamma_{i,0} t_i). \quad (\text{A.9})$$

Appendix B

Temperature Dependence of the Transparency Carrier Density

The transparency carrier density increases with temperature since the effective number of states, N_c and N_v , increases with temperature while the band gap E_g decreases. In a parabolic band approximation, n_o can be obtained from Fermi-Dirac statistics via the condition that the separation of the quasi-Fermi levels should equal the photon energy, that is,

$$E_{cf}(n_o) + E_{vf}(n_o) + E_g = h\nu \quad (\text{B.1})$$

and the condition of charge neutrality $n = p$.

For simplicity one can assume that the electron temperature is equal to the lattice temperature. We also assume that the photon energy, $h\nu$, is equal to the band gap energy at the reference temperature and that the laser wavelength is not affected by the localized heating of the facet. The distance of the quasi-Fermi levels from the band gap edges, E_{fc} and E_{vf} , are for parabolic bands defined by the relations

$$n = \frac{2N_c}{\sqrt{\pi kT}} \int_0^\infty \frac{\sqrt{E}}{1 + \exp[(E - E_{cf})/kT]} dE. \quad (\text{B.2})$$

$$p = \frac{2N_v}{\sqrt{\pi kT}} \int_0^\infty \frac{\sqrt{E}}{1 + \exp[(E - E_{vf})/kT]} dE. \quad (\text{B.3})$$

However, we need the inverse formulas as we are interested in the quasi-Fermi energies as a function of the carrier density. Therefore the following approximate formula can be used

$$E_{cf} \approx kT \ln\left(\frac{n}{N_c}\right) + kT\left(\frac{n}{N_c}\right) \left\{ 64 + 0.05524\left(\frac{n}{N_c}\right) \times \left[64 + \sqrt{\left(\frac{n}{N_c}\right)} \right] \right\}^{-0.25} \quad (\text{B.4})$$

and corresponding for holes in the valence band. This approximation is valid for $E_{cf} < 15kT$.

The transparency carrier density as a function of temperature for the GaAs and InGaAsP lasers can with good accuracy be approximated as

$$n_o = n_{tr,0} + n_{tr,T} \left(\frac{T_\Delta}{T_0} \right)^q \quad (\text{B.5})$$

with $n_{tr,0} = 1.16 \times 10^{18} \text{ cm}^{-3}$, $n_{tr,T} = 6.50 \times 10^{18} \text{ cm}^{-3}$ and $q = 1.37$ for the GaAs laser and $n_{tr,0} = 0.81 \times 10^{18} \text{ cm}^{-3}$, $n_{tr,T} = 3.18 \times 10^{18} \text{ cm}^{-3}$ and $q = 1.32$ for the InGaAsP laser.

Appendix C

Solution of the Diffusion Electrical Equation

Starting from the diffusion Eq. C.1

$$0 = D \frac{d^2 n}{dz^2} + \frac{J}{ed} - \frac{c}{\eta_g} [\Gamma g(n, T)] [I^f + I^b] - \frac{n}{\tau_c(n, T)} \quad (C.1)$$

where the gain is given by

$$g(n, T) = a \left(n - n_0 \times \exp\left(-\frac{E_{int}}{k_B T}\right) \right) \quad (C.2)$$

and the spontaneous carrier lifetime by

$$\tau_c(n) = \left(A_{nr} \times \exp\left(-\frac{E_1}{k_B T}\right) + Bn \times \exp\left(-\frac{E_2}{k_B T}\right) + Cn^2 \times \exp\left(-\frac{E_3}{k_B T}\right) \right)^{-1} \quad (C.3)$$

to shrink notation the following definitions can be made:

$$n_o(T) = n_0 \times \exp\left(-\frac{E_{int}}{k_B T}\right), S = I^f + I^b, A = \frac{c}{\eta_g} a \quad (C.4)$$

and

$$A_{nr}(T) = A_{nr} \times \exp\left(-\frac{E_1}{k_B T}\right), B(T) = B \times \exp\left(-\frac{E_2}{k_B T}\right), C(T) = C \times \exp\left(-\frac{E_3}{k_B T}\right) \quad (C.5)$$

by substituting Eq. C.2 into C.1 we can rewrite it as

$$0 = D \frac{d^2 n}{dz^2} + \left(\frac{J}{ed} + S \Gamma a n_o(T) \right) - \left(\frac{1}{\tau_c} + \Gamma A S \right) n \quad (C.6)$$

defining the source term

$$source = \frac{J}{edD} + \frac{S \Gamma a n_o(T)}{D} \quad (C.7)$$

and substituting C.7 into C.6 we have

$$0 = \frac{d^2 n}{dz^2} + source - \left(\frac{A_{nr}(T)}{D} + \frac{\Gamma AS}{D} \right) n - \frac{B(T)}{D} n^2 - \frac{C(T)}{D} n^3 \quad (C.8)$$

by further defining

$$K_1 = \left(\frac{A_{nr}(T)}{D} + \frac{\Gamma AS}{D} \right), B_D = \frac{B(T)}{D}, C_D = \frac{C(T)}{D} \quad (C.9)$$

Eq. C.8 becomes

$$0 = \frac{d^2 n}{dz^2} + source - K_1 n - B_D n^2 - C_D n^3 \quad (C.10)$$

if we denote an approximation of $n(z)$ by $N(z)$, the exact solution can be expressed as

$$n(z) = N(z) + \delta N(z) \quad (C.11)$$

by substituting C.11 into C.10 we obtain,

$$0 = \frac{d^2}{dz^2} [N(z) + \delta N(z)] + source - K_1 [N(z) + \delta N(z)] - B_D [N(z) + \delta N(z)]^2 - C_D [N(z) + \delta N(z)]^3 \quad (C.12)$$

since $\delta N(z)$ is small, second order terms can be dropped out and the square polynomial can be approximated by

$$[N(z) + \delta N(z)]^2 = N(z)^2 + 2N(z)\delta N(z) \quad (C.13)$$

similarly, keeping only the first order terms the cubic polynomial can be approximated by

$$[N(z) + \delta N(z)]^3 = N(z)^3 + 3N(z)^2\delta N(z) \quad (C.14)$$

and Eq. C.12 becomes

$$0 = \frac{d^2}{dz^2}[N(z) + \delta N(z)] + source - K_1[N(z) + \delta N(z)] - B_D[N(z)^2 + 2N(z)\delta N(z)] - C_D[N(z)^3 + 3N(z)^2\delta N(z)] \quad (C.15)$$

grouping terms for $N(z)$ and $\delta N(z)$ we have

$$\begin{aligned} & -\frac{d^2}{dz^2}\delta N(z) + K_1\delta N(z) + 2B_DN(z)\delta N(z) + 3C_DN(z)^2\delta N(z) = \\ & \frac{d^2}{dz^2}N(z) + source - K_1N(z) - B_DN(z)^2 - C_DN(z)^3 \end{aligned} \quad (C.16)$$

using the following finite differences approximation for the second order derivatives at any discrete node i

$$\frac{d^2}{dz^2}\Phi(z) = \frac{\Phi_{i-1} - 2\Phi_i + \Phi_{i+1}}{h^2} \quad (C.17)$$

Eq. C.16 becomes

$$\begin{aligned} & \frac{-\delta N_{i-1} + 2\delta N_i - \delta N_{i+1}}{h^2} + K_1\delta N_i + 2B_DN_i\delta N_i + 3C_DN_i^2\delta N_i = \\ & \frac{N_{i-1} - 2N_i + N_{i+1}}{h^2} + source - K_1N_i - B_DN_i^2 - C_DN_i^3 \end{aligned} \quad (C.18)$$

to simplify notation once more, we can group terms in the following way

$$BB = 2 + h^2K_1 + 2h^2B_DN_i + 3h^2C_DN_i^2 \quad (C.19)$$

and

$$DD = 2 + h^2K_1 - h^2B_DN_i + h^2C_DN_i^2 \quad (C.20)$$

in this way Eq. C.18 reduces to

$$-\delta N_{i-1} + BB\delta N_i - \delta N_{i+1} = N_{i-1} - DDN_i + N_{i+1} + h^2source \quad (C.21)$$

Eq. C.21 can be written in a tridiagonal matrix form

$$\begin{bmatrix} B_1 & C_1 & 0 & 0 & 0 & 0 & 0 \\ A_2 & B_2 & C_2 & 0 & 0 & 0 & 0 \\ 0 & A_3 & B_3 & C_3 & 0 & 0 & 0 \\ 0 & 0 & \dots & \dots & \dots & 0 & 0 \\ 0 & 0 & 0 & A_i & B_i & C_i & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & A_{N_n} & B_{N_n} \end{bmatrix} \begin{bmatrix} \delta N_1 \\ \delta N_2 \\ \delta N_3 \\ \dots \\ \delta N_i \\ \dots \\ \delta N_{N_n} \end{bmatrix} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ \dots \\ D_i \\ \dots \\ D_{N_n} \end{bmatrix} \quad (C.22)$$

where N_n is the number of nodes and the entries A_i , B_i and D_i are given by

$$A_i = -1, B_i = BB, C_i = -1 \quad (C.23)$$

and

$$D_i = N_{i-1} - DDN_i + N_{i+1} + h^2 source. \quad (C.24)$$

The solution of C.22 can easily be obtained [23],[25].

Analysis of the Boundary Conditions

The diffusion equation must in general be solved subject to the boundary conditions of zero flow of carriers at the edges of the simulated domain (laser facets). Nevertheless, when the extra heating effect at the facets wants to be investigated, the facets behave like a sink of carriers and the following boundary conditions must be applied

$$-D \frac{d}{dz} n(0) = -V_{sur} n(0), \quad -D \frac{d}{dz} n(L) = V_{sur} n(L) \quad (C.25)$$

where D is the diffusion constant, L is the laser length and V_{sur} is the surface recombination velocity at which carriers are sunk. To discretize Eq. C.25 the following finite differences approximation was used

$$\frac{d\Phi_1}{dz} = \frac{\Phi_2 - \Phi_0}{2h} \quad (C.26)$$

where for convenience a fictitious node ($i = 0$) was introduced as shown in Fig. C.1

Figure C.1: One dimensional discretization of the electrical domain.

In this way the boundary condition at $z = 0$ becomes

$$D \frac{n_2 - n_0}{2h} = V_{sur} n_1 \quad (C.27)$$

and

$$n_0 = n_2 - K_2 n_1 \quad (C.28)$$

where

$$K_2 = \frac{2hV_{sur}}{D} \quad (C.29)$$

By discretizing equation C.21 at node 1 ($z = 0$) and replacing n_0 by C.28 the following entries for Eq. C.22 are obtained

$$A_1 = 0, B_1 = BB + K_2, C_1 = -2 \quad (C.30)$$

and

$$D_1 = 2N_2 - DDN_1 - K_2N_1 + h^2 source. \quad (C.31)$$

By carrying out the same procedure at node N_n ($z = L$) the following entries are obtained

$$A_{N_n} = -2, B_{N_n} = BB + K_2, C_{N_n} = 0 \quad (C.32)$$

and

$$D_{N_n} = 2N_{N_n-1} - DDN_{N_n} - K_2N_{N_n} + h^2 source. \quad (C.33)$$

By solving iteratively Eq. C.22 and updating C.11 at each loop, the solution of C.1 is obtained. The iteration is started by setting the carrier density to its threshold value. Convergence is achieved when the solution of C.22 (δN_i) is small (*about* $\sim 1 \times 10^{10}$) and 10 to 15 iterations are usually enough.

Appendix D

Solution of the Non-Diffusion Electrical Equation

Starting from the non-diffusion Eq. D.1

$$0 = \frac{J}{ed} - \frac{c}{\eta_g} [\Gamma g(n, T) [I^f + I^b]] \quad (D.1)$$

$$- n \left[A_{nr} \times \exp\left(-\frac{E_1}{k_B T}\right) + Bn \times \exp\left(-\frac{E_2}{k_B T}\right) + Cn^2 \times \exp\left(-\frac{E_3}{k_B T}\right) \right]$$

and following a procedure similar to that in Appendix C, this equation can be written as

$$0 = source - K_1 n - B(T)n^2 - C(T)n^3 \quad (D.2)$$

with

$$source = \frac{J}{ed} + S\Gamma an_o, \quad K_1 = (A_{nr}(T) + \Gamma AS) \quad (D.3)$$

and

$$A_{nr}(T) = A_{nr} \times \exp\left(-\frac{E_1}{k_B T}\right), \quad C(T) = C \times \exp\left(-\frac{E_3}{k_B T}\right),$$

$$B(T) = B \times \exp\left(-\frac{E_2}{k_B T}\right) \quad (D.4)$$

the other parameters are as in Appendix C.

Eq. D.2 is a cubic polynomial in n and can be solved analytically as follows. Doing a change of notation for simplicity we can rewrite D.2 as

$$ax^3 + bx^2 + cx + d = 0 \quad (D.5)$$

which reduces to

$$y^3 + py + q = 0 \quad (\text{D.6})$$

where

$$y = \frac{b}{3a} + x, \quad p = \frac{c}{a} - \frac{b^2}{3a^2}, \quad q = \frac{2}{27}\left(\frac{b}{a}\right)^3 - \frac{bc}{3a^2} + \frac{d}{a} \quad (\text{D.7})$$

solution:

$$y_1 = u + v, \quad y_2 = -\frac{1}{2}(u + v) + i\frac{1}{2}(u - v)\sqrt{3}, \quad y_3 = -\frac{1}{2}(u + v) - i\frac{1}{2}(u - v)\sqrt{3} \quad (\text{D.8})$$

where

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}, \quad v = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \quad (\text{D.9})$$

if

$$\frac{q^2}{4} + \frac{p^3}{27} > 0 \quad \text{one real solution}$$

$$\frac{q^2}{4} + \frac{p^3}{27} = 0 \quad \text{one real solution for } p = q = 0$$

$$\text{two real solutions for } 27p^3 = -4q^2$$

$$\frac{q^2}{4} + \frac{p^3}{27} < 0 \quad \text{three real solutions}$$

in the case $\frac{q^2}{4} + \frac{p^3}{27} < 0$ the following method for solution may be used:

$$y_1 = -2r \cos \frac{\phi}{3}, \quad y_2 = 2r \cos \left(60^\circ - \frac{\phi}{3} \right), \quad y_3 = 2r \cos \left(60^\circ + \frac{\phi}{3} \right) \quad (\text{D.10})$$

where

$$\cos \phi = \frac{q}{2r^3} \quad \text{and} \quad r = \pm \sqrt{\frac{p}{3}} \quad (\text{D.11})$$

the sign is that of q .