National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

NOVEL DEVICES FOR HYBRID AND MONOLITHIC OPTOELECTRONIC INTEGRATION

BY

CHRISTOPHER FRANK JANZ ©

A thesis submitted to the faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

DEPARTMENT OF ELECTRICAL ENGINEERING

Edmonton, Alberta
Fall, 1995

National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-612-06230-9

Canada

For my Grandparents:

Frank & Anna Janz

Clifford & Helen Waller

Abstract

This thesis is concerned with the development of optical and optoelectronic components which facilitate hybrid and monolithic device integration. The work described addresses two principal issues: compactness, and multi-functionality.

Compactness is a critical requirement for integrated photonic or optoelectronic circuits since many such circuits are implemented in expensive multi-layer semiconductor heterostructures. One factor inhibiting compact photonic circuit realization is the requirement to use large radius waveguide bends to avoid power losses due to radiation. Deeply etched rib waveguides can be sharply bent without suffering such losses; in the first part of this work, the implementation of multi-mode couplers in such waveguides is studied. Couplers are designed for use in a Mach-Zehnder switch, fabricated in an indium phosphide multi-quantum well semiconductor structure. High device performance is achieved using multi-mode couplers which are among the smallest ever reported. A novel mechanism for optimizing multi-mode couplers configured as optical (de)multiplexers, using circular waveguide bends, is also discussed. Devices based on this principle are designed and implemented in a silicon oxynitride glass waveguide system which is compatible with rare-earth doped materials for optical amplifier or laser implementation. (De)multiplexing performance superior to that achieved by other methods is demonstrated with a very compact device.

The second major concern addressed in this work is multi-functionality of optoelectronic components in integrated circuits. Normally, a semiconductor structure cannot operate as both a light source and a modulator at one wavelength. In this work, a multi-quantum well heterostructure ("LAMDA" - for LAser/Modulator/Detector/Amplifier) which overcomes this problem is designed and fabricated. Two quantum well stacks, with different well widths, are used to create active regions with different characteristic energies. Since guided optical fields overlap with both active regions, which can be separately biased, light generation and absorption can be simultaneously achieved at one wavelength. Computer models which were used to predict the optical and electrical characteristics of the device are described, and experimental results demonstrating operation of the device as a single cavity laser / electroabsorption modulator are presented.

Finally, a theoretical consideration of spontaneous emission in the vicinity of a planar dielectric waveguide is discussed. This work is an extension of some of the quantum mechanical theory used in the design of LAMDA, and is used to examine the spontaneous emission characteristics of optically pumped rare-earth doped planar waveguide amplifiers.

Acknowledgements

I would like to express my most sincere thanks to my supervisors, Dr. Barrie Keyworth and Dr. R. Ian MacDonald, for facilitating, guiding and encouraging this work throughout its duration. I would also like to thank Dr. James McMullin for his constant tutelage and steady willingness to discuss scientific matters of all kinds. These gentlemen have been true mentors for almost four years.

I extend my thanks to all the staff and students, past and present, of Telecommunications Research Laboratories, who have not only provided a pleasant and intellectually charged working environment, but have also furnished enormous technical assistance as well as moral support and encouragement over the years. My special acknowledgements in this regard go out to David Clegg, Dino Corazza, Rainer Iraschko, Dobby Lam, Greg May, Tim Neufeld, Reza Paiam, Terrance Rosadiuk and Rohit Sharma.

I thank the staff of the Alberta Microelectronic Centre, in particular Dr. James Broughton, Glen Fitzpatrick, Graham McKinnon and Tran Tran for their considerable technical assistance and guidance. I also thank the staff and professors of the Department of Electrical Engineering, University of Alberta, notably Dr. Ron Lawson, for their help, encouragement and instruction. Special thanks go to Dr. Walter Allegretto, Department of Mathematics, University of Alberta, for his collaboration on aspects of this work.

I gratefully acknowledge the assistance and encouragement of Dr. Claude Rolland, Dr. Tony SpringThorpe, Glen Hillier and other members of the Advanced Optical Devices Group at Bell-Northern Research Ltd., Ottawa, and that of Dr. Francoise Chatenoud, Dr. Andre Delage, Dr. Mahmoud Fallahi and others at the Institute for Microstructural Sciences, National Research Council of Canada, Ottawa. Without their collaboration much of this work would not have been possible.

I would like to thank the Natural Sciences and Engineering Research Council of Canada, Telecommunications Research Laboratories and the University of Alberta for their generous financial support.

Finally, I thank my parents and family for their patience, support and encouragement over the years. Their devotion has made this work, as all my endeavours, possible.

Table of Contents

Chapter Page

List of Tables

List of Figures

Chapter **Page**

List of Symbols, Nomenclature and Abbreviations

a	mode amplitude coefficient (complex)
a', a	creation, annihilation coefficients
A	phase constant (in certain instances); amplitude coefficient (in certain instances)
A, B, C	recombination rate coefficients (in certain instances)
a, b, c, F, G, K	arbitrary or dummy variables or coefficients (in certain instances)
$[A]$, $[B]$	matrices, defined in text
b, B	time-dependent, time-independent magnetic field (complex)
c	speed of light
$[C]$	connectivity matrix
d, D	time-dependent, time-independent electric displacement (complex)
D	optical mode density
$e, E, \mathcal{E}$	time-dependent, time-independent electric field (complex), time-independent electric field (real)
e	electronic charge
E	energy
E_{fc}, E_{fv}	conduction band, valence band quasi-Fermi levels
f_{co}, f_c	conduction band Fermi-Dirac function (with respect to photon, electron energies respectively)
f_{vo}	valence band Fermi-Dirac function
F	field variable
g	bulk gain
g_m	mode gain
G	steady-state carrier recombination rate
$h, H, \mathcal{H}$	time-dependent, time-independent magnetic intensity (complex), time-independent electric field (real)
h	discretization interval (in some instances);

	total lateral displacement of waveguide s-bends
H	unitary Heaviside function
i	current
j	unit imaginary number (i.e. $\sqrt{-1}$); time-dependent current density (vector)
$\hbar$	Planck's constant
k	wave number; Boltzmann's constant
K	transverse propagation constant
l	length
L	length
L_π	beat length
L_π^2	beat length, defined between fundamental and second-order modes
m	mass
M, N	integers with significance with respect to resonant self-imaging in multi-mode waveguides
M	momentum matrix
n	index of refraction; electron density
N	mode normalization factor; transformed refractive index profile of a bent waveguide
p, q, r	Sturm-Liouville equation variables
p	photon number operator
P	power; number of photons; mode type index
q	quantity determined by slab waveguide parameters, eqn. 6.33
Q	transverse propagation constant
r	radius; position operator
r_{12}	dipole magnitude
R	radial variable; reflectivity
r^{sp}, R_{sp}	spontaneous emission rate
R_y	Rydberg constant
s, S	time-dependent, time-independent Poynting vector (complex); sum parameter in certain instances
sep	separation between coupler access waveguides
t	time

T	temperature
u	a field function
v_g^K	group velocity of mode K
V	volume; potential
w	gaussian field diameter
W	multi-mode waveguide width; emission rate
x,y,z	Cartesian spatial coordinates

Subscripts

a	denotes absorption layer
A	denotes area
avg	denotes average
b	denotes a heterostructure barrier
ba	denotes barrier, e.g. electron density in barriers
bar	denotes barrier potential
$bound$	denotes bound modes
c	denotes conduction band; continuum edge
cav	denotes optical cavity
cb	denotes conduction band
cc	denotes a cross-coupler
cl	denotes cladding
cv	denotes valence band
C	denotes a contour
$disjoint$	denotes finite element method quantities before reduction
e	denotes effective index; electron
eba	denotes electrons in barriers
ex	denotes exciton spectral resonance
g	denotes band gap energy; gain layer
h	denotes heavy hole; valence band
i	denotes a discrete counting index
if	denotes input field
in	denotes incident field
l	denotes an optical mode; light hole

$joint$	denotes finite element method quantities after reduction
jb	denotes exciton binding energy
k, m, n, p	denotes mode number, discrete variable
min	denotes minimum
max	denotes maximum
m, p	denotes mode number, continuous variable
N	denotes sub-band quantum number
o	denotes value in free space or in incident field;
	denotes electron rest mass
of	denotes output field
I	denotes the imaginary component of a complex variable
QW	denotes quantum well
r	denotes residual
rad	denotes radiation modes
red	denotes reduced function, e.g. density of states
R	denotes the real component of a complex variable
s	denotes contour component
sp	denotes spontaneous emission
t	denotes transverse component
u, v	denote stoichiometric composition in InGaAsP
un	denotes unconfined electrons
w	denotes a heterostructure well
$well$	denotes quantum well-confined electrons
x,y,z	denote Cartesian spatial components

Superscripts

$\bar{\ }$	denotes a vector quantity
$\hat{\ }$	denotes a vector of unit magnitude
$\overline{\ \ }$	denotes a normalized quantity
$'$	denotes evaluation at a specific point; derivative
cb	denotes conduction band
cv	denotes valence band
e	denotes effective index
j	denotes hole type

N	denotes sub-band quantum number

Greek Letters

α	shape function; absorption coefficient; spontaneous emission factor
β	propagation constant
χ, χ_{so}	doping distribution function; parameters in model for refractive index of InGaAsP
δ	phase constant (in certain instances)
Δ_{so}	parameter in model for refractive index of InGaAsP
ε	electric permittivity; energy (in some instances)
ϕ	angular variable; field envelope function
φ	angular variable
Γ	line width; optical intensity confinement factor
λ	wavelength; Sturm-Liouville equation parameter
μ	magnetic permeability
θ	angular variable
ρ	charge density; radius of curvature; density of states
υ	(subscript) mode number, discrete variable
τ	lifetime
ω	angular frequency
Ω	functional variable
ζ	matrix component, defined in text; dipole distribution function
ξ	radial variable
ψ, Ψ	scalar field or wave function (complex)

Abbreviations

ASE	Amplified Spontaneous Emission
ATM	Asynchronous Transfer Mode
CW	Continuous Wave
DFB	Distributed Feedback

DH	Double Heterostructure
HWHM	Half-Width at Half-Maximum
LAMDA	LAser/Modulator/Detector/Amplifier Structure
MBE	Molecular Beam Epitaxy
MOCVD	Metal-Organic Chemical Vapour Deposition
MQW	Multiple Quantum Well
PECVD	Plasma Enhanced Chemical Vapour Deposition
PIC	Photonic Integrated Circuit
QW	Quantum Well
RC	Resistance times Capacitance
RF	Radio Frequency
SONET	Synchronous Optical Network
S-R-H	Shockley-Read-Hall
TDM	Time Division Multiplexing
TE	Transverse Electric
TEM	Transverse Electric and Magnetic
TM	Transverse Magnetic
WDM	Wavelength Division Multiplexing

Chapter 1

Introduction

Over the last twenty-five years, the application of photonics has revolutionized telecommunications. The development of semiconductor lasers and photodetectors, capable of high-speed operation, followed by the invention of low-attenuation glass fiber, made long-distance transmission using an optical frequency carrier possible for the first time. In the 1980's, long-haul fiber links became a ubiquitous feature of networks in and between developed countries, raising the possibility of all-optical networks in the future, including "fiber-to-the-home." It is envisaged that such networks, exploiting the enormous bandwidth potential of optical carriers, will enable the deployment of a wide variety of high bit-rate services, including video and high-rate data distribution.

The successful implementation of such services must await the surmounting of a number of obstacles. The current network is not "optically transparent;" that is, an optically borne data stream is converted to electronic form, and re-converted back into optical form, a number of times between the path end-points. These electronic "bottlenecks," which limit the system bit-rate to the speed of the electronics in the link, are required for two reasons. First, repeaters must periodically correct for attenuation and dispersion of the fiber-borne optical signal; second, network functions such as framing, multiplexing and switching require digital logic elements that are currently available only in electronic form. Dramatic increases in the capacity of fiber-based networks will require that some, or all, of the functions listed above be accomplished in the optical domain.

Some progress to this end has been achieved in recent years. For example, the new rare-earth doped fiber amplifiers, which boost an optical signal stream directly, eliminate the need for electronic signal amplification. The impact of chromatic dispersion can be minimized by using narrow linewidth distributed feedback lasers, dispersion-shifted fiber and multi-level coding schemes with reduced spectral content. These measures have greatly increased the distance-bit rate product which can be achieved without electronic repeaters. Dispersion, normally a limiting factor, can actually be turned to advantage if sufficient optical power is available. Stable (non-spreading) propagating waves called solitons can be generated by balancing the chromatic group velocity dispersion of a bit pulse against the phase shift induced by the intensity-dependent part of the index of refraction, apparent in silica fibers at powers on the order of milliwatts. Single-channel, repeaterless soliton transmission at 10 Gbit/s over more than 20000 km has been demonstrated using erbium-doped fiber amplifiers and sliding frequency optical bandpass filters to restrict amplified spontaneous emission noise accumulation [1], while 20 Gbit/s over 125000 km has been achieved using active timing jitter suppression [2]. Future main trunks, i.e. trans-oceanic and trans-continental links, are likely to employ soliton transmission. However, the great sensitivities of soliton stability to the characteristics of the source laser are likely to prevent the deployment of soliton links at other network levels for a considerable time.

A number of the impediments which limit current networks have thus been satisfactorily addressed. However, complex processing operations such framing and switching are likely to remain in the electronic domain for the foreseeable future, since practical optical logic circuits are a long way from development. The immediate route to increased network optical transparency (and increased overall data capacity) lies with restricting these electronic operations to the lower levels (i.e. lower data rate portions) of the network. The practicality of such a migration of processing requirements to lower network levels is enhanced by new network protocols, such as the Synchronous Optical Network (SONET), which supports Asynchronous Transfer Mode (ATM) data transport, coupled with improved network synchronization technology. However, the use of optical or optoelectronic multiplexing and demultiplexing techniques is required to support this transition. If optical carriers modulated at rates compatible with electronics can be optically multiplexed, either in the wavelength or the time domain, then the electronic bandwidth bottlenecks can be moved from the main fiber trunks down to the subscriber level. It is for this reason that optical multiplexing techniques are widely seen as the most important focus of contemporary photonics research.

Wavelength and time division optical multiplexing (WDM and TDM, respectively) require a range of passive and active guided wave components. Lasers, modulators, polarization splitters/converters, waveguides, diffractive elements, optical amplifiers, optical switches and photodetectors may all be needed to realize a WDM or TDM transmitter or receiver system. While these components are all commercially available, practical and reproducible system fabrication and packaging requires the realization of such systems in integrated form. The development of photonic integrated circuits (PIC's) is, in many respects, the "second generation" of optical telecommunications technology evolution.

Two approaches to PIC fabrication are possible. The first is the hybrid approach, wherein discrete optoelectronic devices, such as lasers and photodetectors, are mounted on a planar substrate and connected by optical waveguides. With this approach, the individual devices can be of different semiconductor construction. Hybrid integration is currently seen as the most promising means of realizing PIC's in the immediate future. Its inherent advantage is that individual optoelectronic devices, already fabricated and separately optimized for their respective functions, are employed. Glass or polymer waveguides then provide a convenient means of providing optical connections.

Hybrid PIC's have a number of drawbacks, however. One is size: the discrete devices must be totally self-contained, including packaging and biasing terminals, and some space must be left between the devices on the substrate to allow sufficient room to place and bond the devices. Furthermore, it is difficult to make low-loss component-to-waveguide connections. Monolithic integration, wherein all components and optical waveguide interconnections are fabricated by simultaneous processing of a semiconductor substrate, alleviates all of these problems. High component densities and low-loss interconnections are possible. However, monolithic PIC's require that optoelectronic components with different functions be fabricated in the same semiconductor structure. This is problematic, since different functions require different local semiconductor energy characteristics. For example, a laser must emit light at a certain wavelength, while a waveguide must be transparent to light at that same wavelength. Monolithic photonic circuit integration therefore requires some mechanism for spatially selective band-gap energy tailoring, such as patterned epitaxy and selective regrowth processes, which lead to slow, expensive, and low-yield fabrication. For these reasons, monolithic integration is generally seen as being farther from practical applicability than hybrid integration, although some considerable successes have been reported [3]-[5].

In this work, two issues related to hybrid and monolithic photonic integration are considered. One is multi-functional optoelectronic integration, as discussed above; the other is the promotion of compactness by using deeply etched rib waveguides. Photonic circuit compactness is important in both hybrid and monolithic PIC realization. In hybrid PIC's, the complexity of integrable circuit functions is limited by the density of components that can be mounted on a single substrate. Monolithic PIC's are implemented on expensive semiconductor substrates and therefore the size of an integrated photonic circuit has an important effect on cost. A principal impediment to compact PIC implementation is the requirement to use large radius bends in waveguide interconnections, typically millimeters or centimeters, to avoid power losses due to radiation. Much smaller radius bends can be employed if the guided optical fields are presented with large lateral refractive index differences which serve to strongly confine the fields; deeply etched air-clad rib waveguides meet this condition. Considerable attention has been focused on such waveguides in recent years, and low-loss bends as small as 30 μm radius have been reported [6]. Such dimensions facilitate an integrated component density several orders of magnitude better than is typically available. However, the implementation of guided wave components in such waveguides is difficult, since the propagation characteristics of the optical fields are highly sensitive to the physical parameters, e.g. width and depth, of the waveguides.

In Chapter 2, we review the aspects of waveguide theory which are applied in this work. In particular, methods of analyzing the optical fields in deeply etched waveguides are justified and developed. In Chapter 3, we apply this theory to a study of multi-mode waveguide couplers. Such couplers, which are important components of many passive and active guided wave devices, are examined in the domain of strong guiding using a modal analysis. In particular, the effects of using angled waveguide inputs and outputs of multi-mode couplers are studied. Angled access waveguides facilitate the repeatable fabrication of compact multi-mode devices, but their use is expected to degrade, or at least modify, coupler performance, and their impact has not been previously studied in quantitative terms. A 3-dB coupler for use in a Mach-Zehnder switch, implemented in InP for operation at 1550 nm wavelength, is designed. The fabricated device exhibits almost 20 dB of contrast between outputs, using a 3 dB coupler that is among the smallest ever reported (65 μm).

In Chapter 4, a novel technique for implementing wavelength (de)multiplexers in deeply etched waveguides is presented. This technique uses waveguide bends to optimize the performance of multi-mode waveguide couplers, configured as (de)multiplexers, for arbitrary pairs of wavelengths. Such devices are required, for example, to combine and separate pump and signal light in integrated rare-earth doped amplifiers and ring resonator lasers. Analysis techniques for bent waveguides are developed, and a device is designed for implementation in silicon oxynitride glass, which is compatible with erbium-doped glass systems. Modeled and experimental results are compared. Almost 20 dB and 14 dB contrast between output waveguides are obtained in the demultiplexer configuration at 1550 nm and 850 nm, respectively, with a coupler only 145 μm long. This performance compares favorably to that of other integrated (de)multiplexer devices, which are often physically much larger.

In Chapter 5, the issue of multi-functionality in monolithic semiconductor PIC's is addressed. We demonstrate that selective-area epitaxy and regrowth can be avoided by including two active layers, with different energy properties, in the vertical semiconductor structure profile. This is accomplished by using GaAs/AlGaAs quantum wells, with different well widths to offset the excitonic resonance energies. Therefore, the absorption spectrum of one set of quantum wells is normally transparent at the photon emission wavelength of the other. Both quantum well stacks overlap with the guided optical fields, and can be separately biased, hence independently controlled light generation and absorption, via the Quantum Confined Stark Effect, can be simultaneously achieved. This structure is called LAMDA (for LAser/Modulator/Detector/Amplifer) and is unique in that local functionality is determined exclusively by electrical biasing. Models used to predict the electrical and optical characteristics of the structure are presented, along with experimental results that demonstrate the performance of a single-cavity laser/electroabsorption modulator.

The theoretical models used in the design of LAMDA led to an interest in the influence of waveguide structures on spontaneous emission, and the capture of spontaneously emitted radiation by bound waveguide modes. These phenomena in fact have an important influence on the performance of optical amplifiers and lasers. In Chapter 6, a theoretical consideration of spontaneous emission in the vicinity of planar dielectric waveguides is presented. Exact analytical expressions for the complete set of bound and radiating waveguide modes are used in quantum mechanical expressions to directly yield the rates of spontaneous emission into any individual mode, or group of modes. This work is applied to an analysis of spontaneous emission capture in planar rare-earth doped waveguide amplifiers, a subject of considerable current interest.

The LAMDA device was conceived by Dr. B. P. Keyworth, and fabricated at the Institute for Microstructural Sciences, National Research Council of Canada, Ottawa. The Mach-Zehnder switch work was directed by the author, with useful input from Dr. Keyworth, Dr. R. I. MacDonald and Dr. Claude Rolland. Fabrication of the Mach-Zehnder devices took place at Bell-Northern Research Limited, Ottawa, and at the Institute for Microstructural Sciences. The bent multi-mode waveguide (de)multiplexer was conceived by the author and fabricated at the Alberta

Microelectronic Centre, Edmonton. The work on spontaneous emission was a joint effort by the author and Dr. J. N. McMullin.

Chapter 2

Theoretical Aspects of Optical Waveguides

As is the case for all classical electromagnetic phenomena, optical field propagation in waveguides is completely described by the Maxwell equations. However, deriving a clear conceptual picture and useful quantitative estimations from those equations is by no means straightforward. Indeed, approximate and numerical methods of modeling waveguide behaviour are the subject of considerable current research. In this chapter, the theoretical aspects of optical waveguides of importance in the following work are presented. In addition, the implementation of some numerical models which were used in the work is described.

In general, waveguides are structures that are able to confine propagating light, thereby conducting it from one point to another. Their basic principle of operation is total internal reflection, whereby light in one medium is totally reflected at an interface with a medium of lower refractive index. If a high index medium is surrounded on all sides by a lower index medium, then the high index medium is capable of guiding light, since light coupled into it, within a certain off-axis angle, will be totally internally reflected as it propagates. All of the devices considered in this work use waveguides with dimensions on the order of wavelengths to make use of field quantization effects; therefore, an electromagnetic wave analysis is required. The starting point for such an analysis is the Maxwell equations.

The theory presented in sections 2.1-2.7 is well-established; [7]-[10] are good general references for this material.

2.1 Derivation of the Vector Helmholtz Equations In point form, Maxwell's equations are:

$$\nabla \times \bar{e} = -\frac{\partial \bar{b}}{\partial t} \tag{2.1.a}$$

$$\nabla \times \bar{h} = \frac{\partial \bar{d}}{\partial t} + \bar{j} \tag{2.1.b}$$

$$\nabla \bullet \bar{d} = \rho \tag{2.1.c}$$

$$\nabla \bullet \bar{b} = 0 \qquad \textbf{(2.1.d)}$$

where $\bar{e}$ is electric field, $\bar{b}$ is magnetic field, $\bar{h}$ is magnetic intensity, $\bar{d}$ is electric displacement, $\bar{j}$ is current density, and ρ is charge density. The lower case symbols denote time-dependent quantities. We eliminate the time dependence in (2.1) by assuming the time dependence to be completely described by:

$$\bar{e}(x,y,z,t) = \bar{E}(x,y,z)e^{-j\omega t} \qquad \textbf{(2.2)}$$

etc. Upper case symbols are used to denote time-independent quantities. The characterization given in (2.2) describes the field disturbances which give rise to monochromatic radiation of radial frequency ω; any time variation can be represented by a spectrum of similar functions by means of a Fourier series or integral. In any case, (2.2) is a good approximation for optical frequency waves modulated at information frequencies of up to hundreds of GHz. We next assume that the materials of interest are free of current and unbound charge:

$$\bar{j} = \rho = 0 \qquad \textbf{(2.3)}$$

and that they are non-magnetic and have a linear and isotropic electric susceptibility:

$$\bar{B} = \mu_o \bar{H} \qquad \textbf{(2.4.a)}$$

$$\bar{D} = \varepsilon \bar{E} \qquad \textbf{(2.4.b)}$$

where μ_o is the magnetic permeability of free space and ε is the (scalar) electric permittivity. Using (2.2)-(2.4), we can rewrite the Maxwell equations as follows:

$$\nabla \times \bar{E} = j\omega\mu_o \bar{H} \qquad \textbf{(2.5.a)}$$

$$\nabla \times \bar{H} = -j\omega\varepsilon\bar{E} \qquad \textbf{(2.5.b)}$$

$$\nabla \bullet (\varepsilon\bar{E}) = 0 \qquad \textbf{(2.5.c)}$$

$$\nabla \bullet \bar{H} = 0 \qquad \textbf{(2.5.d)}$$

We are now in a position to develop Helmholtz equations (wave equations with the time dependence removed). We take the curl of (2.5.a) and substitute (2.5.b):

$$\nabla\times\nabla\times\bar{E} = -j\omega\mu_o\nabla\times\bar{H} = \omega^2\mu_o\varepsilon\bar{E} \qquad \textbf{(2.6)}$$

Now, $\varepsilon = n^2\varepsilon_o$, where n is the refractive index and ε_o is the permittivity of free space. Hence:

$$\omega^2\mu_o\varepsilon = n^2k_o^2 \qquad \textbf{(2.7)}$$

where the wave number $k_o = 2\pi/\lambda_o$; λ_o is the free space wavelength. We can eliminate $\nabla\times\nabla\times\bar{E}$ in (2.6) using the vector identities:

$$\nabla\times\nabla\times\bar{E} = \nabla(\nabla\bullet\bar{E}) - \nabla^2\bar{E} \qquad \textbf{(2.8.a)}$$

$$\nabla\bullet(n^2\bar{E}) = \bar{E}\bullet\nabla n^2 + n^2\nabla\bullet\bar{E} = 0 \qquad \textbf{(2.8.b)}$$

where (2.8.b) uses (2.5.c). Therefore, equation (2.6) becomes:

$$\nabla^2\bar{E} + \nabla\left(\bar{E}\bullet\frac{\nabla n^2}{n^2}\right) + n^2k_o^2\bar{E} = 0 \qquad \textbf{(2.9.a)}$$

A similar function of the magnetic intensity $\bar{H}$ may be obtained:

$$\nabla^2\bar{H} + n^2k_o^2\bar{H} + \left(\frac{\nabla n^2}{n^2}\times\nabla\times\bar{H}\right) = 0 \qquad \textbf{(2.9.b)}$$

Equations (2.9.a) and (2.9.b) are known as the vector Helmholtz equations. We shall discuss the methods of solution of these equations for various waveguide types later; first, we examine some general properties of those solutions. Such an examination leads directly to some important results which are used throughout this work.

Consider a longitudinally invariant structure such that:

$$n = n(x,y) \qquad \textbf{(2.10)}$$

This longitudinal invariance suggests the existence of spatially harmonic solutions of (2.9.a&b):

$$\vec{E}_m(x,y,z) = \vec{E}_m(x,y)e^{j\beta_m z} \qquad \textbf{(2.11.a)}$$
$$\vec{H}_m(x,y,z) = \vec{H}_m(x,y)e^{j\beta_m z} \qquad \textbf{(2.11.b)}$$

where β_m is a longitudinal propagation constant which describes the fast phase evolution of the fields; the subscripts m indicate that multiple solutions of the form of (2.11.a&b) may exist. Field solutions of this form are called *modes*. We postulate for the moment that mode solutions exist for arbitrary (longitudinally invariant) structures. It is useful to decompose the mode field vectors into transverse and longitudinal components:

$$\vec{E}_{tm} = \left(E_{xm}, E_{ym}, 0\right), \quad \vec{H}_{tm} = \left(H_{xm}, H_{ym}, 0\right) \qquad \textbf{(2.12.a)}$$
$$\vec{E}_{zm} = \left(0, 0, E_{zm}\right), \quad \vec{H}_{zm} = \left(0, 0, H_{zm}\right) \qquad \textbf{(2.12.b)}$$
$$\vec{E}_m = \vec{E}_{tm} + \vec{E}_{zm}, \quad \vec{H}_m = \vec{H}_{tm} + \vec{H}_{zm} \qquad \textbf{(2.12.c)}$$

Using (2.12.a-c) we can rewrite the Maxwell equations (2.5.a&b) as follows:

$$\nabla_t \times \vec{E}_{tm} = j\omega\mu_o \vec{H}_{zm} \qquad \textbf{(2.13.a)}$$
$$\nabla_t \times \vec{H}_{tm} = -j\omega n^2 \varepsilon_o \vec{E}_{zm} \qquad \textbf{(2.13.b)}$$
$$\nabla_t \times \vec{E}_{zm} + j\beta_m \hat{z} \times \vec{E}_{tm} = j\omega\mu_o \vec{H}_{tm} \qquad \textbf{(2.13.c)}$$
$$\nabla_t \times \vec{H}_{zm} + j\beta_m \hat{z} \times \vec{H}_{tm} = -j\omega n^2 \varepsilon_o \vec{E}_{tm} \qquad \textbf{(2.13.d)}$$

where ∇_t denotes a transverse gradient operator:

$$\nabla_t = \left(\partial/\partial x, \ \partial/\partial y, \ 0\right) \qquad \textbf{(2.14)}$$

2.2 Phase Relationships Among Mode Field Components The presence of j in equations (2.13.a-d) suggests some phase differences among the various mode field components; let us first examine this. We begin by defining a time-dependent mode using equations (2.2) and (2.11):

$$\vec{e}_1(x,y,z,t) = \vec{E}_1(x,y,z)e^{-j\omega t} = \vec{E}_1(x,y)e^{j(\beta_1 z-\omega t)} \tag{2.15.a}$$

$$\vec{h}_1(x,y,z,t) = \vec{H}_1(x,y,z)e^{-j\omega t} = \vec{H}_1(x,y)e^{j(\beta_1 z-\omega t)} \tag{2.15.b}$$

If we now reverse the time-axis in (2.15), i.e. take $t \rightarrow -t$, and set:

$$\vec{e}_1(x,y,z,t) = \vec{e}_1(x,y,z,-t) \tag{2.16.a}$$

then we must have:

$$\vec{h}_1(x,y,z,t) = -\vec{h}_1(x,y,z,-t) \tag{2.16.b}$$

to satisfy the Maxwell equations (2.1.a&b). If we consider equations (2.15) to describe a forward-traveling wave, then the time-reversed function may be considered to describe an identical but *backward*-traveling wave. However, we can also construct a backward-traveling wave by reversing z in (2.15). If we set:

$$\vec{e}_{t1}(x,y,z,t) = \vec{e}_{t1}(x,y,-z,t) \tag{2.17.a}$$

then we require:

$$\vec{e}_{z1}(x,y,z,t) = -\vec{e}_{z1}(x,y,-z,t) \tag{2.17.b}$$

$$\vec{h}_{t1}(x,y,z,t) = -\vec{h}_{t1}(x,y,-z,t) \tag{2.17.c}$$

$$\vec{h}_{z1}(x,y,z,t) = \vec{h}_{z1}(x,y,-z,t) \tag{2.17.d}$$

to satisfy the Maxwell equations (2.1.a&b). From (2.16) and (2.17) we have:

$$\vec{e}_{t1}(x,y,z,-t) = \vec{e}_{t1}(x,y,-z,t) \tag{2.18.a}$$

$$\vec{e}_{z1}(x,y,z,-t) = -\vec{e}_{z1}(x,y,-z,t) \tag{2.18.b}$$

$$\vec{h}_{t1}(x,y,z,-t) = \vec{h}_{t1}(x,y,-z,t) \tag{2.18.c}$$

$$\vec{h}_{z1}(x,y,z,-t) = -\vec{h}_{z1}(x,y,-z,t) \tag{2.18.d}$$

We can rewrite (2.18.a) as follows:

$$\left\{\bar{\mathcal{E}}_{t1R}(x,y) + j\bar{\mathcal{E}}_{t1I}(x,y)\right\}e^{j(\beta_1 z+\omega t)} = \left\{\bar{\mathcal{E}}_{t1R}(x,y) + j\bar{\mathcal{E}}_{t1I}(x,y)\right\}e^{-j(\beta_1 z+\omega t)} \tag{2.19}$$

where the script letters indicate real quantities, and the subscripts R and I denote real and imaginary components, respectively, of the corresponding complex variables. From (2.19):

$$\begin{aligned}&\{\bar{\mathcal{E}}_{t1R}(x,y)+j\bar{\mathcal{E}}_{t1I}(x,y)\}\{\cos(\beta_1 z+\omega t)+j\sin(\beta_1 z+\omega t)\}\\&=\{\bar{\mathcal{E}}_{t1R}(x,y)+j\bar{\mathcal{E}}_{t1I}(x,y)\}\{\cos(\beta_1 z+\omega t)-j\sin(\beta_1 z+\omega t)\}\end{aligned}\qquad\textbf{(2.20)}$$

$$\begin{aligned}&\{\bar{\mathcal{E}}_{t1R}(x,y)\cos(\beta_1 z+\omega t)-\bar{\mathcal{E}}_{t1I}(x,y)\sin(\beta_1 z+\omega t)\}+j\{\bar{\mathcal{E}}_{t1I}(x,y)\cos(\beta_1 z+\omega t)+\bar{\mathcal{E}}_{t1R}(x,y)\sin(\beta_1 z+\omega t)\}\\&=\{\bar{\mathcal{E}}_{t1R}(x,y)\cos(\beta_1 z+\omega t)+\bar{\mathcal{E}}_{t1I}(x,y)\sin(\beta_1 z+\omega t)\}+j\{\bar{\mathcal{E}}_{t1I}(x,y)\cos(\beta_1 z+\omega t)-\bar{\mathcal{E}}_{t1R}(x,y)\sin(\beta_1 z+\omega t)\}\end{aligned}$$

(2.21)

Only the real parts of (2.21) have physical significance; therefore (2.21) can be reduced to:

$$\begin{aligned}&\bar{\mathcal{E}}_{t1R}(x,y)\cos(\beta_1 z+\omega t)-\bar{\mathcal{E}}_{t1I}(x,y)\sin(\beta_1 z+\omega t)\\&=\bar{\mathcal{E}}_{t1R}(x,y)\cos(\beta_1 z+\omega t)+\bar{\mathcal{E}}_{t1I}(x,y)\sin(\beta_1 z+\omega t)\end{aligned}\qquad\textbf{(2.22)}$$

Hence:

$$\bar{\mathcal{E}}_{t1I}(x,y)=0\qquad\textbf{(2.23)}$$

and:

$$\bar{E}_{t1}(x,y)=\bar{E}^{*}_{t1}(x,y)\qquad\textbf{(2.24.a)}$$

In a similar fashion we obtain from (2.18.b-d):

$$\bar{E}_{z1}(x,y)=-\bar{E}^{*}_{z1}(x,y)\qquad\textbf{(2.24.b)}$$

$$\bar{H}_{t1}(x,y)=\bar{H}^{*}_{t1}(x,y)\qquad\textbf{(2.24.c)}$$

$$\bar{H}_{z1}(x,y)=-\bar{H}^{*}_{z1}(x,y)\qquad\textbf{(2.24.d)}$$

Therefore,the transverse and longitudinal components of the electric field are out of phase by a quarter-wave; the same is true for the magnetic fields. In other words:

$$\bar{E}_m(x,y) = \pm\bar{\mathcal{E}}_{tm}(x,y) \pm j\bar{\mathcal{E}}_{zm}(x,y) \tag{2.25.a}$$

$$\bar{H}_m(x,y) = \pm\bar{\mathcal{H}}_{tm}(x,y) \pm j\bar{\mathcal{H}}_{zm}(x,y) \tag{2.25.b}$$

Note that (2.25) ensures the modes can transport power only along the axis of propagation. The direction of power transport of mode m is that of the Poynting vector:

$$\bar{s}_m = \bar{e}_m \times \bar{h}_m$$

$$= (e_{my}h_{mz} - e_{my}h_{mz})\hat{x} + (e_{mz}h_{mx} - e_{mx}h_{mz})\hat{y} + (e_{mx}h_{my} - e_{my}h_{mx})\hat{z} \tag{2.26}$$

The x- and y-components of the Poynting vector contain only products of transverse and longitudinal field terms, and these components are always orthogonal in time, e.g.:

$$\bar{e}_m(x,y,z,t) = \bar{E}_m(x,y)e^{j(\beta_m z-\omega t)} = \left(\pm\bar{\mathcal{E}}_{tm}(x,y) \pm j\bar{\mathcal{E}}_{zm}(x,y)\right)e^{j(\beta_m z-\omega t)}$$

$$= \left\{\pm\bar{\mathcal{E}}_{tm}(x,y)\cos(\beta_m z-\omega t) \mp \bar{\mathcal{E}}_{zm}(x,y)\sin(\beta_m z-\omega t)\right\} + j\left\{\pm\bar{\mathcal{E}}_{zm}(x,y)\cos(\beta_m z-\omega t) \pm \bar{\mathcal{E}}_{tm}(x,y)\sin(\beta_m z-\omega t)\right\} \tag{2.27}$$

The time-averaged x- and y-components of the Poynting vector are therefore zero, and it is easy to show that the time-averaged mode Poynting vector is effectively:

$$\bar{S}_m(x,y) = \frac{1}{2}\mathrm{Re}\left\{\bar{E}_{tm}(x,y) \times \bar{H}^*_{tm}(x,y)\right\} \tag{2.28.a}$$

$$= \frac{1}{4}\left\{\bar{E}_{tm}(x,y) \times \bar{H}^*_{tm}(x,y) + \bar{E}^*_{tm}(x,y) \times \bar{H}_{tm}(x,y)\right\} \tag{2.28.b}$$

We shall now proceed to show that the modes of a waveguide comprise an orthogonal set which can be used as a basis for representing arbitrary fields. This analysis begins with the reciprocity theorem.

2.3 The Lorentz Reciprocity Theorem We define two modes, identified by the subscripts 1 and 2. From (2.13a&b) we obtain:

$$\nabla \times \vec{E}_1 = j\omega\mu_o \vec{H}_1 \quad \textbf{(2.29.a)}$$

$$\nabla \times \vec{H}_2^* = j\omega n^2 \varepsilon_o \vec{E}_2^* \quad \textbf{(2.29.b)}$$

Taking the scalar product of (2.29.a) and $\vec{H}_2^*$, and that of (2.29.b) and $\vec{E}_1$, we have:

$$\vec{H}_2^* \bullet \left(\nabla \times \vec{E}_1\right) = j\omega\mu_o \vec{H}_1 \bullet \vec{H}_2^* \quad \textbf{(2.30.a)}$$

$$\vec{E}_1 \bullet \left(\nabla \times \vec{H}_2^*\right) = j\omega n^2 \varepsilon_o \vec{E}_1 \bullet \vec{E}_2^* \quad \textbf{(2.30.b)}$$

Subtracting, we obtain:

$$\vec{H}_2^* \bullet \left(\nabla \times \vec{E}_1\right) - \vec{E}_1 \bullet \left(\nabla \times \vec{H}_2^*\right) = j\omega\left(\mu_o \vec{H}_1 \bullet \vec{H}_2^* - n^2 \varepsilon_o \vec{E}_1 \bullet \vec{E}_2^*\right) \quad \textbf{(2.31)}$$

Using the vector identity:

$$\nabla \bullet \left(\vec{a} \times \vec{b}\right) = \vec{b} \bullet \left(\nabla \times \vec{a}\right) - \vec{a} \bullet \left(\nabla \times \vec{b}\right) \quad \textbf{(2.32)}$$

we can obtain from (2.31):

$$\nabla \bullet \left(\vec{E}_1 \times \vec{H}_2^*\right) = j\omega\left(\mu_o \vec{H}_1 \bullet \vec{H}_2^* - n^2 \varepsilon_o \vec{E}_1 \bullet \vec{E}_2^*\right) \quad \textbf{(2.33.a)}$$

Similarly, we can derive:

$$\nabla \bullet \left(\vec{E}_2^* \times \vec{H}_1\right) = j\omega\left(n^2 \varepsilon_o \vec{E}_1 \bullet \vec{E}_2^* - \mu_o \vec{H}_1 \bullet \vec{H}_2^*\right) \quad \textbf{(2.33.b)}$$

Adding (2.33.a) and (2.33.b), we obtain:

$$\nabla \bullet \left(\vec{E}_1 \times \vec{H}_2^* + \vec{E}_2^* \times \vec{H}_1\right) = 0 \quad \textbf{(2.34)}$$

This is one version of the Lorentz reciprocity theorem.

2.4 Orthogonality of Modes If we now consider two modes, identified by subscripts m and p:

$$\bar{E}_m(x,y,z) = \bar{E}_m(x,y)e^{j\beta_m z} \tag{2.35.a}$$

$$\bar{E}_p(x,y,z) = \bar{E}_p(x,y)e^{j\beta_p z} \tag{2.35.b}$$

and substitute these into (2.34), we obtain:

$$\nabla_t \bullet \left(\bar{E}^* \times \bar{H}_p + \bar{E}_p \times \bar{H}_m^*\right) - j\left(\beta_p - \beta_m\right)\left(\bar{E}_{tp} \times \bar{H}_{tm}^* + \bar{E}_{tm}^* \times \bar{H}_{tp}\right) \bullet \hat{z} = 0 \tag{2.36}$$

If we integrate (2.36) over all space in the x-y plane, we get:

$$\int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} \left(\bar{E}_{tp} \times \bar{H}_{tm}^* + \bar{E}_{tm}^* \times \bar{H}_{tp}\right) dx\,dy = 0 \text{ if } \beta_m \neq \beta_p \tag{2.37}$$

since the integral of the first term in (2.36) vanishes, a result guaranteed by Green's theorem (Gauss' theorem in the plane):

$$\iint_A \left(\nabla_t \bullet \bar{F}\right) dA = \oint_C \left(\bar{F} \bullet \hat{s}\right) ds \tag{2.38}$$

where C is the curve bounding the area A. We anticipate that modes may be of either limited or unlimited transverse extent. For modes of finite transverse extent, which we will later call *bound* modes, the zero value of the contour integral is assured by choosing a contour beyond the region where the modes have non-zero magnitudes. For modes which extend throughout all space, which we call *radiation* modes, the zero value of the contour integral can be obtained by choosing C such that the mode field magnitudes are appropriately periodic on the contour.

Since the integrand of (2.37) has the form of a Poynting vector as in equation (2.28.b), the implication of (2.37) is that the total power transported by a group of modes is the sum of the powers carried by each mode. In other words, there is no "cross-mode" power. Equation (2.37) thus demonstrates the power orthogonality of waveguide modes.

2.5 Mode Normalization The orthogonality of the set of waveguide modes suggests their utility as a basis for describing arbitrary fields; to this end, we seek to obtain an *orthonormal* set of functions by appropriately normalizing the mode fields. We therefore define normalized fields as follows:

$$\hat{\bar{E}}_{tm} = \bar{E}_{tm}\Big/\sqrt{N_m}, \quad \hat{\bar{H}}_{tm} = \bar{H}_{tm}\Big/\sqrt{N_m} \qquad \textbf{(2.39.a,b)}$$

where N_m is a normalization factor. The orthogonality relation (2.37) suggests a normalization with respect to power. Modes of finite transverse extent, which we called bound modes, are power-integrable and therefore we use (2.28.b) to define N_m for these modes as:

$$N_m = \frac{1}{4}\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left(\bar{E}_{tm}\times\bar{H}^*_{tm} + \bar{E}^*_{tm}\times\bar{H}_{tm}\right)dxdy \quad \text{(bound modes)} \qquad \textbf{(2.40)}$$

For modes which extend through all space, the so-called radiation modes, we cannot use (2.40) since then the normalized fields would be identically zero. The standard way to overcome this problem is to normalize the radiation modes with respect to (transverse) power *density* rather than total power, i.e. to write:

$$P_{\text{radiation mode } m} = \frac{1}{4}\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left(\bar{E}_{tp}\times\bar{H}^*_{tm} + \bar{E}^*_{tm}\times\bar{H}_{tp}\right)dxdy = N_m\delta(m-p) \qquad \textbf{(2.41.a)}$$

where δ is the Dirac delta; for finite N_m the radiation mode power is therefore infinite. From (2.41.a) we see that if N_m has units of areal power density, then from the property $\delta(at-at')=\delta(t-t')/a$, m must have dimensions of length^{-2}. We shall see later that in fact m is the product of the x- and y-propagation constants, which define the transverse character of the mode. We can therefore explicitly write N_m for radiation modes as:

$$N_m = \frac{1}{4}\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left(\vec{E}_{tm}\times\vec{H}_{tm}^{*}+\vec{E}_{tm}^{*}\times\vec{H}_{tm}\right)\delta(x-x')\delta(y-y')dxdy \quad \text{(radiation modes)} \quad \textbf{(2.41.b)}$$

$$= \frac{1}{4}\left(\bar{E}_{tm}\times\bar{H}_{tm}^{*}+\vec{E}_{tm}^{*}\times\bar{H}_{tm}\right)\Big|_{x',y'} \quad \textbf{(2.41.c)}$$

where δ are again Dirac delta functions and x' and y' are chosen as follows. Any waveguide consists essentially of a certain transverse area over which the local index varies, surrounded by an essentially semi-infinite region or regions of homogeneous index. Since the average power density of a radiation mode is determined only by the fields in these regions and is not affected by the field amplitude variations in the relatively small vicinity of the waveguide itself, it is appropriate to choose x' and y' deep in homogeneous transverse space (i.e. far away from the waveguide) where the fields are constant-valued.

The fact that bound modes are power-integrable, whereas radiation modes are not, suggests something about the collective natures of these types of modes. Since a given bound mode may transport a finite power, only a finite number of these modes can physically exist, i.e. the set of bound modes must be finite and countable - a discrete set. By contrast, our definitions above provide for infinite power transport by individual radiation modes; therefore, a single radiation mode cannot be physically excited. However, a superposition of such modes can collectively transport a finite power if they comprise an uncountable - i.e. continuous - set. The situation is analogous to the Fourier transform of a finite signal; such a signal can be represented by an integral superposition of infinite sinusoidal functions.

2.6 Field Expansion Using Mode Functions

We have demonstrated that solutions of the vector Helmholtz equations (2.9.a&b) of the form (2.11.a&b) constitute an orthogonal, normalizable set. If we postulate that they are also a *complete* set (this is difficult to prove *a priori* in the general case), then the set of mode functions (2.11) may be used as a basis for representing arbitrary fields. Neglecting for the moment the inferred continuous character of the radiation modes, we anticipate constructions of the form:

$$\bar{E}_t(x,y) = \sum_m a_m \hat{\bar{E}}_{tm}(x,y) \quad \textbf{(2.42.a)}$$

$$\bar{H}_t(x,y) = \sum_m a_m \hat{\bar{H}}_{tm}(x,y) \tag{2.42.b}$$

where the coefficients a_m are complex mode amplitudes. The value of the coefficients a_m is obtained by first taking the vector products of (2.42.a&b) and mode fields n:

$$\bar{E}_t \times \hat{\bar{H}}^*_{tn} = \sum_m a_m \hat{\bar{E}}_{tm} \times \hat{\bar{H}}^*_{tn} \tag{2.43.a}$$

$$\bar{H}_t \times \hat{\bar{E}}^*_{tn} = \sum_m a_m \hat{\bar{H}}_{tm} \times \hat{\bar{E}}^*_{tn} \tag{2.43.b}$$

Adding (2.43.a) and (2.43.b), we get:

$$\bar{E}_t \times \hat{\bar{H}}^*_{tn} + \hat{\bar{E}}^*_{tn} \times \bar{H}_t = \sum_m \left\{ a_m \left(\hat{\bar{E}}_{tm} \times \hat{\bar{H}}^*_{tn} + \hat{\bar{E}}^*_{tn} \times \hat{\bar{H}}_{tm} \right) \right\} \tag{2.44}$$

Integrating over all space, and making use of (2.37), (2.39) and (2.40), we obtain, swapping $n \rightarrow m$:

$$a_m = \frac{1}{4} \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} \left(\bar{E}_t \times \hat{\bar{H}}^*_{tm} + \hat{\bar{E}}^*_{tm} \times \bar{H}_t \right) dx dy \tag{2.45}$$

by assuming that $\beta_m \neq \beta_n$.

Two comments are called for regarding (2.42)-(2.45). First, only forward-propagating waves have been assumed; if the field contains backward-propagating components we must replace the coefficients in (2.42.a) by:

$$a_m + a_{-m} \tag{2.46.a}$$

and those in (2.42.b) by:

$$a_m - a_{-m} \tag{2.46.b}$$

where the subscripts *-m* indicate the backward-propagating modes; (2.45) must also be appropriately modified. Note that the z-inversion properties of (2.17.a-d) have been included in (2.46).

Second, we have anticipated that, unlike bound modes, radiation modes form a continuous set. Therefore, an integral term must be added to equations (2.42.a&b) to account for the continuous nature of the set of radiation modes, e.g.:

$$\bar{E}_t(x,y) = \sum_m a_m \hat{\bar{E}}_{tm}(x,y) + \int_{\mathscr{m}} a(\mathscr{m})\hat{\bar{E}}_{t\mathscr{m}}(x,y)\,d\mathscr{m} \tag{2.47}$$

where the sum includes the bound modes and the integral is over the set of radiation modes; the script $\mathscr{m}$ denotes a continuous variable. Equation (2.42.b) is similarly modified. The normalized fields in the sum are constructed using (2.40), whereas those in the integral use (2.41). Since we have used the same expansion coefficients for the electric and magnetic fields, from (2.37) it is clear that a_m has dimensions of square root power, and comparing (2.40) and (2.41) we see that $a(\mathscr{m})$ must have dimensions of square root power times length. As already mentioned the parameter m (or script $\mathscr{m}$) is in fact the product of two parameters, the lateral propagation constants β_x and β_y, where:

$$\beta_x^2 + \beta_y^2 + \beta^2 = n^2 k_o^2, \tag{2.48}$$

and nk_o=the local phase velocity, so that in fact:

$$d\mathscr{m} = d\beta_x d\beta_y \tag{2.49.a}$$

Both β_x and β_y may be required to fully specify the transverse form of a mode and thus provide a unique mode identifier or index. This is true also for the bound modes and hence the sum index in (2.47) is:

$$m = \beta_x, \beta_y \tag{2.49.b}$$

We can find approximate values of $a(\mathscr{m})$ by discretizing the integral in (2.47):

$$\int_{\mathscr{m}} a(\mathscr{m})\hat{\bar{E}}_{t\mathscr{m}}(x,y)\,d\mathscr{m} \approx \sum_{\mathscr{m}_i} a(\mathscr{m}_i)\hat{\bar{E}}_{t\mathscr{m}_i}(x,y)\,\Delta\mathscr{m} \tag{2.50}$$

We can therefore obtain approximations to $a(m)$ for the radiation modes by replacing:

$$\hat{\bar{E}}_{lm}(x,y) \rightarrow \hat{\bar{E}}_{lm}(x,y)\Delta m, \quad \hat{\bar{H}}_{lm}(x,y) \rightarrow \hat{\bar{H}}_{lm}(x,y)\Delta m \tag{2.51}$$

in equation (2.45).

Using equation (2.28.b), the power transported by the total field is:

$$P = \frac{1}{4}\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left\{\bar{E}_t \times \bar{H}_t^* + \bar{E}_t^* \times \bar{H}_t\right\}dxdy \tag{2.52.a}$$

$$= \frac{1}{4}\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left\{\left(\sum_m a_m \hat{\bar{E}}_{tm}(x,y) + \int_m a(m)\hat{\bar{E}}_{lm}(x,y)dm\right) \times \left(\sum_p a_p^* \hat{\bar{H}}_{tp}^*(x,y) + \int_p a^*(p)\hat{\bar{H}}_{lp}^*(x,y)dp\right)\right.$$

$$\left. + \left(\sum_p a_p^* \hat{\bar{E}}_{tp}^*(x,y) + \int_p a^*(p)\hat{\bar{E}}_{lp}^*(x,y)dp\right) \times \left(\sum_m a_m \hat{\bar{H}}_{tm}(x,y) + \int_m a(m)\hat{\bar{H}}_{lm}(x,y)dm\right)\right\}dxdy \tag{2.52.b}$$

$$= \sum_m \sum_p a_m a_p^* \left[\frac{1}{4}\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left\{\hat{\bar{E}}_{tm} \times \hat{\bar{H}}_{tp}^* + \hat{\bar{E}}_{tp}^* \times \hat{\bar{H}}_{tm}\right\}dxdy\right]$$

$$+ \int_m \int_p a(m)a^*(p)\left\{\frac{1}{4}\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left\{\hat{\bar{E}}_{lm}(x,y) \times \hat{\bar{H}}_{lp}^*(x,y) + \hat{\bar{E}}_{lp}^*(x,y) \times \hat{\bar{H}}_{lm}(x,y)\right\}dxdy\right\}dp\,dm \tag{2.52.c}$$

$$= \sum_m |a_m|^2 + \int_m \int_p a(m)a^*(p)\delta(m-p)\,dp\,dm \tag{2.52.d}$$

$$= \sum_m |a_m|^2 + \int_m |a(m)|^2 dm \tag{2.52.e}$$

where we have used (2.37), (2.39)-(2.41) and (2.47).

The amplitude coefficients a thus contain all information about the power content of the modes. If the total power is normalized to one, i.e. if we set:

$$P = \frac{1}{4}\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left\{\bar{E}_t \times \bar{H}_t^* + \bar{E}_t^* \times \bar{H}_t\right\}dxdy = 1 \tag{2.53}$$

then the coefficients a become dimensionless, and we have:

$$= \sum_m |a_m|^2 + \int_m |a(m)|^2 dm = 1 \tag{2.54}$$

This is a particularly useful formalism since now the mode amplitude coefficients squared tell us simply the fraction of the total power transported by each mode.

2.7 Waveguide Transitions The technique just discussed of field expansion using mode functions can be used to describe field effects at waveguide transitions along the axis of propagation. Let us assume that a waveguide's transverse structure changes suddenly at z=0. The fields at $z=0^-$ must equal those at $z=0^+$. Therefore, if the pre- and post-transition waveguides are sufficiently similar that power reflected at the transition can be neglected, then the amplitude coefficients of the post-transition waveguide modes are simply given by (2.45) where the total fields are taken as those at $z=0^-$, i.e. in the pre-transition waveguide. For bound modes, using (2.39), (2.40), (2.45) and normalizing the pre-transition field using (2.53), we have:

$$a_m = \frac{\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left(\bar{E}_o \times \bar{H}_{tm}^* + \bar{E}_{tm}^* \times \bar{H}_o\right)dxdy}{\sqrt{\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left(\bar{E}_o \times \bar{H}_o^* + \bar{E}_o^* \times \bar{H}_o\right)dxdy}\sqrt{\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left(\bar{E}_{tm} \times \bar{H}_{tm}^* + \bar{E}_{tm}^* \times \bar{H}_{tm}\right)dxdy}} \tag{2.55}$$

where:

$$\bar{E}_o,\ \bar{H}_o \tag{2.56}$$

are the fields at $z=0^-$. A corresponding result can be obtained for the radiation modes; though (2.55) can be approximately applied to these modes by appropriately discretizing the radiation mode continuum, using (2.50) and (2.51), and replacing the second integral in the denominator with (2.41.c).

Equation (2.55) is a highly important result which is applied many times in this work. Since the pre-transition field is normalized, the coefficients a_m describe the fraction of the incident power which is coupled into mode m, i.e.:

$$\frac{P_{\text{mode m}}}{P_{\text{incident}}} = |a_m|^2 \qquad \textbf{(2.57)}$$

2.8 Modes of Planar Waveguides We have seen that modal decomposition of propagating fields is a useful method of analyzing waveguide behaviour. We now take a closer look at modes, examining their form and properties in various types of waveguides, and how they may be calculated. As a first step, we examine planar waveguides, which not only illustrate the general characteristics of waveguide modes, but which, as we shall see, can also serve as a useful approximate model for many three-dimensional structures.

Planar, or slab, waveguides are structures that are two-dimensional, in the sense that the refractive index distribution is invariant along one axis:

$$n = n(x,z) \qquad \textbf{(2.58)}$$

For such structures, one complete set of solutions of the vector Helmholtz equations (2.9.a&b) has the property that the fields are invariant with respect to y, i.e.:

$$\frac{\partial F}{\partial y} = 0 \qquad \textbf{(2.59)}$$

for any field F. We restrict the following discussion to the case of such solutions. It is simple to show that under these conditions, (2.9.a&b) become partially uncoupled

[11] and hence two independent (i.e. orthogonal in the sense of (2.37)) sets of solutions exist: transverse electric (TE), and transverse magnetic (TM):

$$\text{TE Solutions} \quad \begin{cases} \bar{E}(x,z) = (0, E_y, 0) \\ \bar{H}(x,z) = (H_x, 0, H_z) \end{cases} \qquad \textbf{(2.60.a)}$$

$$\text{TM Solutions} \quad \begin{cases} \bar{E}(x,z) = (E_x, 0, E_z) \\ \bar{H}(x,z) = (0, H_y, 0) \end{cases} \qquad \textbf{(2.60.b)}$$

If we assume a longitudinally invariant structure so that:

$$n = n(x) \qquad \textbf{(2.61)}$$

and we substitute solutions of the form of (2.11.a&b) into (2.9.a&b), then, given (2.58)-(2.60), we have:

$$\frac{d^2 E_{ym}(x)}{dx^2} + \left(k_o^2 n^2(x) - \beta_m^2\right) E_{ym}(x) = 0 \qquad \textbf{(2.62)}$$

for TE modes, and:

$$n^2(x) \frac{d}{dx}\left(\frac{1}{n^2(x)} \frac{dH_{ym}(x)}{dx}\right) + \left(k_o^2 n^2(x) - \beta_m^2\right) H_{ym}(x) = 0 \qquad \textbf{(2.63)}$$

for TM modes. The remaining field components are obtained directly from the Maxwell equations (2.5.a-d) as:

$$H_{xm}(x) = \frac{-\beta_m}{\omega\mu_o} E_{ym}(x), \quad H_{zm}(x) = \frac{-j}{\omega\mu_o} \frac{\partial E_{ym}(x)}{\partial x} \qquad \textbf{(2.64.a,b)}$$

for TE modes, and:

$$E_{xm}(x) = \frac{\beta_m}{\omega\varepsilon_o n^2(x)} H_{ym}(x), \quad E_{zm}(x) = \frac{j}{\omega\varepsilon_o n^2(x)} \frac{\partial H_{ym}(x)}{\partial x} \qquad \textbf{(2.65.a,b)}$$

for TM modes. Note that the phase relationships predicted in (2.25) are in fact obtained. The properties of the solutions of (2.62) and (2.63) may be drawn directly from the well-known Sturm-Liouville theory, discussed in detail in this context in [12], which deals with differential equations of the form.

$$\frac{d}{dx}\left\{p(x)\frac{d\psi(x)}{dx}\right\}+\{q(x)+\lambda r(x)\}\psi(x)=0 \qquad \textbf{(2.66)}$$

From (2.62) we see that the respective functions in (2.66) are, for TE modes:

$$p(x)=r(x)=1, \quad q(x)=k_o^2n^2(x) \qquad \textbf{(2.67)}$$

while for TM modes, we have, from (2.63):

$$p(x)=r(x)=1/n^2(x), \quad q(x)=k_o^2 \qquad \textbf{(2.68)}$$

and:

$$\lambda=-\beta_m^2 \qquad \textbf{(2.69)}$$

for both TE and TM modes. If the refractive index distribution *n(x)* is real, bounded and piecewise-continuous over an x-interval (a,b), then the conditions of the Sturm-Liouville theory are satisfied and we can therefore immediately draw the following conclusions from the theory:

i) The non-trivial solutions of (2.62) and (2.63) are eigenpairs, i.e. a set of eigenfunctions $E_{ym}(x)$ for TE modes and $H_{ym}(x)$ for TM modes, with corresponding eigenvalues λ_m (different for TE and TM) as in equation (2.69). If the interval (a,b) is the infinite x-interval, then this set of eigenvalues consists of a continuum which can extend to $+\infty$, and may be followed by a set of discrete eigenvalues bounded by some negative number. Given (2.69) this translates to a discrete set of real β_m followed by a continuum of lower (absolute) values of real β_m, followed again by a continuum of imaginary β_m. We shall see that the discrete β_m solutions correspond to bound modes, as we have already inferred, while the continuous set of real β_m corresponds to radiation modes. The continuum of imaginary β_m solutions are called *evanescent* modes; these are similar to radiation modes in transverse character but from (2.11) are exponential rather than periodic under propagation. For reasons that will be explained, these modes have been neglected in our analysis.

ii) The sets of eigenfunctions $E_{ym}(x)$, $H_{ym}(x)$ are orthonormal in the sense of the scalar product:

$$\int_a^b \{E/H_{ym}(x)\}\{E/H_{yn}(x)\} r(x)dx = \delta_{m,n} \tag{2.70}$$

where E/H in (2.70) means E or H, while δ is the Dirac delta for continuum modes and the Kronecker delta for discrete modes. This confirms our previous general conclusion that waveguide modes are an orthogonal, normalizable set where the normalization requires the Dirac delta for radiation modes.

iii) The sets of eigenfunctions $E_{ym}(x)$, $H_{ym}(x)$ are complete in the sense that any field function which decays sufficiently rapidly near the x-interval endpoints can be represented, to an arbitrary precision, by an expansion of the form of (2.42).

The results (i)-(iii) above stem directly from the mathematical properties of Sturm-Liouville equations, without any resort to purely physical arguments. Results (ii) and (iii) confirm our earlier general inferences, while (i) confirms the existence of discrete and continuous eigensolutions of the Helmholtz equations.

An important class of slab waveguides is structures composed of layers of uniform refractive index separated by sharp boundaries. In such cases, in order to avoid singular derivative evaluations at the layer interfaces, equations (2.62) and (2.63) require continuity of:

$$E_{ym},\ \frac{dE_{ym}}{dx} \quad \text{for TE modes; and} \tag{2.71.a}$$

$$H_{ym},\ \frac{1}{n^2(x)}\frac{dH_{ym}}{dx} \quad \text{for TM modes.} \tag{2.71.b}$$

In fact, from (2.64) and (2.65) we can recognize (2.71.a&b) as the well-known requirement that the tangential components of $\bar{E}$ and $\bar{H}$ be continuous across a dielectric boundary [13]. Therefore, the modes of uniform-layered slab structures may be found by writing solutions of:

$$\frac{d^2E_{ym}(x)}{dx^2}+\left(k_o^2n^2(x)-\beta_m^2\right)E_{ym}(x)=0 \tag{2.72}$$

for TE modes, and:

$$\frac{d^2H_{ym}(x)}{dx^2}+\left(k_o^2n^2(x)-\beta_m^2\right)H_{ym}(x)=0 \tag{2.73}$$

for TM modes, in each layer of uniform refractive index, and then using the appropriate continuity relations (2.71.a or b) to join the solutions in each layer. It is apparent from (2.72) and (2.73) that if, in a given layer:

$$n_{em}\equiv\frac{\beta_m}{k_o}<n(x) \tag{2.74}$$

then $E_{ym}(x)$ or $H_{ym}(x)$ is oscillatory in that layer; otherwise, $E_{ym}(x)$ or $H_{ym}(x)$ is exponential. The new parameter n_e is called the *effective index* of the mode. Based on this, we can examine qualitatively the characteristics of the mode solutions as a function of effective index. See Figure 2.1; here, we show the various types of TE modes for a symmetric three-layer slab waveguide. The index of the center ("core" or "guide") layer is n_2, while that of the semi-infinite surrounding ("cladding") layers is n_1; $n_2>n_1$. For $n_{em}>n_2$, (or rather, $|n_{em}|>n_2$ to take account of backward-propagating modes) the fields E_{ym} are exponential in all three layers, and the continuity relations (2.71.a) therefore require that the fields are exponentially increasing in the claddings. Such fields are not power-normalizable, either individually or upon superposition, and are therefore disallowed by physical arguments. Furthermore, Sturm-Liouville theory dictates that all solutions of (2.66) must be bounded over the whole solution interval; hence such solutions are also prohibited by purely mathematical arguments. For solutions such that $n_2>n_{em}>n_1$, the fields E_{ym} are exponential in the claddings and sinusoidal in the core; the continuity relations therefore require solutions of the form shown. These energy of these modes is confined to a specific transverse area in the vicinity of the core and they therefore correspond to the bound modes referred to previously. These bound modes form a discrete set in the interval $n_2>n_{em}>n_1$, according to both our earlier physical arguments and Sturm-Liouville theory.

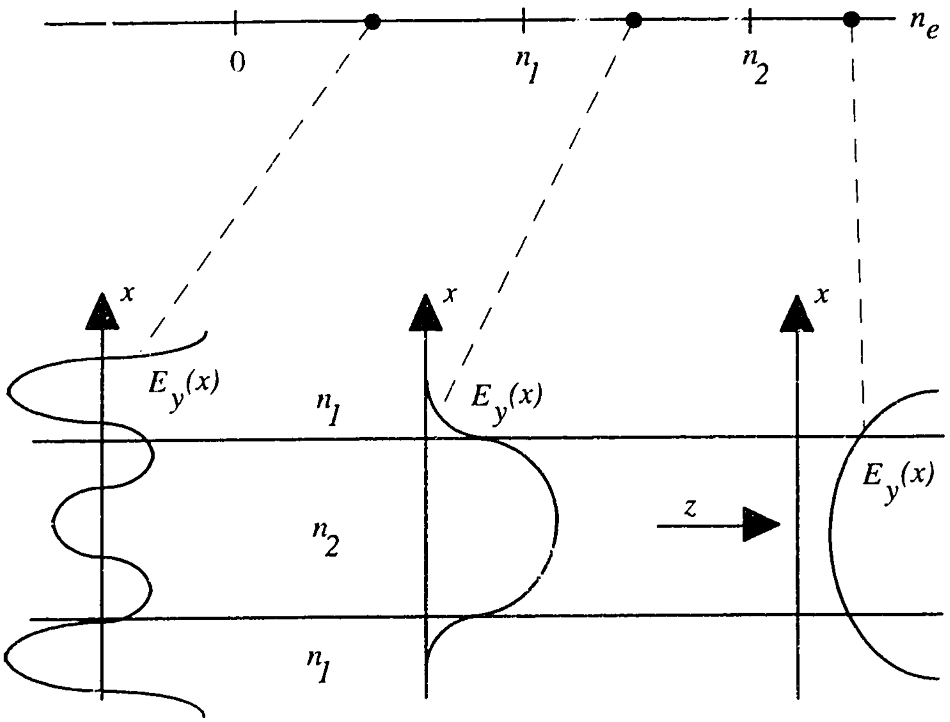


Figure 2.1 Parametric characterization of three-layer, symmetric planar waveguide TE modes with respect to mode effective index n_e. Solutions for $n_e>n_2$ are exponential everywhere and are disallowed. A discrete set of bound modes may exist in the interval $n_2>n_e>n_1$, while a continuous set of radiation modes of infinite transverse extent exists in the interval $n_1>n_e>0$. Evanescent modes with imaginary n_e may also exist; these have transverse field characteristics similar to the radiation modes.

For solutions in the range $n_1>n_{em}>0$, the solutions are sinusoidal in all layers. These are the radiation modes, which extend over all transverse space. From the comments (i) above, these modes form a continuum. Evanescent modes such that $n_{em}=\pm j\eta$, where η is real, are in general allowed by Sturm-Liouville theory, as discussed above. Such modes are also transversally sinusoidal in all layers, but by (2.11) they are exponential rather than periodic with respect to propagation and therefore they may be disallowed by longitudinal boundary conditions. For example, if the waveguide is of infinite length, i.e. if it extends over $-\infty < z < \infty$, then no evanescent modes can exist as they would blow up exponentially at large $\pm z$. This argument is used to exclude evanescent modes from the consideration of spontaneous emission in Chapter 6. Throughout this work, we also exclude evanescent modes from our consideration of waveguide transitions (section 2.7). This is justified if the post-transition waveguide is always long, prohibiting significant excitation of modes such that n_{em}=+jη, while modes such that n_{em}=-jη will decay to negligible amplitudes at large z and therefore are of little practical importance.

The parametric behaviour of TM modes is similar, but the solutions satisfy the interface conditions (2.71.b). For simple planar waveguides, the solutions of (2.71)-(2.73) may be given explicitly; this is done in Chapter 6 for the case of a symmetric three-layer structure. For more complicated structures, a numerical approach is better-suited. In general, the continuity conditions (2.71.a) render TE fields more amenable to numerical solution than TM fields. Any of the standard numerical methods for fixed-boundary-value solutions of ordinary differential equations may be used to solve (2.72) with (2.71.a). In particular, the finite difference methods are relatively simple to implement and have good stability characteristics [14]. Consider the following centered-differences approximation to the second derivative of E_{ym}:

$$\frac{d^2E_{ym}(x_i)}{dx^2} = \frac{1}{h^2}\left[E_{ym}(x_{i+1}) - 2E_{ym}(x_i) + E_{ym}(x_{i-1})\right] + O\left(h^2\right) \qquad \textbf{(2.75)}$$

where h is a discretization interval and $x_i = ih$; *O( )* means "to the order of" and provides an error upper-bound. The approximation (2.75) is derived using Taylor's expansion theorem which requires the continuity of the expanded variable *and* its derivatives, up to third order in this case. Therefore, a similar approximation cannot be obtained for TM fields; we shall return to this point shortly. Applying (2.75) to (2.72) we have:

$$\psi_{i-1}^m + \left\{n_i^2k_o^2h^2 - 2\right\}\psi_i^m + \psi_{i+1}^m = \beta_m^2h^2\psi_i^m \qquad \textbf{(2.76)}$$

where we have set $\psi_i^m = E_{ym}(x_i)$ and $n_i = n(x_i)$. If we divide (2.76) through by $k_o^2 h^2$ then we obtain:

$$\frac{1}{\bar{h}^2}\psi_{i-1}^m + \left\{n_i^2 - \frac{2}{\bar{h}^2}\right\}\psi_i^m + \frac{1}{\bar{h}^2}\psi_{i+1}^m = n_{em}^2\psi_i^m \tag{2.77}$$

where $\bar{h} = k_o h$. Equation (2.77) is suitably scaled for computation. If we assume that the "window" of computation $x_{i=0} \rightarrow x_{i=n+1}$ is sufficiently large to properly contain all of the bound mode fields, then we can take $x_0 = x_{n+1} = 0$ and apply (2.77) at every node i ($i=1, \ldots, n$) to obtain the matrix eigenequation:

$$[A]\{\psi^m\} = n_{em}^2\{\psi^m\} \tag{2.78}$$

where $\{\psi^m\}$ is the vector consisting of the values of $\psi_i^m = E_{ym}(x_i)$, and the matrix $[A]$ is given by:

$$[A] = \begin{bmatrix} \zeta_1 & \frac{1}{\bar{h}^2} & 0 & \cdots & \cdots & 0 \\ \frac{1}{\bar{h}^2} & \zeta_2 & \frac{1}{\bar{h}^2} & 0 & \cdots & 0 \\ & & & \cdots & & \\ 0 & \cdots & \cdots & 0 & \frac{1}{\bar{h}^2} & \zeta_n \end{bmatrix} \tag{2.79.a}$$

where :

$$\zeta_i = n_i^2 - \frac{2}{\bar{h}^2} \tag{2.79.b}$$

The symmetric positive definite matrix $[A]$ has n eigensolutions with real eigenvalues n_{em}^2. Those with effective index eigenvalues $n_{em} > n_{min}$, where n_{min} is the larger of the indices of the upper and lower cladding layers, are approximations to the bound mode fields, as per our comments on page 26. The remainder may be considered a sampling of the set of radiation modes. The numerical solution of the matrix eigenequation (2.78) is the most difficult aspect of this formulation. Iterative procedures such as the power method are simple to implement but are not well suited to finding more than one eigensolution. In this work we have used the package *LANZ* [15], which employs the Lanczos algorithm with appropriate spectral transformations [16]. *LANZ* is able to quickly find a large number of matrix eigenpairs and furthermore is able to find solutions to more general eigenequations of the form:

$$[A]\{\psi^m\} = \lambda[B]\{\psi^m\} \tag{2.80}$$

where the matrix $[B]$ may be positive semi-definite. This capability is required for the finite element solution of non-planar waveguide modes described in section 2.11, and for the work of Chapter 4, where curved waveguides are analyzed.

The method described above is not suited to TM fields due to the interface conditions (2.71.b); iterative methods based on transfer matrices are therefore usually used to find TM modes. However, we note here that if:

$$\frac{1}{n^2(x)}\frac{dn^2(x)}{dx}\frac{dH_{ym}(x)}{dx} << \frac{d^2H_{ym}(x)}{dx^2} \tag{2.81}$$

everywhere, i.e. if the changes in refractive index are small, then (2.63) reduces approximately to an equation of the same form as (2.62) for the TE fields [17], and hence for both TE and TM fields we can solve a *scalar* Helmholtz equation of the form:

$$\frac{d^2\psi}{dx^2} + \left(k_o^2n^2 - \beta^2\right)\psi = 0 \tag{2.82}$$

where ψ is a field function representing E_y for TE fields and H_y for TM fields; the mode index subscripts m have been dropped for simplicity, as have the explicit x-dependences of ψ and the refractive index n. The solutions of (2.82) are exactly correct for TE fields but only approximately correct for TM fields. We can obtain an indication of the error yielded by (2.82) in the propagation constants β for TM modes, by first writing (2.82) in variational form. Multiplying by ψ^* and integrating over all x-space, we have for bound modes:

$$\beta^2 = \frac{\int\limits_{x=-\infty}^{\infty} \left\{ k_o^2 n^2 |\psi|^2 - \left| \frac{d\psi}{dx} \right|^2 \right\} dx}{\int\limits_{x=-\infty}^{\infty} |\psi|^2 dx} \tag{2.83}$$

where we have used integration by parts. It can be shown [18] that (2.83) is a true variational expression in that the exact solutions to (2.82) yield stationary maxima of (2.83), i.e. a reasonable approximation to a field solution ψ of (2.82) introduces only a second order error in β when applied in (2.83), and this incorrect β is less than the exact value. If we apply a similar process to equation (2.63) we can derive:

$$\beta_{TM}^2 = \frac{\int\limits_{x=-\infty}^{\infty} \left\{ k_o^2 n^2 |H_y|^2 - \left| \frac{dH_y}{dx} \right|^2 - \frac{1}{n^2} \frac{dn^2}{dx} \frac{dH_y}{dx} H_y^* \right\} dx}{\int\limits_{x=-\infty}^{\infty} |H_y|^2 dx} \tag{2.84}$$

where we have used the subscript TM to indicate that (2.84) yields the correct propagation constants for the TM modes. Replacing H_y in (2.84) with the scalar solutions ψ of (2.82) as a first order approximation, and subtracting (2.84) from (2.83), we obtain:

$$\beta^2 - \beta_{TM}^2 \le \frac{\int\limits_{x=-\infty}^{\infty} \frac{1}{n^2} \frac{dn^2}{dx} \frac{d\psi}{dx} \psi^* dx}{\int\limits_{x=-\infty}^{\infty} |\psi|^2 dx} \tag{2.85}$$

Noting that for solutions of (2.82) the product:

$$\frac{dn^2}{dx} \frac{d\psi}{dx} \psi^* \tag{2.86}$$

is always positive, we can derive, in the limit of a piecewise-uniform refractive index distribution $n(x)$:

$$\beta^2 - \beta_{TM}^2 \le \frac{2\sum \frac{n_H - n_L}{n_L} \left| \frac{d\psi}{dx} \psi^* \right|}{\int\limits_{x=-\infty}^{\infty} |\psi|^2 dx} \tag{2.87}$$

where the summation is over the layer interfaces, at which the quantities in the numerator are evaluated. The parameters n_H and n_L are the high and low refractive indices, respectively, on either side of each interface. Since (2.87) uses the scalar fields as approximations to the true TM fields H_y, we can expect (2.87) to yield accurate estimates of β_{TM} only as long as the index variations are small. We can refine (2.87) by attempting to account for the correct derivative continuity condition on H_y. From (2.71.b), we relate the derivatives of H_y at either side of an index interface:

$$\frac{1}{n_L^2} \frac{dH_y}{dx}\Big|_L = \frac{1}{n_H^2} \frac{dH_y}{dx}\Big|_H \tag{2.88}$$

We now define an "average" value of the derivative across the interface as:

$$\frac{dH_y}{dx}\Big|_{avge} = \frac{1}{2}\left(\frac{dH_y}{dx}\Big|_H + \frac{dH_y}{dx}\Big|_L \right) = \frac{1}{2}\left(1 + \frac{n_L^2}{n_H^2} \right) \frac{dH_y}{dx}\Big|_H \tag{2.89}$$

Replacing $\frac{d\psi}{dx} \approx \frac{dH_y}{dx}\Big|_H$ in (2.87) with $\frac{dH_y}{dx}\Big|_{avge}$ we get:

$$\beta^2 - \beta_{TM}^2 \approx \frac{\sum \frac{(n_H - n_L)(n_H^2 + n_L^2)}{n_H^2 n_L} \left| \frac{d\psi}{dx} \psi^* \right|}{\int\limits_{x=-\infty}^{\infty} |\psi|^2 dx} \tag{2.90}$$

where (2.90) is no longer strictly variational.

In Figure 2.2 are shown calculated values of the fundamental mode effective indices of a symmetric three-layer planar waveguide with core index 1.50 and thickness 2.5 μm, for a wavelength of 1.55 μm, as a function of the refractive index of the semi-infinite cladding layers. The exact TE (scalar) and TM effective indices are compared with those obtained from the scalar solutions using (2.87) and (2.90), indicated by "M1" and "M2" respectively. Equation (2.90) is seen to yield much better estimates of $n_e^{TM} = \beta_{TM} / k_o$ than (2.87) when the core/cladding index difference is large.

Equations (2.87) and (2.90) provide convenient estimations of the propagation constants of bound TM-polarized planar waveguide modes, obtained from the solutions of the scalar formulation (2.82), where the scalar fields and their derivatives are taken as continuous with respect to x. The important point here is that for planar waveguides, the scalar Helmholtz equation (2.82), which is well-suited to numerical solution when field and field derivative continuity properties are associated, in fact directly yields the TE mode solutions, and also provides a good starting point for examining the TM modes. This, to a large extent, justifies an exclusively scalar analysis of planar waveguide modes. We shall see that similar arguments pertain to non-planar waveguides in the limit of strong field confinement in a transverse region of uniform vertical index distribution ("strong guiding"), so that a scalar treatment is justified in that case also. The work of Chapters 3 and 4 deals with precisely this kind of waveguide.

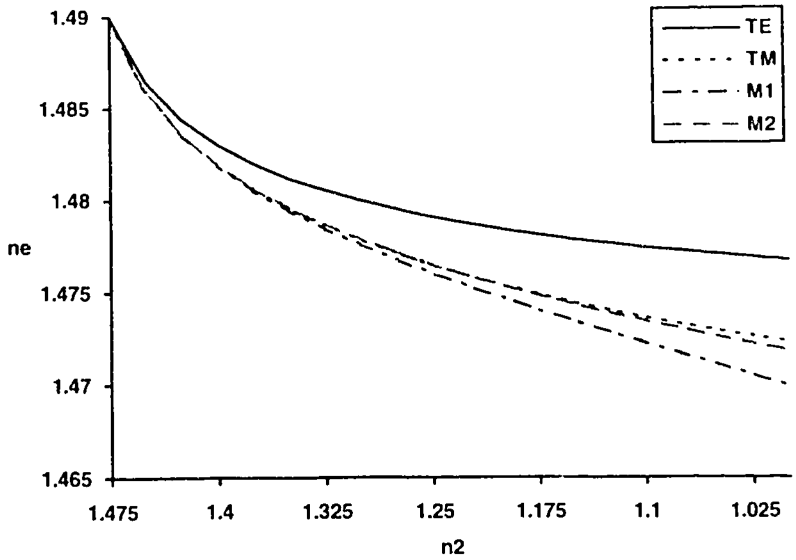
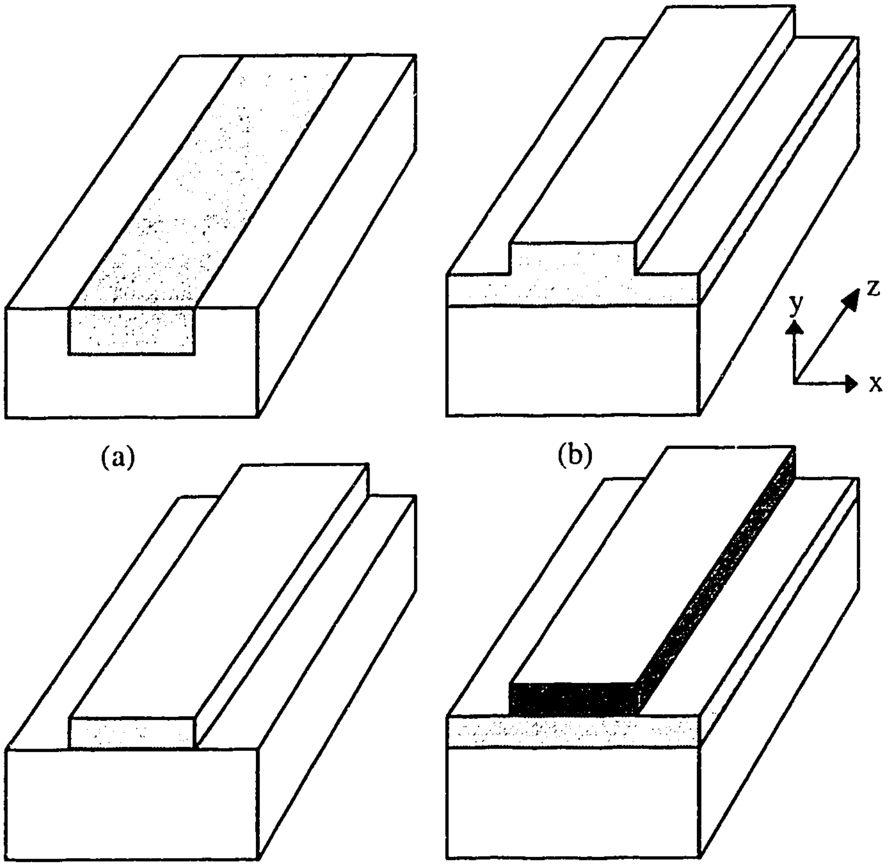


Figure 2.2 Comparison of the fundamental mode effective indices, "ne", of a planar waveguide with index distribution n2/1.50/n2 and core thickness 2.5 μm, for a wavelength of 1.55 μm. Curves "TE" and "TM" show the exact values for the TE (scalar) and TM modes respectively, while "M1" and "M2" are calculated from the scalar solutions using (2.87) and (2.90) respectively.

2.9 Modes of Non-Planar Waveguides We have seen that the scalar Helmholtz equation, characterized by a single field component and associated with field and field derivative continuity properties, is particularly easily adapted to numerical solution. In planar waveguides, the convenient separability of fields into orthogonal TE and TM polarizations facilitates the application of a scalar treatment. However, for non-planar waveguides, i.e. those for which the refractive index distribution varies along both transverse axes (see Figure 2.3), the situation is much less clear. In this case orthogonal polarization states do not exist, so that the modes of non-planar waveguides are always *hybrid*, meaning that all three components of both the electric and magnetic field vectors are non-zero. Six scalar field components are thus required to fully describe a non-planar waveguide mode, versus three for planar waveguides. We saw that for a planar waveguide mode, we need only calculate one field component, since the remaining two components are functions of the first (see (2.60), (2.64) and (2.65)). The hybrid polarization (lack of orthogonal polarization states) of non-planar waveguide modes requires us to find *two* transverse field components: it is necessary to simultaneously calculate both of either, for example, E_x and E_y, or H_x and H_y. The remaining four field components are then accessible using the Maxwell equations [19]. This situation considerably complicates the task of numerical mode solution, since no single field component can be selected for computation as in the planar case, and hence a scalar treatment is, in general, inherently unsuitable.



(a) (b)

(c) (d)

Figure 2.3 Some standard non-planar waveguide configurations for integrated optics; shading differences indicate material (refractive index) differences: (a) channel waveguide; (b) ridge waveguide; (c) rib waveguide (a limiting case of the ridge configuration); (d) strip-loaded waveguide.

Historically, the first numerical treatments of non-planar dielectric waveguides were fully vectorial, in the sense that they solved equations for the transverse components of $\bar{E}$ or $\bar{H}$, obtained by substituting (2.11.a) or (2.11.b) into one of the vector Helmholtz equations (2.9.a) or (2.9.b). The usual choice is (2.9.b) for the magnetic intensity, since for non-magnetic materials all components of $\bar{H}$ are continuous across arbitrarily oriented dielectric interfaces or gradients. However, the field component derivatives are not necessarily continuous and hence a two-dimensional version of the centered-differences scheme outlined previously is unsuitable. The complicated form of (2.9.b) requires a powerful numerical method for solution; finite element methods based on variational formulations of (2.9.b) and not restricted by derivative continuity conditions have been widely used [20],[21]. However, such formulations are highly complex and furthermore often lead to spurious, non-physical solutions which are not easily differentiated from the true modes.

A considerable simplification of the problem is obtained by assuming the non-planar waveguide modes to be "quasi-polarized". This means that while a mode is, in fact, hybrid and therefore consists of six non-zero field components, the six components are often not equally significant, so that for practical problems we may be able to neglect one of the two transverse field components. This reasoning leads to the class of so-called "semi-vectorial" methods [22],[23]. Consider a general electromagnetic wave in a homogeneous medium. It is simple to show that such a wave is always TEM (transverse electric *and* magnetic, i.e. $E_Z=H_Z=0$) and thus consists of only four non-zero field components [13]. For conceptual purposes, a TEM wave may be considered the sum of a "quasi-TE" wave and a "quasi-TM" wave, each with two non-zero field components. Real waveguides are generally near-homogeneous in the sense that refractive index variations are relatively small except at air interfaces, where the strong index difference prohibits significant field penetration. Therefore, for most non-planar waveguide structures, a field description in terms of quasi-polarized modes (where the quasi-polarizations are, generally, not degenerate as in homogeneous space) is expected to be quite accurate. These modes are defined as follows:

$$\text{Quasi-TE Solutions} \quad \begin{cases} \bar{E}(x,y,z)=\left(E_x,0,E_z\right) \\ \bar{H}(x,y,z)=\left(H_x,H_y,H_z\right) \end{cases} \qquad \textbf{(2.91.a)}$$

$$\text{Quasi - TM Solutions} \quad \begin{cases} \bar{E}(x,y,z)=\left(E_x,E_y,E_z\right) \\ \bar{H}(x,y,z)=\left(H_x,0,H_z\right) \end{cases} \qquad \textbf{(2.91.b)}$$

For quasi-TE modes, we expect that $E_x>>E_z$ and $H_y>>H_z>>H_x$, while for quasi-TM modes we anticipate that $H_x>>H_z$ and $E_y>>E_z>>E_x$. Each quasi-polarized mode has only one independent field component, E_x (or H_y) in the quasi-TE case and H_x (or E_y) in the quasi-TM case, which are used as the bases for computation in the semi-vectorial methods. These methods thus exploit the fact that either $E_x>>E_y$ and $H_y>>H_x$, or $E_y>>E_x$ and $H_x>>H_y$ for the hybrid modes of real non-planar dielectric waveguide structures to achieve a considerable simplification of the mathematical formalism. For example, in [23] the vector Helmholtz equation (2.9.a) is reduced to the following form:

$$\left(\frac{\partial}{\partial x}\frac{1}{n^2}\frac{\partial n^2}{\partial x}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}+k_o^2n^2\right)E_x=0 \qquad \textbf{(2.92.a)}$$

for quasi-TE modes, and:

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial}{\partial y}\frac{1}{n^2}\frac{\partial n^2}{\partial y}+\frac{\partial^2}{\partial z^2}+k_o^2n^2\right)E_y=0 \qquad \textbf{(2.92.b)}$$

for quasi-TM modes, where (2.92.b) uses $E_x \approx 0$. Note that the arguments justifying the semi-vectorial methods are neither general nor rigorous; they have gained acceptance only by demonstrating results close to those of fully vectorial treatments for a range of standard optical waveguide configurations.

If, for a given waveguide, the transverse index variations are small, then the semi-vectorial equations (2.92) reduce approximately to a single scalar equation, where by scalar we again mean that the equation contains only a single field component variable, and that both the field and its derivative are continuous across all transverse index boundaries or gradients. Such a formulation is obviously a large approximation in the general case, but may be a very good one, at least for one quasi-polarization, in certain instances. For example, consider the ridge waveguide of Figure 2.3(b). For quasi-TE modes, where we assume $E_y \approx 0$, the only index interfaces for which $\nabla_t \bar{E}$ is not continuous are the vertical sidewalls of the ridge. However, if the ridge is air-clad, the large index discontinuity will prevent significant lateral field penetration out of the ridge, so that the effect of the ridge sidewalls may be insignificant. This argument is examined in [24] using a variational analysis similar to equations (2.83)-(2.90). From equation (2.92.a) above and the two-dimensional scalar modal Helmholtz equation:

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+k_o^2 n^2-\beta^2\right)\psi=0, \tag{2.93}$$

we can derive the following expression using a procedure similar to that described earlier (equations (2.83)-(2.90)):

$$\beta^2-\beta_{TE}^2 \approx \frac{\dfrac{n_2^2-1}{n_2^2}\displaystyle\int_{sidewall}\left|\frac{d\psi}{dx}\psi^*\right|dy}{\displaystyle\int_{y=-\infty}^{\infty}\int_{x=0}^{\infty}|\psi|^2\,dx\,dy} \tag{2.94}$$

where n_2 is the index in the ridge and the lateral center of the ridge is at x=0. Air-cladding and lateral symmetry of the ridge have been assumed and the integral is along the vertical ridge sidewall. Equation (2.94) provides an estimate of the propagation constants of the quasi-TE modes using the solutions of the scalar equation (2.93). A similar formulation may be obtained for the quasi-TM modes; the correction in this case involves integration of:

$$\left|\frac{d\psi}{dy}\psi^*\right| \tag{2.95}$$

over all the horizontal layer interfaces.

Equation (2.94) implies that if the integrand in the numerator is insignificant, i.e. if the scalar field and/or its derivative are virtually zero along the ridge sidewall, then the scalar Helmholtz equation (2.93) directly yields the quasi-TE field solutions. This condition is produced as the ridge depth becomes large (i.e. tending towards the rib configuration of Figure 2.3(c)), since the mode fields tend to be constrained within the raised ridge itself, and to lie less under the ridge where they can spread laterally. It is this lateral field spread that leads to non-zero field magnitudes and derivatives toward the bottom of the ridge sidewall as is easy to understand. If E_x immediately under the ridge spreads laterally beyond the x-position of the sidewall, then it has (generally) a non-zero value at that position. But since E_x must be continuous across horizontal layer interfaces, it must also have a non-zero value in the air immediately above the ridge bottom and just outside the ridge wall. It was shown in [24] that in the limit of large ridge depth the correction term in (2.94) does indeed approach zero, and the scalar equation (2.93) produces the same mode fields and propagation constants as the quasi-TE solutions of both semi- and fully vectorial methods, for a typical refractive index distribution. We refer to waveguides for which this condition holds as "strongly guiding" or "strongly confining", since a large ridge depth tends to strongly confine the mode fields laterally, within the bounds of the ridge.

Note that the condition of strong mode confinement implies a separability of the mode fields with respect to the transverse axes:

$$E_x(x,y) = \psi(x,y) = F(x)G(y) \qquad \textbf{(2.96)}$$

since effectively the fields are confined to a region of uniform vertical refractive index profile, i.e. under the ridge. Equation (2.96) implies that a separation of (2.93) itself may be possible. This reasoning leads us to an examination of the application of the well-known "effective index method" to strongly confining waveguides.

2.10 Application of the Effective Index Method to Strongly Confining Waveguides

Let us assume that the solutions to the two-dimensional scalar Helmholtz equation (2.93) can be written in the following form:

$$\psi(x,y) = F(x)G(x,y) \qquad \textbf{(2.97)}$$

where $G(x,y)$ is a piecewise-uniform function of x. For a ridge/rib waveguide that is laterally symmetric about x=0, and where the ridge/rib width is a, we thus have:

$$G(x,y)=\begin{cases}G_1(y) & |x|\le a/2\\ G_2(y) & |x|>a/2\end{cases} \tag{2.98}$$

Using (2.97) the scalar Helmholtz equation (2.93) becomes:

$$G\frac{d^2F}{dx^2}+2\frac{\partial G}{\partial x}\frac{dF}{dx}+F\left\{\frac{\partial^2 G}{\partial x^2}+\frac{\partial^2 G}{\partial y^2}\right\}+\left(k_o^2n^2-\beta^2\right)FG=0 \tag{2.99}$$

Equation (2.99) is "piecewise-separable" according to (2.98) so that:

$$\frac{1}{F}\frac{d^2F}{dx^2}-\beta^2=-\frac{1}{G_1}\frac{d^2G_1}{dy^2}-k_o^2n^2(x,y)=-K_1 \quad |x|\le a/2 \tag{2.100.a}$$

$$\frac{1}{F}\frac{d^2F}{dx^2}-\beta^2=-\frac{1}{G_2}\frac{d^2G_2}{dy^2}-k_o^2n^2(x,y)=-K_2 \quad |x|>a/2 \tag{2.100.b}$$

where K_1 and K_2 are constants. From (2.100) we therefore have the following:

$$\frac{d^2G_1}{dy^2}+\left\{k_o^2n^2\left(|x|\le a/2,y\right)-K_1\right\}G_1=0 \tag{2.101.a}$$

$$\frac{d^2G_2}{dy^2}+\left\{k_o^2n^2\left(|x|>a/2,y\right)-K_2\right\}G_2=0 \tag{2.101.b}$$

$$\frac{d^2F}{dx^2}+\left\{K(x)-\beta^2\right\}F=0 \tag{2.101.c}$$

where:

$$K(x)=\begin{cases}K_1 & |x|\le a/2\\ K_2 & |x|>a/2\end{cases} \tag{2.101.d}$$

Equations (2.101.a-c) have the same form as the scalar planar Helmholtz equation (2.82) and therefore discrete and continuum solutions are expected. If the real structure in fact supports a bound mode then (2.101.a) always has at least one discrete solution; the number of such solutions indicates the number of vertical orders of modes which exist for the two-dimensional structure. In any case, the bound solution(s) of (2.101.a) should be chosen for application in (2.101.c). If discrete solutions of (2.101.b) also exist, then these correspond to vertically bound fields outside the rib and likewise should be selected for application in (2.101.c). Applied this way, the method described by equations (2.101.a-d) is well-known as the so-called effective index method [25].

However, (2.101.b) may have only continuum solutions, even for waveguides which do in fact support bound modes, and the best way to proceed in such cases has been the subject of some contention. We note that if we choose:

$$K_2 = k_o^2 \qquad \textbf{(2.102)}$$

then :

$$G_2(y) = 0, \quad \text{for } y \text{ along sidewall} \qquad \textbf{(2.103)}$$

is a valid non-trivial solution of (2.101.b). From (2.94) this corresponds to the condition under which a scalar analysis yields the correct solutions for quasi-TE modes. Applying (2.102) in (2.101) we therefore arrive at the following formulation:

$$\frac{d^2G_1}{dy^2} + \left\{k_o^2 n^2\left(|x| \le a/2, y\right) - K_1\right\}G_1 = 0 \qquad \textbf{(2.104.a)}$$

$$\frac{d^2F}{dx^2} + \left\{K(x) - \beta^2\right\}F = 0 \qquad \textbf{(2.104.b)}$$

where:

$$K(x) = \begin{cases} K_1 & |x| \le a/2 \\ k_o^2 & |x| > a/2 \end{cases} \qquad \textbf{(2.104.c)}$$

From the preceeding arguments we expect equations (2.104.a-c) to produce the correct solutions for bound quasi-TE fields as long as the condition (2.103) is satisfied for the true mode fields. Note that (2.103) in fact means that, within the ridge, the mode fields are *strictly* separable according to (2.96). We can thus check the applicability of (2.104) by verifying the validity of (2.96), for (x,y) within the ridge, using the calculated two-dimensional scalar fields. As an even stronger condition we can check that (2.96) applies everywhere. As a general rule, (2.96) is satisfied for true rib waveguides that are air-clad. This is convincingly demonstrated in [26].

We have therefore come to the quite remarkable conclusion that for strongly confining rib waveguides, the effectively one-dimensional, scalar treatment described above yields the correct quasi-TE bound mode solutions. This is of considerable value since the simple finite differences method described earlier (equations (2.75)-(2.79)) is then sufficient to examine such structures. It is necessary, though, to restrict the set of solutions of (2.104.b) to those with effective indices greater than the substrate index, since, by our earlier parametric arguments, other discrete solutions cannot correspond to bound modes of the two-dimension structure and may be considered spurious.

Note that the above arguments also imply that a considerable simplification of equation (2.55), which gives the amplitudes of the waveguide modes following an abrupt longitudinal transition, is possible for strongly confining structures. Where the description of the propagating fields as quasi-polarized and the separation (2.96) are valid for both the pre- and post-transition fields, (2.55) reduces to:

$$a_m = \frac{\int_{-\infty}^{\infty} F_o(x) F_m^*(x)\,dx}{\sqrt{\int_{-\infty}^{\infty} |F_o(x)|^2 dx}\sqrt{\int_{-\infty}^{\infty} |F_m(x)|^2 dx}} \qquad \textbf{(2.105)}$$

since the electric fields and magnetic intensities are related by an approximately uniform impedance which is the same for the pre- and post-transition fields.

2.11 Finite Element Solution of the Two-Dimensional Scalar Helmholtz Equation

Though we have demonstrated that the semi-one-dimensional scalar analysis of equations (2.104.a-c) provides an accurate analysis of the quasi-TE bound mode fields of non-planar waveguides in the limit of strong guiding, a method of solving the two-dimensional scalar Helmholtz equation (2.93) is nevertheless required for the work of Chapter 5, where the strong guiding condition is not satisfied. In addition, it is desirable to be able to solve (2.93) directly to verify that the separability condition (2.96) is satisfied in cases where the strong guiding condition is expected. The formulation (2.104.a-c) can subsequently be used to obtain accurate solutions for high-order modes and to quickly examine the effects of design variations.

A two-dimensional version of the centered-differences scheme applied earlier is applicable to (2.93). However, in its basic form, such a scheme requires a structured (uniform) grid which, for acceptable precision of the computed solutions, generally leads to unacceptably large matrices in eigenequations of the form of (2.78). Although finite difference schemes based on non-uniform grids can be derived [27], the class of variational schemes using finite elements generally offers the possibility of higher solution precision using relatively small-dimension matrices [21],[28],[29]. The trade-off is that the finite element methods themselves are inherently more complex than the finite difference schemes.

To develop an applicable finite element method, we take ψ as real in (2.93), multiply (2.93) by ψ and use integration by parts to obtain, for bound modes:

$$\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left\{k_o^2n^2\psi^2-\beta^2\psi^2-\left(\frac{\partial\psi}{\partial x}\right)^2-\left(\frac{\partial\psi}{\partial y}\right)^2\right\}dxdy=0 \qquad \textbf{(2.106)}$$

As we saw earlier, (2.106) is a variational expression in that actual solutions ψ of (2.93) render stationary a functional of the form:

$$\Omega(\psi)=\int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}\left\{k_o^2n^2\psi^2-\beta\;\psi^2-\left(\frac{\partial\psi}{\partial x}\right)^2-\left(\frac{\partial\psi}{\partial y}\right)^2\right\}dxdy \qquad \textbf{(2.107)}$$

i.e. $\partial\Omega(\psi)/\partial\psi\neq 0$ if ψ is not a solution of (2.93). We divide the solution interval into triangles and define the field function within each triangle, or element, as:

$$\psi(x,y)=\sum_{i=1}^{3}\psi_i\alpha_i(x,y) \qquad \textbf{(2.108)}$$

where ψ_i are the field values at the element nodes and $\alpha_i(x,y)$ are linear shape functions which interpolate the field value over the whole element and are given by, for example:

$$\alpha_1(x,y)=\frac{1}{2A}\left\{(x_2y_3-x_3y_2)+(y_2-y_3)x+(x_3-x_2)y\right\} \tag{2.109}$$

where A is the element area, x_i, y_i are the element nodes and the other α_i are obtained by cyclic interchanges of the subscripts. Higher-order interpolations may be used; we shall comment on this later. The shape functions are defined to have unity value at the nodes, i.e.:

$$\alpha_i(x_i,y_j)=\begin{cases}0 & i\neq j\\ 1 & i=j\end{cases} \tag{2.110}$$

We now write the terms of Ω in each element using (2.108). For example:

$$\frac{\partial\psi}{\partial x}=\sum_{i=1}^{3}\psi_i\frac{\partial\alpha_i}{\partial x}=\frac{y_2-y_3}{2A}\psi_1+\frac{y_3-y_1}{2A}\psi_2+\frac{y_1-y_2}{2A}\psi_3 \tag{2.111}$$

and:

$$\iint_{\text{element}}\left(\frac{\partial\psi}{\partial x}\right)^2 dxdy=\sum_{i=1}^{3}\sum_{j=1}^{3}\psi_i\iint_{\text{element}}\frac{\partial\alpha_i}{\partial x}\frac{\partial\alpha_j}{\partial x}dxdy\,\psi_j$$

$$=\{\psi\}^T\iint_{\text{element}}\frac{\partial\{\alpha\}^T}{\partial x}\frac{\partial\{\alpha\}}{\partial x}dxdy\{\psi\} \tag{2.112}$$

where { } denotes a three-vector and the superscript T denotes a transpose operation. Similar expressions are obtained for the other terms in (2.107). We assemble the resulting functionals on each element to obtain:

$$\Omega_{\text{disjoint}} = \sum_{\text{elements}} \{\psi\}^T \iint_{\text{element}} \left(\left(k_o^2 n^2 - \beta^2\right)\{\alpha\}^T\{\alpha\} - \frac{\partial\{\alpha\}^T}{\partial x}\frac{\partial\{\alpha\}}{\partial x} - \frac{\partial\{\alpha\}^T}{\partial y}\frac{\partial\{\alpha\}}{\partial y} \right) dxdy \{\psi\}$$

$$= (\psi)^T_{\text{disjoint}} \left([A]_{\text{disjoint}} - \beta^2 [B]_{\text{disjoint}} \right) (\psi)_{\text{disjoint}} \qquad \textbf{(2.113)}$$

where (ψ) is a global vector comprised of the field values at all the nodes, and:

$$[A]_{\text{disjoint}} = \sum_{\text{elements}} \iint_{\text{element}} \left(k_o^2 n^2 \{\alpha\}^T\{\alpha\} - \frac{\partial\{\alpha\}^T}{\partial x}\frac{\partial\{\alpha\}}{\partial x} - \frac{\partial\{\alpha\}^T}{\partial y}\frac{\partial\{\alpha\}}{\partial y} \right) dxdy \text{ ,} \textbf{(2.114)}$$

$$[B]_{\text{disjoint}} = \sum_{\text{elements}} \iint_{\text{element}} \{\alpha\}^T\{\alpha\} dxdy \qquad \textbf{(2.115)}$$

The subscript "disjoint" indicates that the elements are not yet connected in any way. Therefore $(\psi)_{\text{disjoint}}$ has $3N$ components, where N is the number of elements, while $[A]_{\text{disjoint}}$ and $[B]_{\text{disjoint}}$ are square matrices of order $3N$. The redundancy of the shared nodes on adjacent elements is eliminated by the use of a connectivity matrix $[C]$, i.e.:

$$(\psi)_{\text{disjoint}} = [C](\psi)_{\text{joint}} \qquad \textbf{(2.116)}$$

where $[C]$ forces the identity of field values of the disjoint vector which correspond to the same node, and also enforces a homogeneous Dirichlet boundary condition on the edges of the computational window. The use of the reduced dimension conjoint vector thus ensures both field continuity throughout the solution domain, and the appropriate boundary condition for guided modes. Note that (2.116) implies that we use $[C]$ to obtain the disjoint vector from the joint vector; in fact, the connectivity matrix is used to modify the disjoint matrices $[A]$ and $[B]$, to obtain an eigenvalue equation which is solved directly for $(\psi)_{\text{joint}}$. From (2.113) and (2.116), we have:

$$\Omega_{\text{joint}} = (\psi)^T_{\text{joint}} \left([A]_{\text{joint}} - \beta^2 [B]_{\text{joint}} \right) (\psi)_{\text{joint}} \qquad \textbf{(2.117)}$$

where:

$$[A]_{\text{joint}} = [C]^T [A]_{\text{disjoint}} [C], \tag{2.118}$$

$$[B]_{\text{joint}} = [C]^T [B]_{\text{disjoint}} [C] \tag{2.119}$$

Enforcing stationarity of Ω_{joint}, i.e. setting:

$$\frac{\partial \Omega_{\text{joint}}}{\partial (\psi)_{\text{joint}}} = 0 \tag{2.120}$$

we obtain the matrix eigenequation:

$$[A]_{\text{joint}} (\psi)_{\text{joint}} - \beta^2 [B]_{\text{joint}} (\psi)_{\text{joint}} = 0 \tag{2.121}$$

In this work, a pre-existing FORTRAN code [30], which implements an algorithm similar to that just described, was modified to produce the correct matrices $[A]_{\text{joint}}$ and $[B]_{\text{joint}}$, after scaling (2.93) to new coordinates:

$$x \rightarrow xk_o, \quad y \rightarrow yk_o \tag{2.122}$$

which are better suited for computation. This transformation changes the eigenvalue equation (2.121) to:

$$[A]_{\text{joint}} (\psi)_{\text{joint}} - n_e^2 [B]_{\text{joint}} (\psi)_{\text{joint}} = 0 \tag{2.123}$$

with (2.114) and (2.115) appropriately modified. The package *LANZ*, which is used to solve (2.78) for planar waveguides, is also used to solve (2.123). The matrices $[A]$ and $[B]$ are not as sparse as in the planar waveguide case, and the eigenvalues found by *LANZ* seem to be generally less accurate, or at least quite sensitive to the distribution of the elements throughout the problem region, especially for high-order modes (non-maximum eigenvalues). Since *LANZ* seems to be more robust in the face of increased matrix size than reduced sparseness, the linear shape function described above, in conjunction with a large number of small elements, is probably better-suited than a higher-order interpolation scheme used with a smaller number of elements. Similar conclusions have been mentioned by other authors [20].

2.12 Modal Analysis versus Beam Propagation Methods The work of Chapter 5 deals with longitudinally invariant structures, and hence the determination of bound modes using finite elements is sufficient to characterize the stable propagating fields. In Chapters 3 and 4, however, we examine structures which are longitudinally varying and hence no one set of bound modes provides a sufficient characterization.

The waveguide transition relations (2.55) and (2.105) hint at one way of examining longitudinally varying waveguides. If the structure is "piecewise-invariant" then it may be convenient to represent the propagating field throughout the structure in terms of local modes, having complex amplitudes determined by the transition relations at each junction between invariant sections. Such an examination is referred to as a "modal" analysis. For a modal analysis to be applicable two conditions must be fulfilled. First, the structure must be adequately representable by a few longitudinally invariant sections, each supporting a relatively small number of bound modes, to keep the required calculations to a practical number. Note that longitudinal-invariance may be defined with respect to a local coordinate system which may differ between sections. Second, the propagating fields of interest must be adequately represented by bound modes only, since the transition calculations (2.55) or (2.105) are convenient only for the bound modes. Furthermore, the numerical methods we have introduced cannot properly calculate radiation modes. Including only the bound waveguide modes means that any power carried by or coupled into radiation modes is neglected, or at least assumed absent at subsequent transitions. This is normally acceptable since it is usually the power retained in bound modes at the end of a waveguide structure that is of interest. However, a modal analysis is not suitable for applications where the distribution of power radiated from a structure is important.

A second way to analyze longitudinally varying structures involves solving a Helmholtz equation where strictly harmonic z-dependence of the fields is not assumed. Let us presume that a scalar treatment is valid and assume a field of the form:

$$\psi(x,y,z)=\phi(x,y,z)e^{j\beta z} \qquad (2.124)$$

where ψ on the right-hand side of (2.124) is made a weak function of z which may be considered a field "envelope". This envelope has a fast phase variation with respect to z given by the exponential, with a uniform phase constant β for the whole field. Using (2.124) in a scalar approximation to (2.9.a) or (2.9.b) we obtain:

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}+k_o^2n^2-\beta^2\right)\phi=-2j\beta\frac{\partial\phi}{\partial z} \qquad (2.125)$$

A simplified "slowly varying envelope" approximation to (2.125), often used in practice, is obtained by setting the second derivative with respect to z to zero. Methods which solve equations of the form of (2.125) as an initial value problem are widely used and are known as "beam propagation methods" (BPM) [31]-[35]. These methods approximate the behaviour of near-axial radiating fields quite well and furthermore can be applied to continuously varying waveguide structures.

However, solving (2.125) in three dimensions is a formidable numerical task, and the validity of methods which purport to effectively reduce non-planar structures to two dimensions for propagation analysis is unclear. Where discrete solutions to the vertical slab Helmholtz equations exist everywhere, the effective index method leads to a good representation of bound mode fields, but the relationship of the continuum modes of the resulting lateral structure to the true radiation modes of the corresponding non-planar structure is uncertain, particularly where strong lateral effective index differences exist. Furthermore, as we have already mentioned, in strong guiding situations the effective index method described by (2.104.a-c) leads to spurious bound mode solutions, which are easily identified by physical arguments, but which cannot be excluded in solving a general propagation equation. A technique has been proposed to overcome this limitation [36], but it relies on a somewhat subjective fitting procedure which is difficult to apply to waveguides which support more than one bound mode. Furthermore there is no reason to believe that the technique provides an accurate representation of radiation modes. Indeed, a comparison of the technique applied to a two-dimensional BPM treatment of a bent waveguide, compared to both a fully three-dimensional analysis and experiment [37], suggests that it does not. We therefore must conclude that, at least for strongly guiding structures, a reliable BPM analysis requires a three-dimensional formulation. Since the work of Chapters 3 and 4 deals with such structures, and the conditions of applicability of a modal analysis are met, a modal analysis is used rather than a more complicated BPM.

Chapter 3

Analysis and Design of Multi-Mode Couplers in Deeply Etched Waveguides and Implementation in an InP-Based Mach-Zehnder Switch

Directional couplers are an important component in integrated optics. In general, these devices rely on the beating of multiple excited modes to shift the transverse intensity distribution under propagation. By choosing the propagation length carefully, a desired spatial transfer of the guided optical energy can be achieved. Furthermore, phase modulation can be used to change the phase relationships of the various guided field components, so that active devices using directional couplers can achieve controlled spatial optical routing. The conventional or "synchronous" directional coupler implementation places two single-mode waveguides in sufficient proximity that their evanescent fields overlap, creating a coupled waveguide system; see Figure 3.1. This lifts the degeneracy of the modes of the individual waveguides and creates symmetric and anti-symmetric "supermodes" of the coupled system, with separated propagation constants. An input to one of the guides excites both supermodes, which pass in and out of phase as they propagate; the lateral distribution of the total guided field therefore changes as a function of propagation distance. By properly choosing the total device length, a bar, cross, or 3-dB coupler state can be selected.

Synchronous couplers can achieve excellent performance. The close match between the modes of the individual waveguides and the lobes of the supermodes leads to low excess loss, and high power transfer efficiency can be achieved. However, since the interaction of the coupled waveguides is weak and hence the difference in propagation constants of the supermodes is small, these devices generally must be very large, typically on the order of millimeters or centimeters. Furthermore, synchronous couplers cannot be implemented with deeply etched rib waveguides, since lateral evanescent fields effectively do not exist. Both of these limitations can be overcome using "zero-gap" or multi-mode waveguide couplers. These devices replace coupled single-mode waveguides with a single multi-mode waveguide, which, like the coupled guides, offers the possibility of exciting symmetric and anti-symmetric modes. Multi-mode couplers offer the prospect of high-performance and compact size and have generated considerable interest in recent years. We now describe the general theory that governs the operation of these multimode resonance couplers.

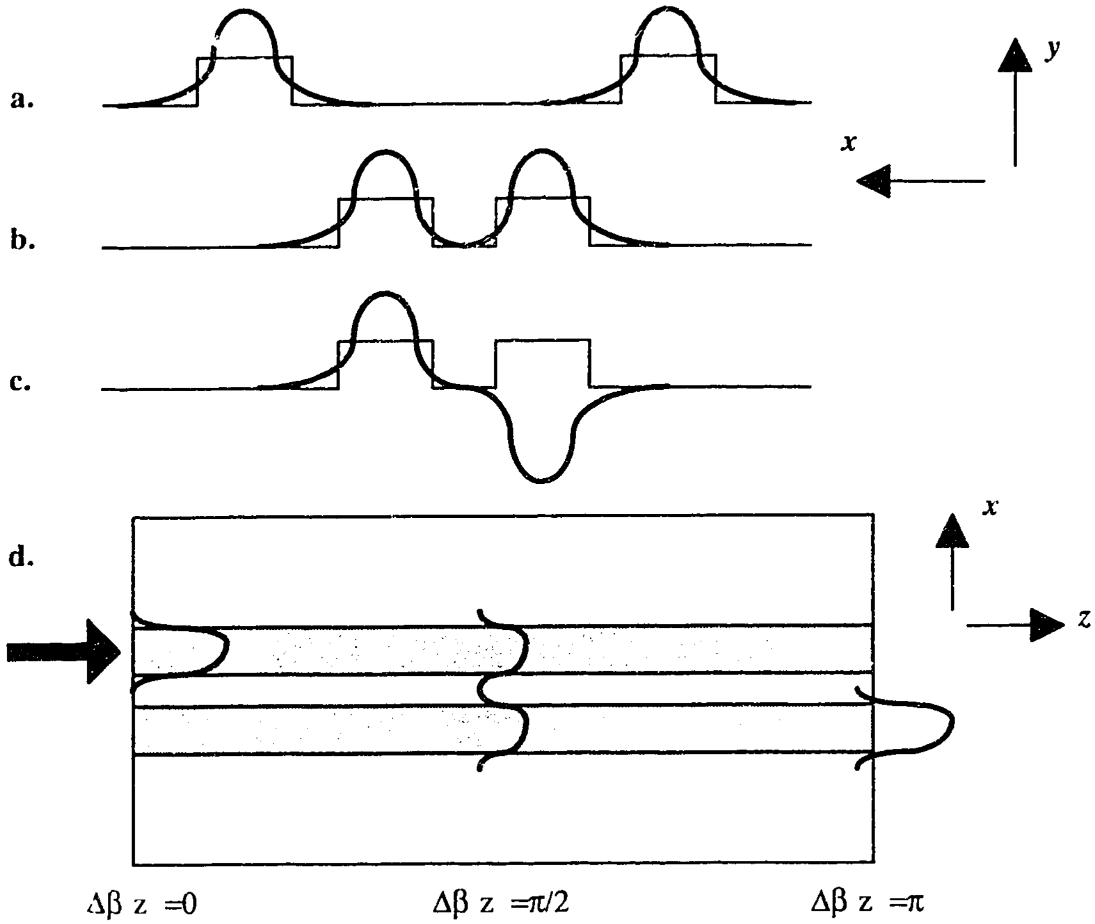


Figure 3.1 Conventional directional coupler. Two waveguides with identical propagation constants, β (a.), are brought close together so that their evanescent mode fields overlap. This lifts the eigenstate degeneracy and creates symmetric (b.) and anti-symmetric (c.) supermodes, with different propagation constants, $\beta_s - \beta_{as} = \Delta\beta > 0$. An input at one guide excites both supermodes (d.), which then interfere as they propagate. The resultant intensity is therefore transferred between the two guides.

3.1 Theory of Resonant Self-Imaging in Multi-Mode Waveguides Multi-mode waveguide couplers operate using a resonant self-imaging process which occurs in homogeneous waveguides [38],[39]. (Homogeneous in this sense refers to a structure wherein a single effective index is sufficient to characterize the "core" region of the waveguide.) These devices can be configured as cross-, bar- and 3-dB couplers, as well as power splitters and combiners, and generally are very compact and have large optical bandwidths and low polarization sensitivities. For these reasons, multi-mode devices are being increasingly applied in integrated photonic devices.

Several theoretical analyses of multi-mode resonance effects have appeared in the literature; the following discussion uses elements of both the general treatment of Bachmann et al. [40] and the consideration presented by Soldano et al. [41]. It is interesting to note that this theory is rigorously correct only in the limit of strong guiding, whereas most practical implementations have used weakly confining waveguide structures.

In a strongly confining structure, the lateral guided mode profiles are well-described by sine functions of the form:

$$\psi_i(x) = \sin\left[\pi(i+1)\frac{x}{W}\right] \qquad \textbf{(3.1)}$$

where i is the mode order and W is the width of the waveguide; the waveguide core extends from x=0 to x=W. The validity of (3.1) is apparent from (2.104.b) and the separation relation (2.96), with the "hard boundary" condition (2.103). For strongly confining rib waveguides, W in (3.1) is the true rib width; for more weakly guiding structures, an "effective" width larger than the true lateral core size can often be defined for application in the following derivations. Following the discussion of strongly confining waveguides in Chapter 2, the functions (3.1) may be directly taken as the electric field x-components of the quasi-TE modes; extension to TM fields will be considered later.

An integral number of half-periods are contained within the waveguide, and the transverse mode propagation constants are therefore:

$$k_{xi} \cong (i+1)\pi / W \qquad \textbf{(3.2)}$$

The longitudinal propagation constants are then given by:

$$\beta_i = n^2 k_o^2 - k_{xi}^2 \tag{3.3}$$

where n in (3.3) is taken as the *effective* index of the waveguide core. For strongly guiding rib structures, this is given by:

$$n = \frac{\sqrt{K_1}}{k_o} \tag{3.4}$$

where K_1 is given by equation (2.104.a). Using the paraxial approximation $k_{xi} << \beta_i \quad \forall i$, we obtain from (3.3):

$$\beta_i \cong nk_o - k_{xi}^2 / 2nk_o \tag{3.5}$$

and therefore, using (3.2), we have:

$$\Delta\beta \equiv \beta_0 - \beta_1 \cong \frac{3\pi^2}{2nk_o W^2} \cong \frac{3\pi\lambda_o}{4nW^2} \tag{3.6.a}$$

$$\beta_o \cong nk_o - \frac{\Delta\beta}{3} \tag{3.6.b}$$

$$\beta_i \cong \beta_0 - i(i+2)\frac{\Delta\beta}{3} \tag{3.6.c}$$

From (3.6.c) we see that the mode propagation constants are a quadratic function of the mode order; this dependence establishes a definite phase relationship among all the modes of a homogeneous waveguide and is the basis of resonant self-imaging. Let us now assume that the waveguide modes are excited by some input field ψ_o at z=0. If this field is intensity-normalized, then the excitation coefficients of the guided modes are, from equations (2.105) and (3.1):

$$a_i = \frac{2}{W} \int_{x=0}^{W} \psi_o(x) \psi_i^*(x) dx \tag{3.7}$$

From (2.11) and (2.47) the guided field distribution in the coupler at a propagation distance z is then:

$$\Psi(x,z)=\sum_{i=0}^{i\max} a_i \psi_i(x)\exp(j\beta_i z) \tag{3.8}$$

where $i$max+1 is the number of guided modes. Let us consider multi-mode waveguides of a fixed length defined by:

$$L_N^M \equiv \frac{M}{N}3L_\pi \tag{3.9}$$

where L_π is a characteristic beat length defined by:

$$L_\pi \equiv \frac{\pi}{\Delta\beta}=\frac{\pi}{\beta_0-\beta_1} \tag{3.10}$$

and, to permit the representation of arbitrary device lengths, M and N are any positive integers without a common divisor. These numbers have a physical significance which will become clear later. We can now use (3.6.c) and (3.8)-(3.10) to write the guided optical field at the coupler output:

$$\Psi\left(x,L_N^M\right)=\sum_{i=0}^{i\max} a_i A_i \psi_i(x) \tag{3.11}$$

where A_i describes the phase of the modes and is given by:

$$A_i=\exp\left[j\beta_0 L_N^M - j\pi\frac{M}{N}i(i+2)\right] \tag{3.12}$$

Using (3.12) and (3.6.a-c), equation (3.11) becomes:

$$\Psi\left(x, L_N^M\right) = \exp\left(j\pi \frac{M}{N}\left\{\frac{3nk_o}{\Delta\beta} - 1\right\}\right)\sum_{i=0}^{i\max} a_i \exp\left(-j\pi \frac{M}{N} i(i+2)\right)\psi_i(x) \qquad \textbf{(3.13)}$$

If we now replace $\psi_i(x)$ with (3.1) and use $\sin x = \left(\exp[j\pi x] - \exp[-j\pi x]\right)/2j$, equation (3.13) becomes:

$$\Psi\left(x, L_N^M\right) = \exp\left(j\pi \frac{M}{N}\left\{\frac{3nk_o}{\Delta\beta} - 1\right\}\right)\frac{1}{2j}\sum_{i=0}^{i\max} a_i \left\{\exp\left(j\pi(i+1)\frac{x}{W} - j\pi\frac{M}{N}i(i+2)\right) - \exp\left(-j\pi(i+1)\frac{x}{W} - j\pi\frac{M}{N}i(i+2)\right)\right\}$$

$$= \exp\left(j\pi \frac{M}{N}\left\{\frac{3nk_o}{\Delta\beta} - 1\right\}\right)\frac{1}{2j}\sum_{i=0}^{i\max} a_i \left\{\exp\left(j\pi\frac{(i+1)x - \frac{M}{N}i(i+2)W}{W}\right) - \exp\left(-j\pi\frac{(i+1)x + \frac{M}{N}i(i+2)W}{W}\right)\right\}$$

$$\textbf{(3.14)}$$

We first examine the case where M=N=1; the coupler length is therefore $3L_\pi$. In this case (3.14) becomes:

$$\Psi\left(x, L_1^1\right) = \exp\left(j\pi\left\{\frac{3nk_o}{\Delta\beta} - 1\right\}\right)\frac{-1}{2j}\sum_{i=0}^{i\max} a_i \left\{\exp\left(j\pi\frac{(i+1)(W-x)}{W}\right)\exp\left(-j\pi\left[i^2+3i+1\right]\right)\right.$$

$$\left. -\exp\left(-j\pi\frac{(i+1)(W-x)}{W}\right)\exp\left(-j\pi\left[i^2+i-1\right]\right)\right\} \qquad \textbf{(3.15)}$$

$$= \exp\left(j\pi\left\{\frac{3nk_o}{\Delta\beta} - 1\right\}\right)\frac{-1}{2j}\sum_{i=0}^{i\max} a_i \exp\left(-j\pi\left[i(i+1)\right]\right)\left\{\exp\left(j\pi\frac{(i+1)(W-x)}{W}\right)\exp\left(-j\pi\left[2i+1\right]\right)\right.$$

$$-\exp\left(-j\pi\frac{(i+1)(W-x)}{W}\right)\exp(j\pi)\Bigg\} \tag{3.16}$$

$$=\exp\left(j\pi\left\{\frac{3nk_o}{\Delta\beta}-1\right\}\right)\frac{1}{2j}\sum_{i=0}^{i\max}a_i\left\{\exp\left(j\pi\frac{(i+1)(W-x)}{W}\right)-\exp\left(-j\pi\frac{(i+1)(W-x)}{W}\right)\right\} \tag{3.17}$$

$$=\exp\left(j\pi\left\{\frac{3nk_o}{\Delta\beta}-1\right\}\right)\sum_{i=0}^{i\max}a_i\sin\left\{\pi\frac{(i+1)(W-x)}{W}\right\} \tag{3.18}$$

$$=\exp\left(j\pi\left\{\frac{3nk_o}{\Delta\beta}-1\right\}\right)\sum_{i=0}^{i\max}a_i\psi_i(W-x) \tag{3.19}$$

Therefore, neglecting the phase pre-factor, we see that at this coupler length an image of the input field, laterally mirrored about the waveguide center, is produced. This effectively corresponds to a cross-coupling resonance. Note that this effect does not depend on the values of the mode excitation coefficients a_i and therefore occurs irrespective of the form of the input field. For this reason, the phenomenon described by the above equations is called a *general* self-imaging resonance. General resonances occur at length multiples of $3L_\pi$ as described by equation (3.9). We shall see later that there are also *restricted* resonances which occur at multiples of L_π if $a_i = 0$ for certain i.

We now examine the case where M=1 and N=2; i.e. the coupler length is $3L_\pi/2$. From (3.14), the guided optical field at the coupler output is:

$$\Psi(x,L_2^1)=\exp\left(j\pi\frac{1}{2}\left\{\frac{3nk_o}{\Delta\beta}-1\right\}\right)\frac{1}{2j}\sum_{i=0}^{i\max}a_i\left\{\exp\left(j\pi\frac{x-\frac{1}{2}i(i+2)W}{W}\right)-\exp\left(-j\pi\frac{(i+1)x+\frac{1}{2}i(i+2)W}{W}\right)\right\} \tag{3.20}$$

$$= \exp\left(j\pi \frac{1}{2}\left\{ \frac{3nk_o}{\Delta\beta} - 1 \right\} \right)\frac{1}{2}\times$$

$$\left[\frac{1}{2j}\sum_{i=0}^{i\max} a_i \exp\left[j\frac{\pi}{2} i(i+2) \right]\left\{ \exp\left(j\pi \frac{(i+1)x - i(i+2)W}{W} \right) - \exp\left(-j\pi \frac{(i+1)x + i(i+2)W}{W} \right) \right\} \right.$$

$$\left. + \frac{1}{2j}\sum_{i=0}^{i\max} a_i \exp\left[-j\frac{\pi}{2} i(i+2) \right]\left\{ \exp\left(j\pi \frac{(i+1)x}{W} \right) - \exp\left(-j\pi \frac{(i+1)x}{W} \right) \right\} \right] \quad \textbf{(3.21)}$$

Using (3.15)-(3.19), we can rewrite (3.21) as follows:

$$\Psi(x, L_2^1) = \exp\left(j\pi \frac{1}{2}\left\{ \frac{3nk_o}{\Delta\beta} - 1 \right\} \right)\left\{ \sum_{i,even} \frac{a_i\psi_i(x) + a_i\psi_i(W-x)}{2} + \sum_{i,odd} \frac{ja_i\psi_i(x) - ja_i\psi_i(W-x)}{2} \right\}$$

$$\textbf{(3.22)}$$

$$= \exp\left(j\pi \frac{1}{2}\left\{ \frac{3nk_o}{\Delta\beta} - 1 \right\} \right)\left\{ \sum_{i} \frac{a_i\psi_i(x) + a_i\psi_i(W-x)}{2} + \sum_{i} \frac{ja_i\psi_i(x) - ja_i\psi_i(W-x)}{2} \right.$$

$$\left. - \sum_{i,odd} \frac{a_i\psi_i(x) + a_i\psi_i(W-x)}{2} - \sum_{i,even} \frac{ja_i\psi_i(x) - ja_i\psi_i(W-x)}{2} \right\} \quad \textbf{(3.23)}$$

The last two terms of (3.23) are zero because of the symmetry properties of the mode fields. Therefore, we have:

$$\Psi(x, L_2^1) = \exp\left(j\pi \frac{1}{2}\left\{ \frac{3nk_o}{\Delta\beta} - 1 \right\} \right)\left\{ \frac{1+j}{2}\sum_{i=0}^{i\max} a_i\psi_i(x) + \frac{1-j}{2}\sum_{i=0}^{i\max} a_i\psi_i(W-x) \right\} \textbf{(3.24)}$$

At this coupler length we thus have a superposition of two images: the input field and its lateral mirror about the waveguide center, each with a magnitude of $1/\sqrt{2}$ that of the input field, and separated in phase by $\pi/2$. This is effectively a 3-dB coupler state.

Finally, we examine the case M=2, N=1. From equation (3.14) the field at the coupler output is:

$$\Psi\left(x,L_1^2\right)=\exp\left(j2\pi\left\{\frac{3nk_o}{\Delta\beta}-1\right\}\right)\frac{1}{2j}\sum_{i=0}^{i\max}a_i\left\{\exp\left(j\pi\frac{(i+1)x-2i(i+2)W}{W}\right)-\exp\left(-j\pi\frac{(i+1)x+2i(i+2)W}{W}\right)\right\}$$

(3.25)

$$=\exp\left(j2\pi\left\{\frac{3nk_o}{\Delta\beta}-1\right\}\right)\frac{1}{2j}\sum_{i=0}^{i\max}a_i\exp[-j\pi 2i(i+2)]\left\{\exp\left(j\pi\frac{(i+1)x}{W}\right)-\exp\left(-j\pi\frac{(i+1)x}{W}\right)\right\}$$

(3.26)

$$=\exp\left(j2\pi\left\{\frac{3nk}{\Delta\beta}-1\right\}\right)\sum_{i=0}^{i\max}a_i\psi_i(x) \qquad \textbf{(3.27)}$$

Therefore, the input field is reproduced at the output: a bar-state resonance. The physical significance of the numbers M and N can now be guessed. In general, a multi-mode waveguide of length L_N^M given by (3.9) will contain, at the output plane, N images of the input field. The locations of these images are partly determined by M; different configurations exist for values of M such that the ratio M/N cannot be reduced. The locations and phases of the N images are derived, for the general case, in [40]. While the presentation given here describes the operation of 2X2 couplers (where bar, cross and 3-dB states are possible), multi-mode waveguides can in fact be used to implement general NXN devices. This possibility has startling ramifications; for example, NXN Mach-Zehnder switches have already been proposed and analyzed [42]. In addition, it has been shown that 1XN imaging is produced by a centered, symmetric excitation such that no odd modes of the coupler are excited [43].

We mentioned earlier the existence of a restricted self-imaging resonance phenomenon; let us now examine this. We postulate that the exciting field at the coupler input is such that:

$$a_i=0 \quad \text{for} \quad \mathrm{mod}_3[i]=2 \qquad \textbf{(3.28)}$$

In other words, modes 2, 5, 8, 11, etc. of the coupler are not excited. In the limit of strong guiding, this condition is in fact easily produced. From (3.1) and (3.7) we find that (3.28) is satisfied by a symmetric field distribution $\psi_o(x)$ centered at $x=W/3$ or $x=2W/3$, since all the "forbidden" mode fields are anti-symmetric about these positions. We now consider a coupler length defined by:

$$L_N^M = \frac{M}{N} L_\pi \tag{3.29}$$

Therefore from (3.8) the guided field profile at the coupler output is:

$$\Psi(x, L_N^M) = \exp\left(j\pi \frac{M}{N} \left\{ \frac{nk_o}{\Delta\beta} - \frac{1}{3} \right\} \right) \frac{1}{2j} \sum_{i=0}^{i\max} a_i \left\{ \exp\left(j\pi \frac{(i+1)x - \frac{M}{N} \frac{i(i+2)W}{3}}{W} \right) - \exp\left(-j\pi \frac{(i+1)x + \frac{M}{N} \frac{i(i+2)W}{3}}{W} \right) \right\} \tag{3.30}$$

We now decompose the summand of (3.30) into two parts using a new sum index:

$$\Psi(x, L_N^M) = \exp\left(j\pi \frac{M}{N} \left\{ \frac{nk_o}{\Delta\beta} - \frac{1}{3} \right\} \right) \frac{1}{2j} \sum_{p=0}^{S} \left[a_{3p} \left\{ \exp\left(j\pi \frac{(3p+1)x - \frac{M}{N} p(3p+2)W}{W} \right) - \exp\left(-j\pi \frac{(3p+1)x + \frac{M}{N} p(3p+2)W}{W} \right) \right\} \right.$$

$$\left. + a_{3p+1} \left\{ \exp\left(j\pi \frac{(3p+2)x - \frac{M}{N}(3p+1)(p+1)W}{W} \right) - \exp\left(-j\pi \frac{(3p+2)x + \frac{M}{N}(3p+1)(p+1)W}{W} \right) \right\} \right] \tag{3.31}$$

where:

$$S \equiv \operatorname{int}\left(\frac{i\max}{3} \right) \tag{3.32}$$

The summations over $3p$ and $3p+1$ in (3.31) include all the guided modes of the coupler, except the forbidden modes defined by equation (3.28). Let us now consider the case where $M=N=1$. Equation (3.31) becomes:

$$\Psi(x, L_1') = \exp\left(j\pi\left\{\frac{nk_o}{\Delta\beta}-\frac{1}{3}\right\}\right)\frac{1}{2j}\sum_{p=0}^{S}\left[a_{3p}\left\{\exp\left(j\pi\frac{(3p+1)x-p(3p+2)W}{W}\right)-\exp\left(-j\pi\frac{(3p+1)x+p(3p+2)W}{W}\right)\right\}\right.$$

$$\left.+a_{3p+1}\left\{\exp\left(j\pi\frac{(3p+2)x-(3p+1)(p+1)W}{W}\right)-\exp\left(-j\pi\frac{(3p+2)x+(3p+1)(p+1)W}{W}\right)\right\}\right]$$

(3.33)

$$=\exp\left(j\pi\left\{\frac{nk_o}{\Delta\beta}-\frac{1}{3}\right\}\right)\frac{-1}{2j}\sum_{p=0}^{S}\left[a_{3p}\left\{\exp\left(j\pi\frac{(3p+1)(W-x)}{W}\right)\exp\left(-j\pi\left[3p^2+5p+1\right]\right)-\exp\left(-j\pi\frac{(3p+1)(W-x)}{W}\right)\exp\left(-j\pi\left[3p^2-p-1\right]\right)\right\}\right.$$

$$\left.+a_{3p+1}\left\{\exp\left(j\pi\frac{(3p+2)(W-x)}{W}\right)\exp\left(-j\pi\left[3p^2+5p+1\right]\right)-\exp\left(-j\pi\frac{(3p+2)(W-x)}{W}\right)\exp\left(-j\pi\left[3p^2+p-1\right]\right)\right\}\right]$$

(3.34)

$$=\exp\left(j\pi\left\{\frac{nk_o}{\Delta\beta}-\frac{1}{3}\right\}\right)\frac{-1}{2j}\sum_{p=0}^{S}\exp\left[-j\pi\left(3p^2+5p+1\right)\right]\left[a_{3p}\left\{\exp\left(j\pi\frac{(3p+1)(W-x)}{W}\right)-\exp\left(-j\pi\frac{(3p+1)(W-x)}{W}\right)\exp\left(j2\pi[3p+1]\right)\right\}\right.$$

$$\left.+a_{3p+1}\left\{\exp\left(j\pi\frac{(3p+2)(W-x)}{W}\right)-\exp\left(-j\pi\frac{(3p+2)(W-x)}{W}\right)\exp\left(j2\pi[2p+1]\right)\right\}\right] \quad \textbf{(3.35)}$$

$$=\exp\left(j\pi\left\{\frac{nk_o}{\Delta\beta}-\frac{1}{3}\right\}\right)\frac{1}{2j}\sum_{p=0}^{S}\left[a_{3p}\left\{\exp\left(j\pi\frac{(3p+1)(W-x)}{W}\right)-\exp\left(-j\pi\frac{(3p+1)(W-x)}{W}\right)\right\}\right.$$

$$\left.+a_{3p+1}\left\{\exp\left(j\pi\frac{(3p+2)(W-x)}{W}\right)-\exp\left(-j\pi\frac{(3p+2)(W-x)}{W}\right)\right\}\right] \quad \textbf{(3.36)}$$

$$=\exp\left(j\pi\left\{\frac{nk_o}{\Delta\beta}-\frac{1}{3}\right\}\right)\sum_{p=0}^{S}\left[a_{3p}\Psi_{3p}(W-x)+a_{3p+1}\Psi_{3p+1}(W-x)\right] \quad \textbf{(3.37)}$$

Neglecting the phase pre-factor, equation (3.37) is simply a superposition of all of the excited modes of the coupler, laterally mirrored about the waveguide center. Therefore, as in the general resonance case for $M=N=1$, a cross-coupled state is achieved, only now the required device length is reduced by a factor of three. Similar derivations can be followed for $M=1$, $N=2$ and $M=2$, $N=1$; in these cases, 3-dB and bar-coupler states, respectively, are obtained.

The various cases of resonant self-imaging in multi-mode waveguides, such that $M, N \leq 2$ (i.e. describing a 2X2 coupler) are summarized in Table 3.1.

3.2 Polarization Dependence of Resonant Self-Imaging The analysis of resonant self-imaging in homogeneous multi-mode waveguides presented above is valid for quasi-TE fields in strongly confining waveguides. We discuss here the applicability of the preceding derivations to quasi-TM fields.

In general, the propagation constants of TM modes are not the same as those of TE modes. However, resonant self-imaging is not a function of the *absolute* values of the mode propagation constants, but rather depends only on their *relative* values. Therefore, presuming that (3.1) is a reasonable description of the lateral TM guided mode fields, if:

$$\beta_i^{TM} - \beta_{i-1}^{TM} = \beta_i^{TE} - \beta_{i-1}^{TE} \quad \forall i, \ i > 0 \qquad \textbf{(3.38)}$$

then the resonant self-imaging described in the preceding section will be polarization-independent. The condition (3.38) is in fact obtained in the limit of strong guiding; this can be appreciated by considering the discussion of section 2.9. In the limit of strong guiding, the quasi-TE modes are directly given by a scalar analysis, according to equation (2.94). The quasi-TM mode propagation constants can be obtained, from the scalar values, by a similar equation involving the integration of (2.95) over all horizontal layer interfaces. In the limit of strong guiding, where the separation (2.96) is valid, the vertical field profile is the same for all lateral mode orders and hence the resulting correction to the propagation constants is the same for all guided modes. Equation (3.38) is therefore satisfied in this case. The polarization-independence of resonant self-imaging in the limit of strong guiding has been mentioned by other authors [44] without intuitive justification.

In practice, polarization-independent operation is more nearly obtained as the number of coupler modes increases; this is easy to understand from the preceding arguments. The highest order modes of a waveguide are the least laterally confined and the separation (2.96) may be less valid than for the better-confined low order modes. The scalar-to-TM corrections to the high-order mode propagation constants therefore may not be uniform. If a coupler sustains only two modes, then the propagation constant correction is apt to be different for the fundamental and first-order modes, making the beat length defined by (3.10) polarization-sensitive. If many modes are sustained, then the scalar-to-TM corrections to the propagation constants of the two lowest-order modes will be closer in value, leading to a smaller difference in beat lengths.

Resonance Type	**Coupler Length, q=0, 1, 2, . . .**	**Coupler State**	**Required Input Condition**
General	$(2q)(3L_\pi)$	Bar	None
General	$(2q+1)(3L_\pi)$	Cross	None
General	$\left(q+\frac{1}{2}\right)(3L_\pi)$	3-dB	None
Restricted	$(2q)L_\pi$	Bar	$a_i=0, \quad \mathrm{mod}_3[i]=2$
Restricted	$(2q+1)L_\pi$	Cross	$a_i=0, \quad \mathrm{mod}_3[i]=2$
Restricted	$\left(q+\frac{1}{2}\right)L_\pi$	3-dB	$a_i=0, \quad \mathrm{mod}_3[i]=2$

Table 3.1 Self-imaging resonance types, and corresponding device lengths, [illegible] implementation of 2X2 couplers.

3.3 Coupler Implementation Using Resonant Self-Imaging: Operational Considerations

The performance parameters of a 2X2 coupler are contrast (imbalance) and insertion loss, usually stated in decibels, and defined as follows:

$$\text{Contrast (Imbalance)} = 10\log_{10}\frac{P_1}{P_2} \text{ (dB)} \qquad \textbf{(3.39.a)}$$

$$\text{Insertion Loss} = -10\log_{10}\frac{P_1}{P_{in}} \text{ (dB)} \qquad \textbf{(3.39.b)}$$

where P_1 and P_2 are the guided optical powers in the two output waveguides, and P_{in} is the power in the input waveguide, as shown in Figure 2.2. In the bar or cross states, a high value of (3.39.a) is desired, and the parameter is referred to as contrast. In the 3-dB state, a zero value of (3.39.a) is sought, so in this case the parameter is called imbalance.

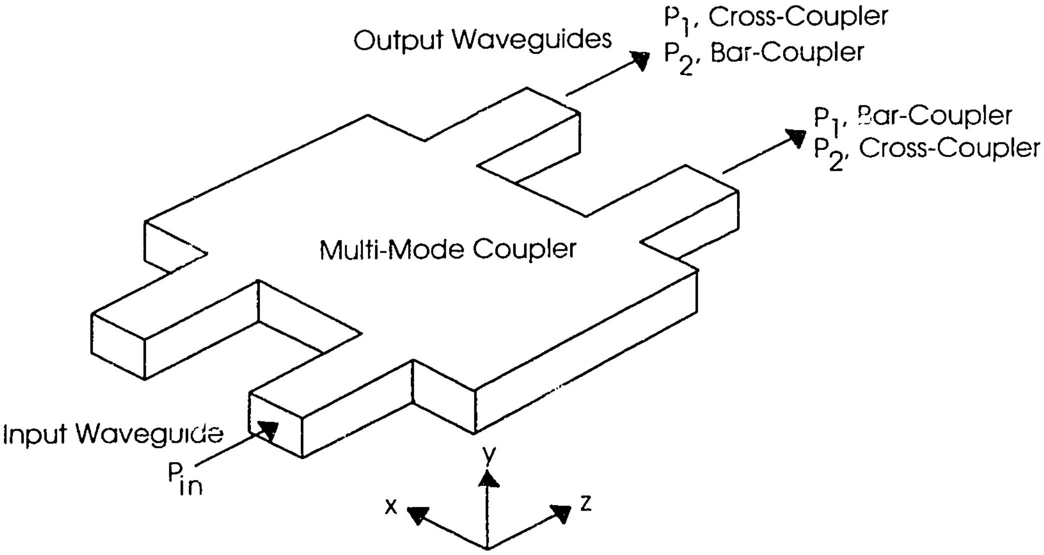


Figure 3.2 Schematic of a 2X2 multi-mode coupler.

For a 3-dB coupler, either output may be designated as 1 or 2 for the purposes of calculating imbalance, while P_1 in (3.39.b) should be replaced with $P_1 + P_2$.

Achieving high contrast (low imbalance) and low insertion loss are important considerations in the realization of multi-mode couplers. In addition, we identify two other concerns: device compactness, and fabrication feasibility/reproducibility. We therefore consider a total of four operational factors which, to some extent, are not simultaneously optimizable. For example, from (3.6.a) and (3.10) we note that:

$$L_\pi \cong \frac{4nW^2}{3\lambda_o} \tag{3.40}$$

The beat length is proportional to the square of the coupler width. Compactness is therefore promoted by making *W* small, but using a narrow coupler has a negative impact on both insertion loss and reproducibility. A narrow coupler supports relatively few guided modes, and therefore few terms are available for the series expansion of the input field. This potentially leads to large insertion losses. Furthermore, for a narrow coupler, the access waveguides at the input and output may have to be offset, at an angle to the coupler propagation axis, to promote repeatable fabrication. This is because a separation of parallel waveguides of less than ~1 μm cannot be reliably defined by conventional photolithographic pattern transfer, leading to some "filling in" of the space between the waveguides. Making the access waveguides diverge at an angle reduces this filling in, or at least restricts it to some relatively repeatable length; see Figure 3.3. However, setting the access waveguides at an angle reduces the efficiency of power transfer to the guided modes of the coupler, for reasons that will be made clear shortly. Again, this leads to an insertion loss penalty.

Early multi-mode couplers employed narrow waveguides sustaining only two modes, and suffered from the fabrication difficulties and insertion loss penalties expected from the arguments above [45],[46]. In recent years, as the principles of resonant self-imaging in multi-mode waveguides have become better understood, the tendency has been to use wider couplers which sustain a large number of modes. This allows the access waveguides to be widely separated and therefore eliminates the need for an angular offset of these waveguides. With no junction angle and many coupler modes, much reduced insertion losses can be obtained; values on the order of 0.5 dB are now typical for well-designed devices [41],[47].

However, as mentioned earlier, the required length of a coupler increases rapidly as it is made wider. Therefore, if compactness is a concern, some optimization of performance versus coupler width may be sought. In this regard we make two observations. First, the insertion loss of a coupler is not strictly a simple function of the number of guided modes sustained by the coupler; the width and location of the input and output waveguides has some effect. For example, a coupler with convergent access waveguides and sustaining two modes may achieve more efficient coupling, and lower losses, than a wider coupler, sustaining three modes, but with access waveguides located at 1/3 and 2/3 of the coupler width so that the second-order mode is not excited. Furthermore, although a larger number of modes leads *generally* to more efficient coupling, beyond a certain point, adding further modes will not have a significant impact. If a coupler sustains three guided modes, and these are able to capture 96% of the optical power from an input waveguide, then making the coupler infinitely wide to capture the remaining 4% reduces the total insertion loss of the device by only 0.35 dB.

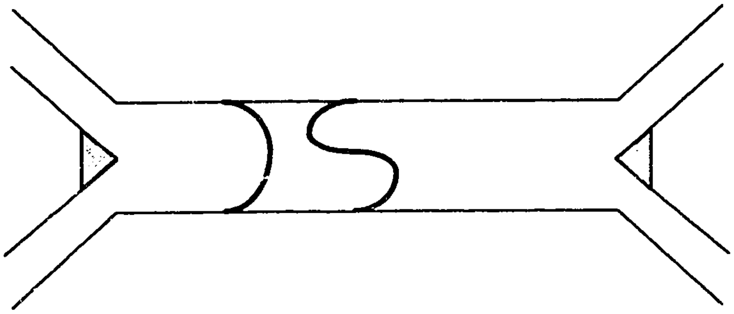


Figure 3.3 Typical narrow multi-mode coupler configuration. The shaded regions indicate irreproducible "fill-in" due to finite photolithographic resolution. This fill-in can reduced by using large junction angles, at the expense of increased coupling loss at the access waveguide-coupler junctions.

Second, the impact of using angled access waveguides in conjunction with multi-mode, as opposed to two-mode, couplers has not yet been quantitatively evaluated, even though the insertion loss penalty with increased junction angle falls rapidly as more coupler modes are sustained. Therefore, using large junction angles, in conjunction with a relatively narrow coupler sustaining more than two modes, and operating on the general resonance where the distribution of power among the waveguide modes is irrelevant, may allow the fabrication of relatively compact couplers. This is the aim of the work presented in this chapter: a theoretical and experimental evaluation of multi-mode couplers with such a configuration.

3.4 Numerical Modeling of Multi-Mode Couplers Following the discussion of Section 2.12, we use a modal analysis to model the performance of multi-mode coupler devices. The coupler geometry illustrated in Figure 3.4 is assumed. The field at the input junction plane is constructed using the mode field(s) of the input waveguides, with specified relative amplitudes and phases. Since strong guiding is assumed, the possibility of field coupling between the access waveguides can be ignored. A specified separation of the access waveguides is used in the construction of the input field, and some branching angle of the access waveguides is taken into account as follows. As was pointed out in section 2.12, in a modal analysis the modes in each section of waveguide can be defined with respect to a local coordinate system. Here, we calculate the guided modes of the straight access waveguides and use a coordinate rotation transformation to bring the mode fields of the access waveguides and the coupler into a uniform system of coordinates for calculating the amplitudes of the coupler modes. Assume that the (lateral) mode fields of one of the access waveguides are defined with respect to coordinates $(\hat{x},\hat{z})$. At the coupler input/output, the equivalent fields expressed in the normal coordinates (x,z) of the coupler are obtained by rotating by $\varphi=(\pm)\theta/2$, such that the off-axis displacement

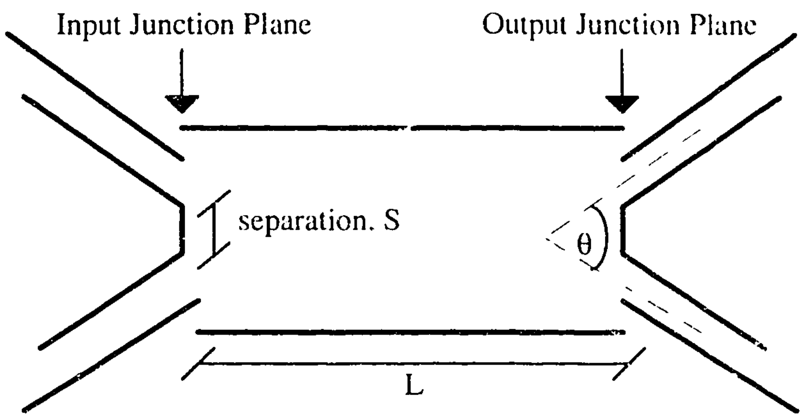


Figure 3.4 Assumed coupler geometry for computer model. The separation and full angle between the access waveguides are separately defined for the input and output.

of the waveguide is assumed to be half of the total branching angle θ. See Figure 3.5 below.

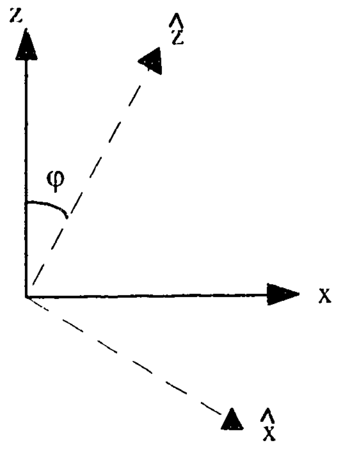


Figure 3.5 Coordinate rotation transformation for calculation of mode fields of access waveguides at coupler junction planes.

The coordinate transformation is given by:

$$\begin{cases} \hat{x} = x\cos\varphi - z\sin\varphi \\ \hat{z} = z\cos\varphi + x\sin\varphi \end{cases} \tag{3.41}$$

The mode fields of the access waveguide are defined by:

$$\psi_i(\hat{x})e^{j\beta_i\hat{z}} \tag{3.42}$$

and therefore they become:

$$\psi_i(x\cos\varphi - z\sin\varphi)e^{j\beta_i(z\cos\varphi + x\sin\varphi)} \tag{3.43}$$

Setting z=0 at the junction plane, the transformed mode fields are, finally:

$$\psi_i(x\cos\varphi)e^{j\beta_i x\sin\varphi} \tag{3.44}$$

From (3.44) we can understand why increasing φ leads to increased coupling to high-order coupler modes. We rewrite (3.44) as follows:

$$\psi_i(x\cos\varphi)e^{j\beta_i} \qquad = \psi_i(x\cos\varphi)\cos[\beta_i x\sin\varphi] + j\psi_i(x\cos\varphi)\sin[\beta_i x\sin\varphi] \tag{3.45}$$

The transformed field therefore has both symmetric and anti-symmetric components which are used in the calculation of the coupler mode amplitudes. Furthermore, these components have an enhanced periodicity (increased spatial frequency) determined by the product $\beta_i \sin\varphi$. The overlap with low-order modes, which have low-order periodicity, is thus reduced.

The modes of the coupler and access waveguides are determined by the effective index method described in section 2.10. Equation (2.105) is used to determine the amplitudes of the coupler modes, where the input modes are transformed using (3.44). The guided optical field at the coupler output is constructed using:

$$\Psi'(x,\quad) = \sum_i a_i \psi_i(x) e^{j\beta_i L} \tag{3.46}$$

The amplitudes of the modes of the output waveguides are determined using (3.46) and (2.105), where the output waveguide modes are also appropriately transformed using (3.44). If the input waveguide fields are normalized, then the amplitude coefficients of the coupler modes and the output waveguide modes are sufficient to calculate the performance parameters of the device. Figure 3.6 shows a flow-chart of the computer model, which was implemented in *Mathematica*.

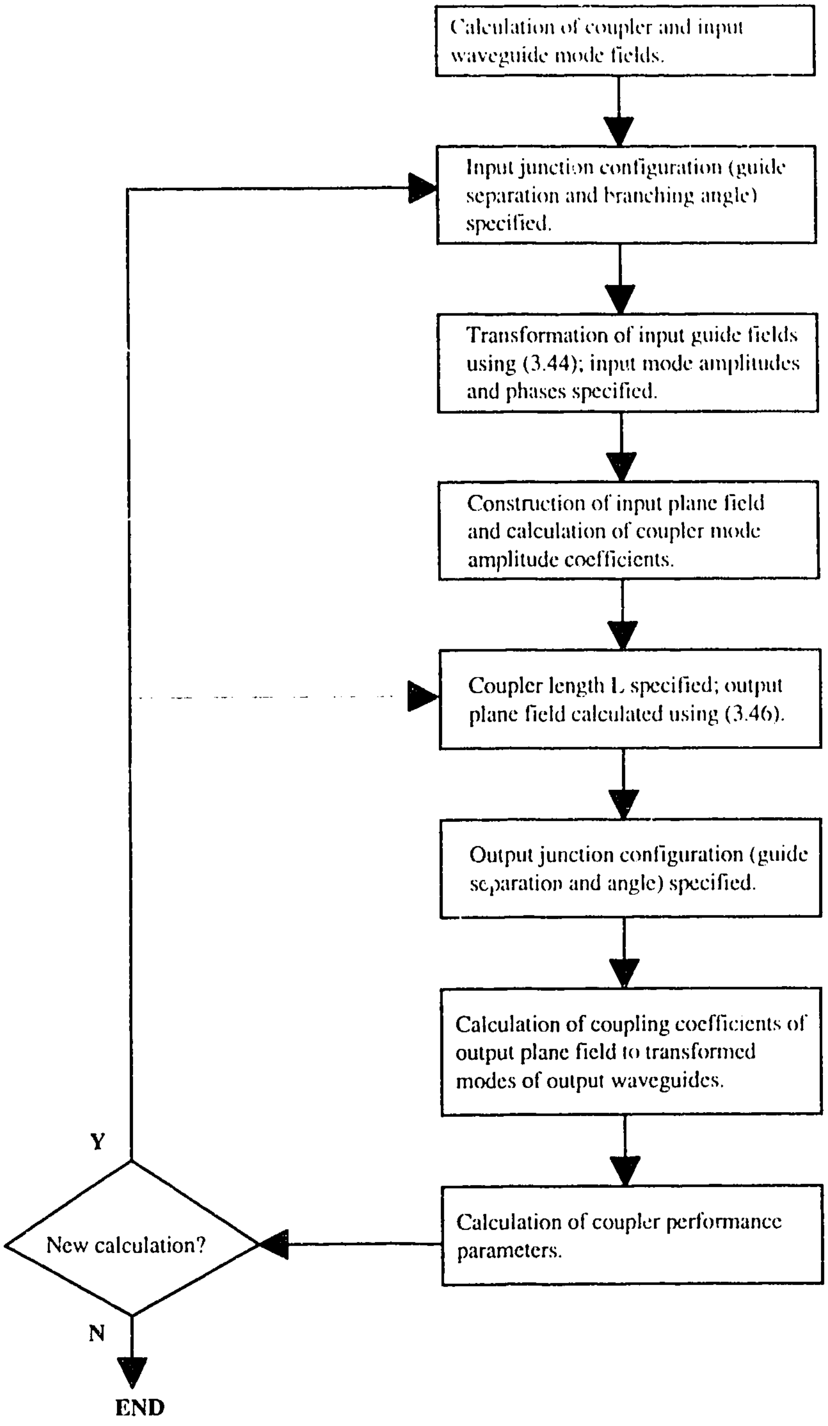


Figure 3.6 Algorithm flow chart for computer model.

3.5 Application: Mach-Zehnder Switch Implementation An important application of waveguide couplers is in electro-optic switches. All electro-optic switches are essentially cross-couplers, either multi-mode or coupled waveguides, with provision for phase modulation to modify the guided mode profiles or the phase situation in the coupler, thereby achieving switching to the bar state. Conventional directional coupler ("COBRA") switches use coupled waveguides to achieve cross-coupling in the passive state, and induce an index change in one guide to destroy the symmetry of the supermodes and reduce or eliminate the cross-coupled intensity [48]. A more efficient device, from an index modulation efficiency standpoint, is the Mach-Zehnder switch [49]. This device is essentially a cross-coupler that is split into two sections, between which is inserted a Mach-Zehnder modulator; see Figure 3.7. At one-half the cross-coupling length, the couplers at the input and output operate in the 3-dB configuration. Light from a single input is therefore evenly divided between the two arms of the Mach-Zehnder, and separated in phase by a quarter-wave. If no phase modulation occurs, then the second 3-dB coupler oper [illegible] ciprocally and directs the full intensity to the cross-port. Looked at alternatively, with no phase modulation, the device is simply a cross-coupler. However, when a π phase delay is induced in one arm of the modulator, the phase situation at the input of the second coupler is reversed, and the device switches to the bar state. The required phase modulation is less than the $\sqrt{3}\pi$ required for the COBRA switch.

The Mach-Zehnder switch is one device where all of the advantages of strongly guiding waveguides can work to advantage. The possibility of high field confinement in the active region of the guide offers the prospect of small modulation voltages and short electrodes, which not only reduce the overall length of the switch but obviate the need for traveling-wave electrodes and high-voltage RF drivers. Compactness is also enhanced by the feasibility of small bend radii in the various access guides. Further size optimization may then be achieved by reduction of the coupler length.

In the present work, coupler design was considered for use in a Mach-Zehnder switch, implemented in an InP/InGaAsP structure [50], shown in Figure 3.8. Multiple quantum wells are used to achieve efficient phase modulation; the physics of this mechanism is considered in Chapter 5. The lower InGaAs layer is included to absorb stray light and hence attenuate any vertically leaky modes [51]. All waveguides are formed by reactive ion etching using CH_4-H_2 to a depth of 3 μm and therefore are strongly guiding, allowing tight waveguide bends. With single-mode waveguides 2 μm wide and TE-polarized light, the voltage-length product for π phase modulation is 2.7 V-mm, while that for $\pi/2$ modulation is 1.8 V-mm. Modulation of TM light requires much higher voltages and is not considered. Small signal modulation 3-dB electrical bandwidths of up to 15.5 GHz have been achieved in this structure without traveling-wave electrodes. Two types of switch configurations are implemented: the conventional, four-port Mach-Zehnder arrangement described above, and a three-port device using a 1X2 multi-mode coupler at the input; see Figure 3.9. The latter arrangement can be made more compact as the required input coupler length is $3ML_{\pi}/8$, as given in [43] and [52], rather than $3ML_{\pi}/2$. In this case, the propagating modes in the modulator arms are in phase, and hence $\pm\pi/2$ phase modulation is required to achieve output selection.

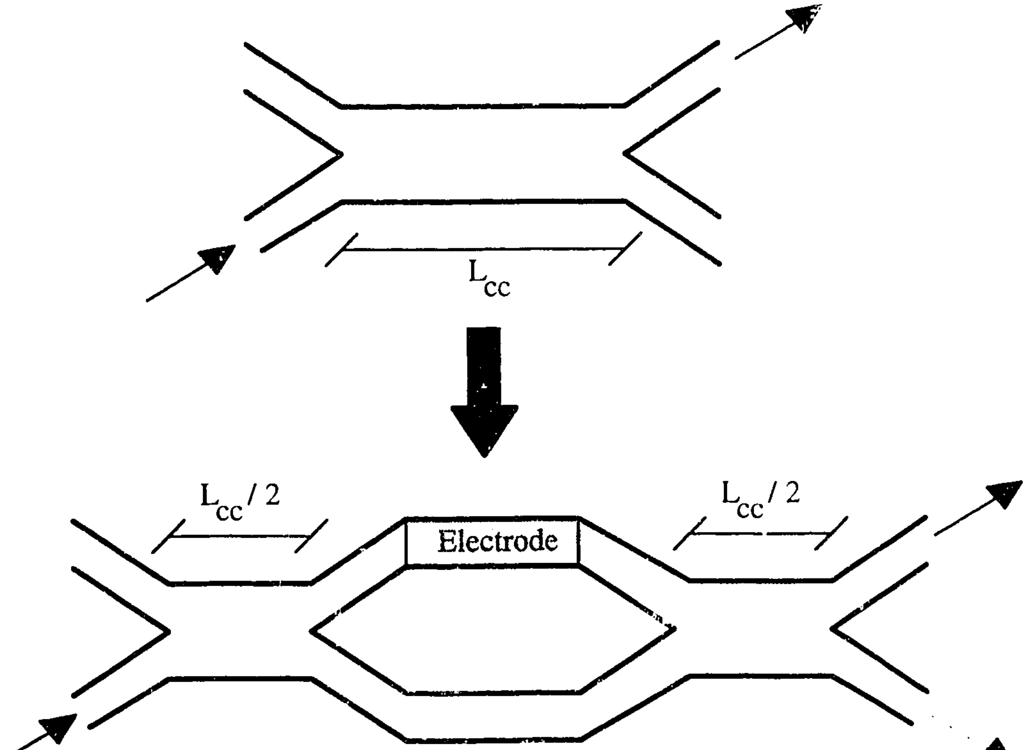


Figure 3.7 Mach-Zehnder electro-optic switch. A cross-coupler is modified to include a Mach-Zehnder modulator, which achieves switching to the bar state by reversing the phase situation at the input to the output 3-dB coupler.

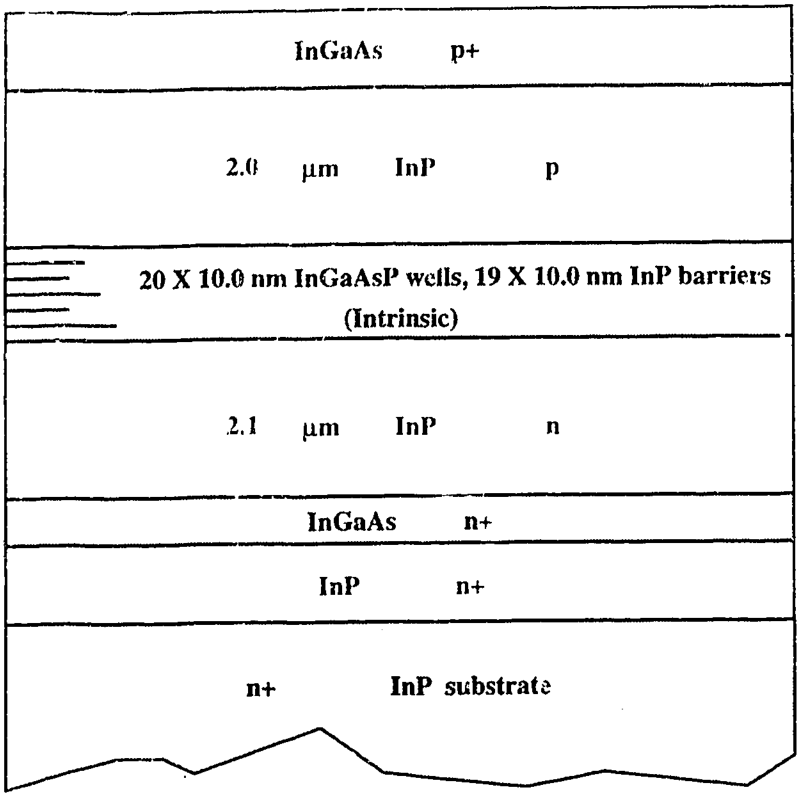


Figure 3.8 Wafer structure for Mach-Zehnder switch device. Layer thicknesses are not drawn to scale.

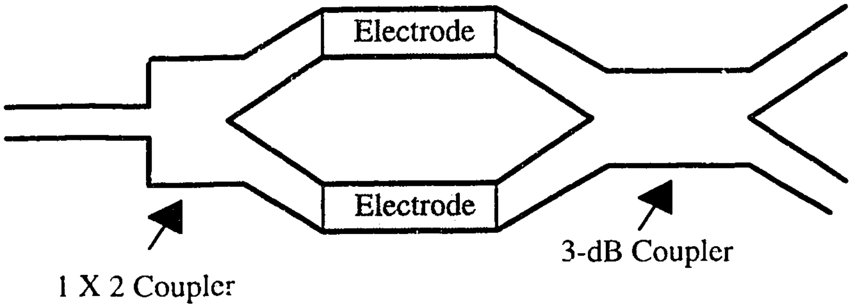


Figure 3.9 Schematic of three-port Mach-Zehnder switch configuration.

The calculation of mode profiles requires a knowledge of the refractive index distribution in the structure. Refractive index is a function of material composition and wavelength; the appropriate expressions for the lattice-matched material system $In_uGa_{1-u}As_{1-v}P_v$ are available in the literature [53]-[55]. Lattice matching requires that $u \cong 0.53+0.47v$. For photon energies below the material band edge, the refractive index is given by:

$$n(\lambda_o) = \sqrt{A(v)\left\{f(\chi)+\frac{1}{2}\left[E_o/(E_o+\Delta_o)\right]^{3/2} f(\chi_{so})\right\}+B(v)} \qquad \textbf{(3.46.a)}$$

where:

$$f(\chi)=\chi^{-2}\left[2-(1+\chi)^{1/2}-(1-\chi)^{1/2}\right], \qquad \textbf{(3.46.b)}$$

$$\chi=\frac{hc_o}{\lambda_o qE_o}, \qquad \textbf{(3.46.c)}$$

$$\chi_{so}=\frac{hc_o}{\lambda_o q(E_o+\Delta_o)}, \qquad \textbf{(3.46.d)}$$

$$E_o=1.35-0.72(1-v)+0.12(1-v)^2,\ (\mathrm{eV}) \qquad \textbf{(3.46.e)}$$

$$E_o+\Delta_o=1.038+0.299(1-v)+0.129(1-v)^2,\ (\mathrm{eV}) \qquad \textbf{(3.46.f)}$$

$$A(v)=8.40-3.40v, \qquad \textbf{(3.46.g)}$$

and:

$$B(v)=6.60+3.40v; \qquad \textbf{(3.46.h)}$$

where h is Planck's constant. The quantum well stack may be considered a single layer for purposes of refractive index specification, with an index approximated by:

$$n = \sqrt{\frac{n_w^2 t_w + n_b^2 t_b}{t_w + t_b}} \qquad (3.47)$$

for TE fields, where n_w, n_b, t_w and t_b are the refractive indices and thicknesses of the wells and barriers, respectively [56]. The quantum well composition is $In_{0.91}Ga_{0.09}As_{0.19}P_{0.81}$ for a photoluminescence peak at 1.45 μm to provide strong electrorefraction at 1.55 μm. The refractive indices for the InP layers and the quantum well stack are calculated as 3.166 and 3.369, respectively, at 1.55 μm.

As a preliminary step, we check the validity of the separation relation (2.96). Two-dimensional mode profiles are calculated using the scalar finite element formulation described in section 2.11. As the index of refraction of InGaAs is complex at 1.55 μm, these layers are taken as InP for calculation purposes; an acceptable approximation since these layers are thin. The vertical field profiles at various lateral points, for various rib widths, are factored from the two-dimensional mode fields. In Figure 3.10, the vertical mode profiles *G(y)*, normalized to unit peak magnitude, are shown for nine different cases: at 10%, 50% and 90% of the rib width for the fundamental and first-order modes of a 4 μm-wide rib, and the second-order mode of a 6 μm-wide rib. The *y*-distance in Figure 3.10 runs from the bottom of the structure up, and the InP thickness beneath the quantum wells has been taken as 2.3 μm. The center of the MQW stack is therefore at ~2.5 μm, and the air boundary at the top of the rib is at ~4.9 μm.

The nine curves in Figure 3.10 are almost exactly superimposed, indicating that the separation relation (2.96) is valid, to a good approximation, in this case. From the two-dimensional mode fields, the lateral field dependence can be obtained by taking the square root of the vertically integrated mode intensity, as follows:

$$\sqrt{\int_{-\infty}^{\infty} |\psi_\upsilon(x,y)|^2 dy} = |F_\upsilon(x)| \sqrt{\int_{-\infty}^{\infty} |G(y)|^2 dy} \qquad (3.48)$$

The lateral field *F(x)* for a 2 μm-wide rib (centered at x=2 μm), obtained from the two-dimensional mode using (3.48), and normalized to unit peak magnitude, is shown in Figure 3.11. The mode profile calculated using equations (2.104.a-c) in section 2.10 is not shown but is virtually indistinguishable from the curve in Figure 3.11, indicating the validity of the (modified) effective index method.

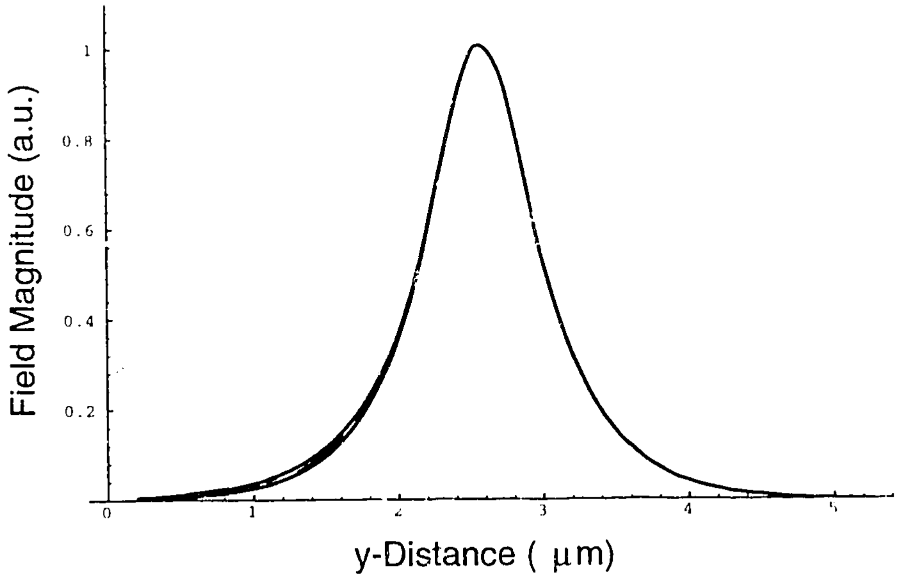


Figure 3.10 Normalized fields *G(y)* taken at 10%, 50% and 90% of rib width for fundamental and first-order modes of 4 μm-wide rib, and second-order mode of 6 μm-wide rib.

Note further the very limited evanescent fields outside the rib boundaries; two of these guides can be brought very close together without significant field coupling. This is confirmed by solving for the modes of a system of two 2 μm-wide waveguides wherein the ribs are separated by progressively smaller distances. The mode degeneracy is not lifted until the separation is ~0.6 μm; since the limit of conventional photolithographic resolution is on the order of 0.5 μm, we can effectively design without concern for inter-guide field coupling.

3.6 Modeled Results and Coupler Design We now present modeled results for 3-dB couplers implemented as components of the Mach-Zehnder switch configurations described earlier. We first consider the three-port device of Figure 3.9. In this case, we model a 2X2 coupler, for implementation at the output of the device. We assume that equal optical intensities are incident in each input waveguide,

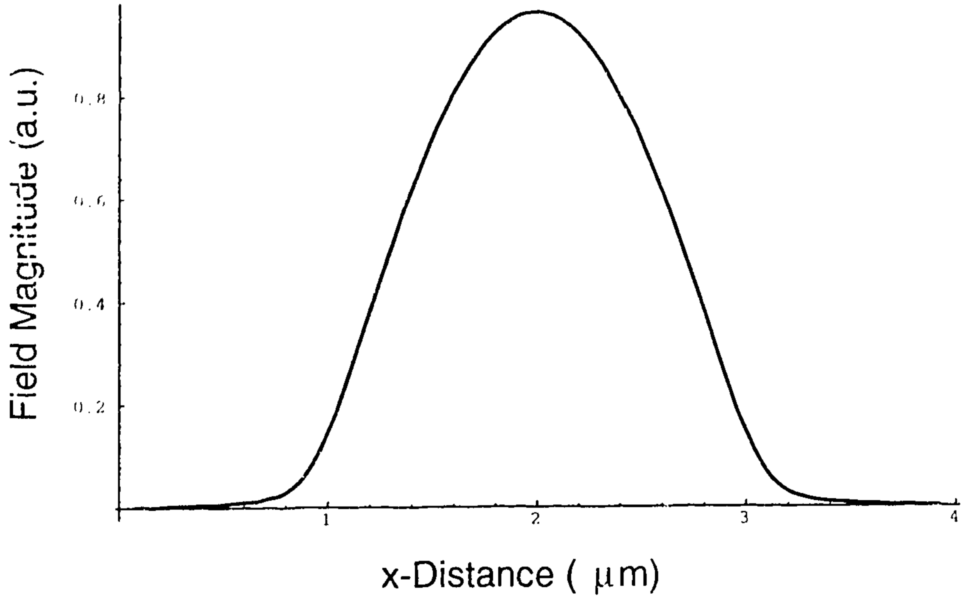


Figure 3.11 Compressed (x-dependent) mode field for a 2 μm rib.

and that these input fields are separated in phase by a quarter-wave. This arrangement is expected to produce maximum contrast when the output coupler is in a true 3-dB state. Using (2.57) and the definitions (3.39.a&b), the performance parameters of the coupler are given by:

$$\text{Contrast} = 10\log_{10}\left\{\frac{P_1}{P_2}\right\} = 10\log_{10}\left\{\frac{|a_{of\to1}|^2}{|a_{of\to2}|^2}\right\} \text{ (dB)}, \qquad \textbf{(3.49.a)}$$

$$\text{Insertion Loss} = -10\log_{10}\left\{\frac{P_1}{P_{in}}\right\} = -10\log_{10}\left\{|a_{of\to1}|^2 \times \sum_{i,\ \text{coupler}} |a_{if\to i}|^2\right\} \text{ (dB)}\textbf{(3.49.b)}$$

where the subscript $of \to 1,2$ denotes the amplitudes of the offset and angled modes of outputs 1,2 excited by the field at the coupler output, and $if \to i$ indicates the amplitude of coupler mode *i* excited by the input plane field, constructed from the offset, angled and phase-shifted modes of the input waveguides.

Our goal is to design the most compact coupler consistent with acceptable performance parameters and realizable fabrication. To reduce potential insertion losses to an acceptable level, the coupler must sustain at least three modes, while compactness dictates using a small width. A 4 μm-wide waveguide supports three modes, and is only twice the width of the single-mode access waveguides, a practical minimum. For these reasons we take 4 μm as our nominal coupler width. We first examine the artificial case of zero separation and zero angle between the access waveguides, at both the input and output junctions. Figure 3.12 shows the modeled performance parameters as a function of the coupler length.

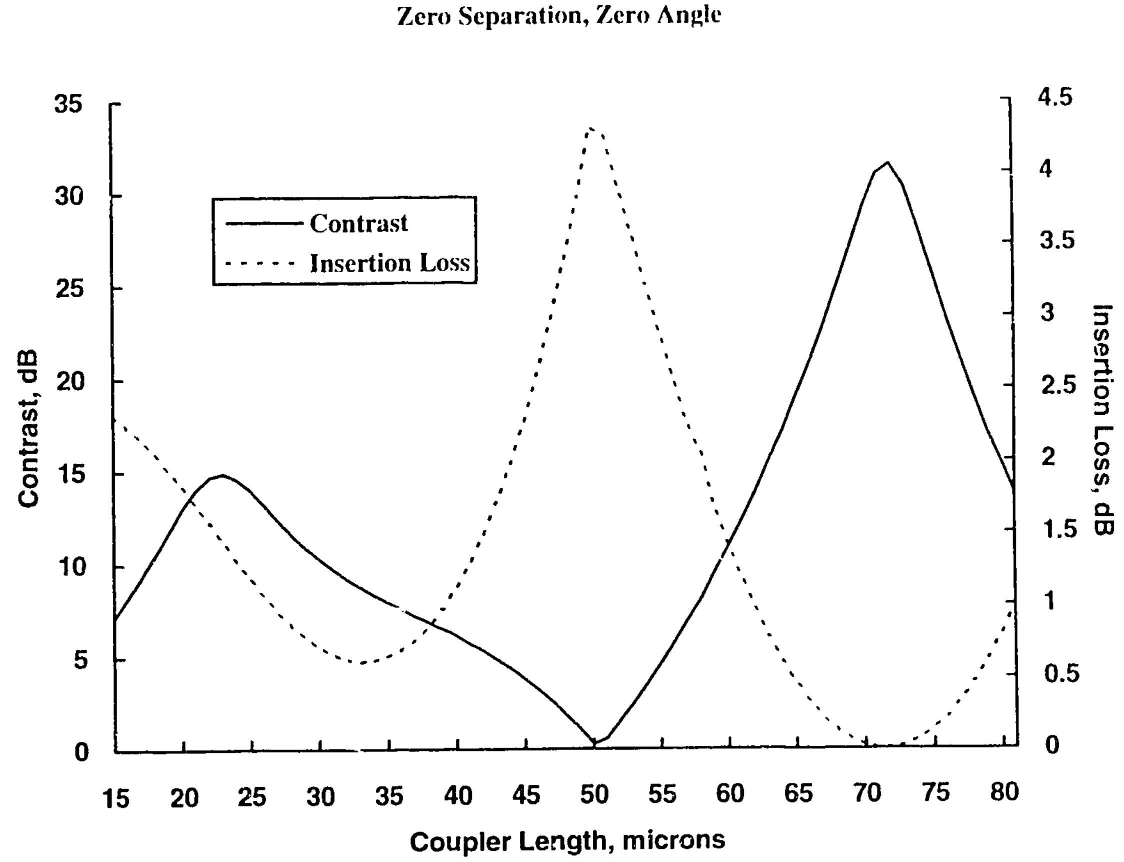


Figure 3.12 Performance of ideal 4 μm-wide coupler as a function of coupler length.

The "selected" output was always taken as the output with the greatest coupled power; the discontinuity about L=50 μm therefore represents a change of selected outputs. The beat length $L_\pi \cong 48$ μm for the 4 μm-wide coupler. The contrast peak at 24 μm thus corresponds to the restricted resonance, while that at 72 μm corresponds to the general resonance. The restricted resonance is degraded by the coupling into the ("forbidden") second-order mode, which impairs the achievable contrast and insertion loss, and separates the coupler lengths for which each is optimized. The general resonance, on the other hand, offers near-ideal performance. The placing of the access waveguides leads to a high overlap between the input modes and the coupler modes and permits an optimum coupler insertion loss of <0.05 dB.

We next examine the more realistic case of 0.5 μm separation and 10 degree branching angle between the input and output waveguides; this configuration is consistent with realizable fabrication. The calculated performance parameters for this design are shown in Figure 3.13. The optimum performance is somewhat degraded. A minimum insertion loss of 1.5 dB is obtained at a coupler length of 65 μm, with a contrast of 21.4 dB. Insertion losses of less than 1.7 dB and contrasts in excess of 20 dB are obtained over a coupler length range of 62 μm to 67 μm.

The specific effect of the branching angle can be seen in Figure 3.14. Here, we have used a 0.5 μm access guide separation at the input and output, and calculated the optimum insertion loss and corresponding contrast as a function of the branching angle. The same branching angle was used at both the input and output junctions in each case.

Neglecting a few numerical anomalies, we see that contrast decreases, roughly linearly, with increasing branching angle, while insertion loss increases, approximately quadratically. The maximum tolerable junction angle, for a given contrast and/or insertion loss specification, can be ascertained from these curves. For an insertion loss of less than 0.5 dB and/or a contrast in excess of 25 dB, the junction angles must be kept below five degrees.

The other coupler configuration that we consider seeks to employ the restricted resonance. For proper restricted resonance operation, we must ensure that the modes defined by (3.28) are not excited. As discussed previously, this can be accomplished by centering the access waveguides at 1/3 and 2/3 of the coupler width. Again assuming an access guide separation of 0.5 μm, and given that the access waveguides are 2 μm wide, this dictates a coupler guide width of 7.5 μm. At this coupler width six guided modes are sustained. In accordance with (3.40) the beat length is increased, to ~162 μm; the restricted 3-dB resonance is therefore expected at a coupler length of 81 μm. The optimum performance of this structure as a function of access guide branching angle is shown in Figure 3.15.

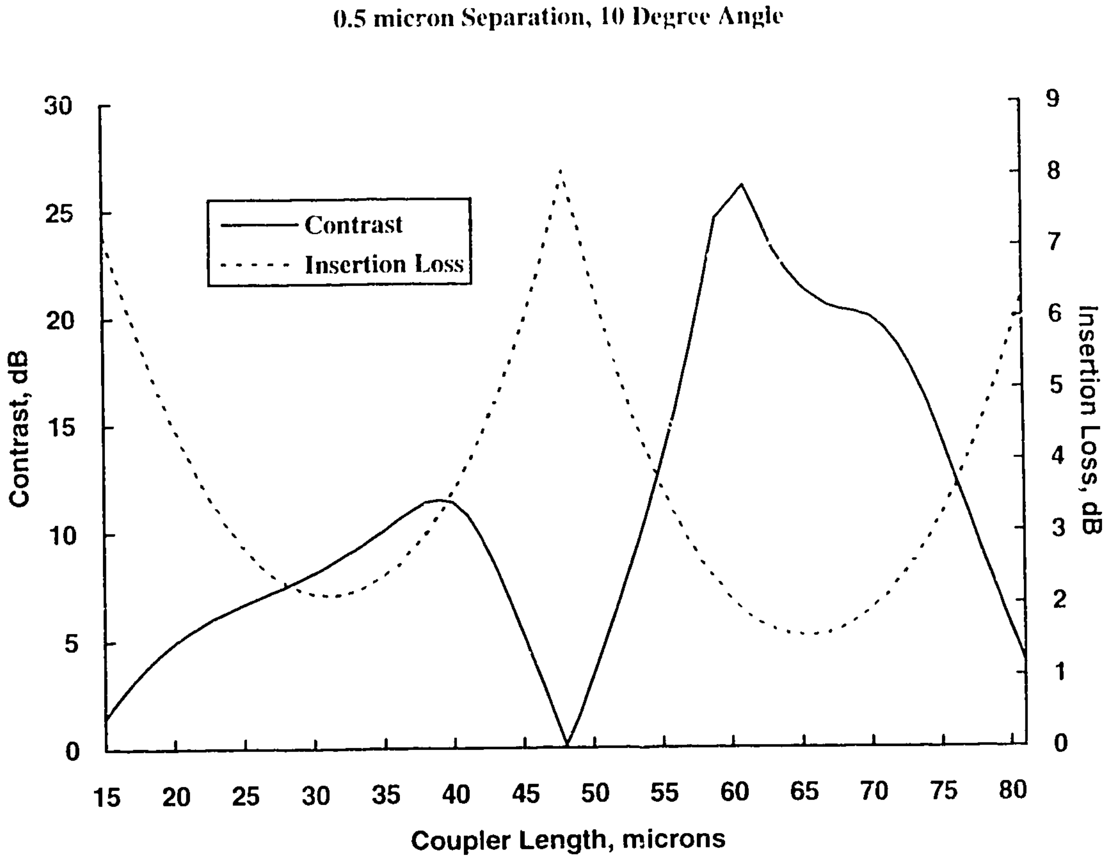


Figure 3.13 Performance parameters as a function of coupler length for 4 μm-wide coupler with 0.5 μm, 10 degree access guide separation.

In this case, the degradation of both contrast and insertion loss with increasing branching angle is worse than in the previous, general resonance case. This is due to a rapid increase in coupling into the forbidden modes, from a normalized power (fraction of the total incident power) of -24 dB at 0 degrees to -10 dB at 20 degrees branching angle. Furthermore, we note that the coupler length, although we are employing the restricted rather than the general resonance, is greater than that for the 4 μm coupler, due to the effect of (3.40). We therefore conclude that the narrower device based on the general resonance principal offers the best length-economized coupler design for our purposes. With this coupler design, from Figure 3.14, we predict that we can achieve 0.5 dB insertion loss, with a device length of ~65 μm, if the junction angles are restricted to 5 degrees. This performance compares favourably with previous multi-mode coupler implementations. For example, in [47], a four-mode deeply etched waveguide 3-dB coupler was reported which achieved 0.5 dB insertion loss. This device was 250 μm long. Another

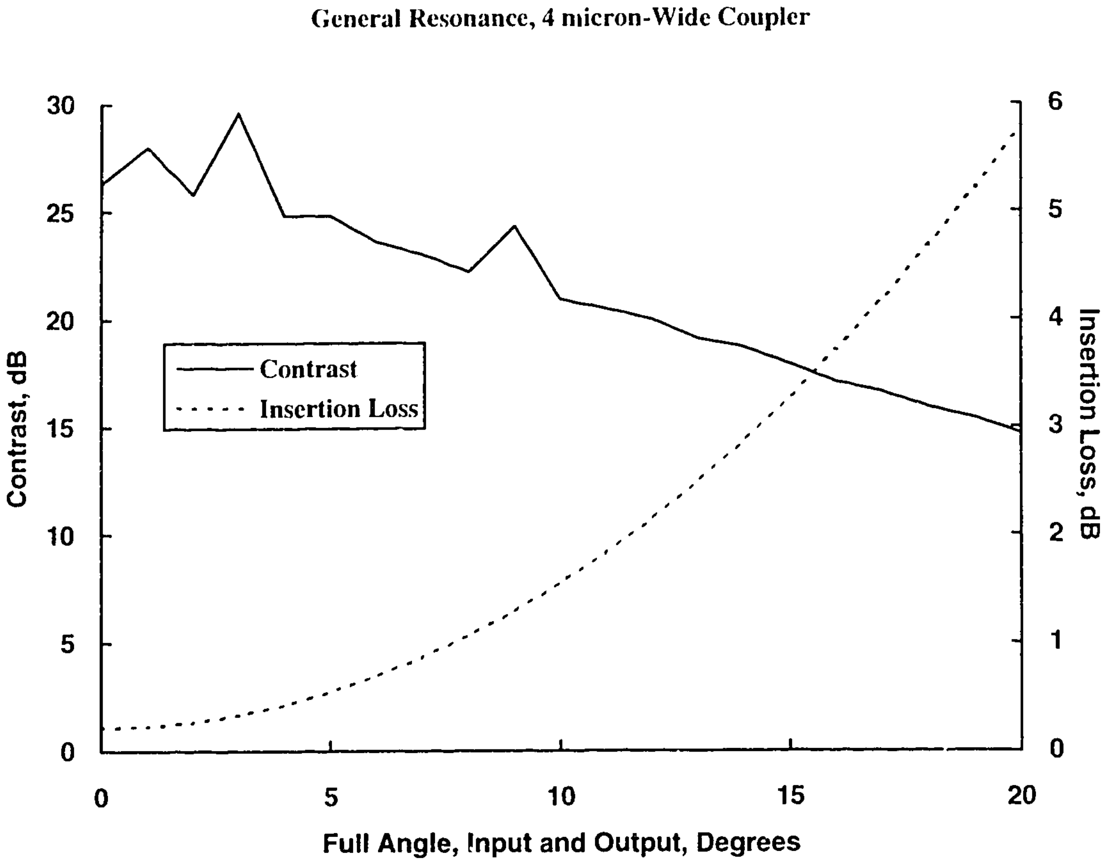


Figure 3.14 Performance as a function of access guide branching angle for 4 μm-wide coupler, general resonance (Optimum coupler lengths, ~70 μm, for minimized insertion losses).

recently reported 3-dB coupler, also using deeply etched ribs, achieved 0.7 dB insertion loss with a 111 μm-long device [6]. These devices incurred a length penalty by using zero branching angles at the junction, with access guide separations of ~2 μm. We have shown that the length of a coupler can be reduced, without significant performance penalty, by introducing a small angle at the input and output junctions.

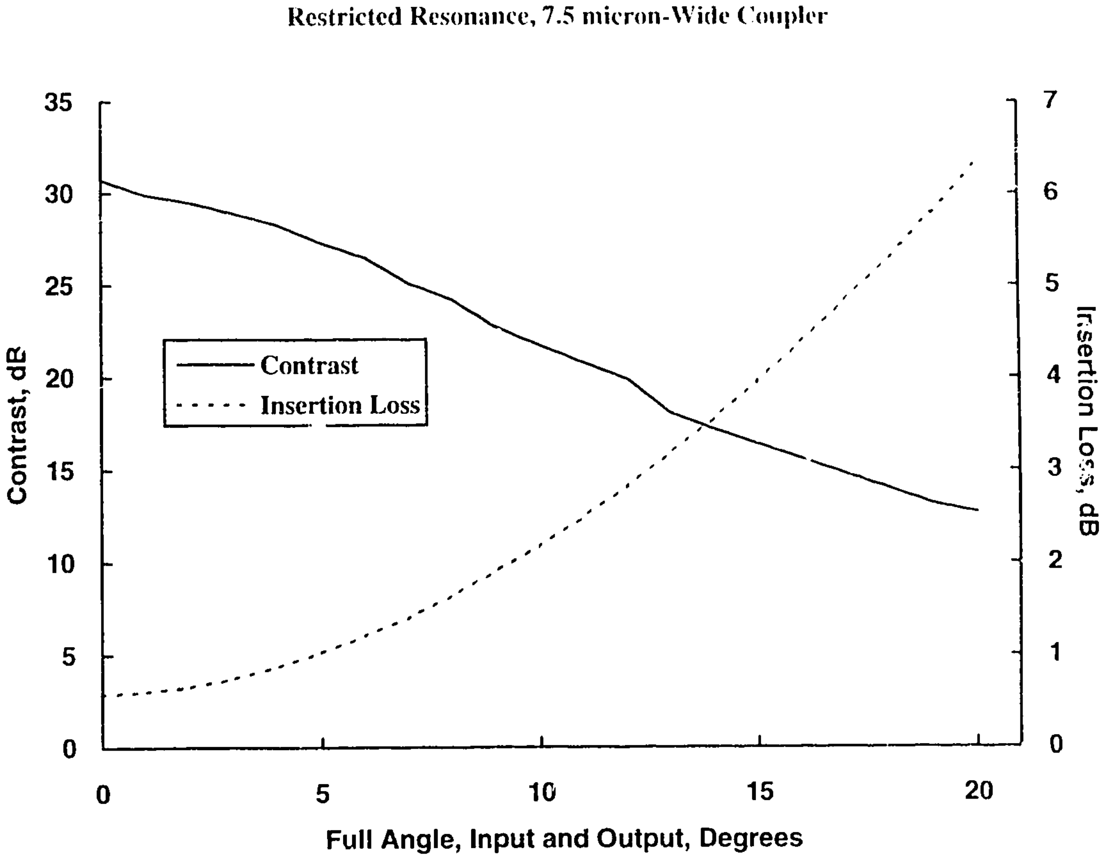


Figure 3.15 Performance as a function of access guide branching angle, 7.5 μm-wide coupler, restricted resonance (Optimum coupler lengths, ~80 μm, for minimized insertion losses).

The smallest junction angle which can be used in practice is determined by the repeatability of the junction fabrication. An uncertainty δS in the realized access waveguide separation at a junction (refer to Figure 3.4) introduces an uncertainty in the length of the coupler, given by:

$$\delta L = \frac{\delta S}{2\tan\left(\frac{\theta}{2}\right)} \tag{3.50}$$

where, again, θ is the full junction angle. In practice, δS is likely to be 0.1-0.3 μm. In Table 3.2 below, we show the total coupler length uncertainties introduced for various junction angles (assumed the same at the coupler input and output), at different values of δS.

Junction Angle, Degrees	**δS=0.1 μm**	**δS=0.2 μm**	**δS=0.3 μm**
1	11.46 μm	22.92 μm	34.38 μm
2	5.73 μm	11.46 μm	17.19 μm
5	2.29 μm	4.58 μm	6.87 μm
10	1.14 μm	2.29 μm	3.43 μm

Table 3.2 Maximum length variation of a 2X2 waveguide coupler, for various uncertainties in the blunted junction tip width (δS), as a function of the full junction angle at the input and output.

The minimum usable junction angle is therefore dictated by the tolerable uncertainty in the coupler length. If δS can be restricted to 0.2 μm in the pattern transfer process, then, from Figure 3.12 and Table 3.2, a contrast of 20 dB can be repeatably achieved if the junction angle is 10 degrees. However, an insertion loss penalty of up to 1.2 dB is incurred compared with the longer devices of [6] and [47]. As smaller junction angles are used, better performance can be obtained, but either the fabrication yield will drop, or better control of the pattern transfer process is required.

A width tolerance analysis for the nominally 4 μm-wide coupler is depicted in Figure 3.16. Here, access guide separations of 0.5 μm and branching angles of 0 and 10 degrees are assumed; the modeled insertion losses for coupler lengths of 71 μm and 65 μm respectively are calculated as a function of the coupler width. For the 10 degree case, control of the coupler width to within slightly better than 0.2 μm is required to keep the total insertion loss below 2 dB. In practice, width definition on the order of ±0.1 μm accuracy can usually be achieved.

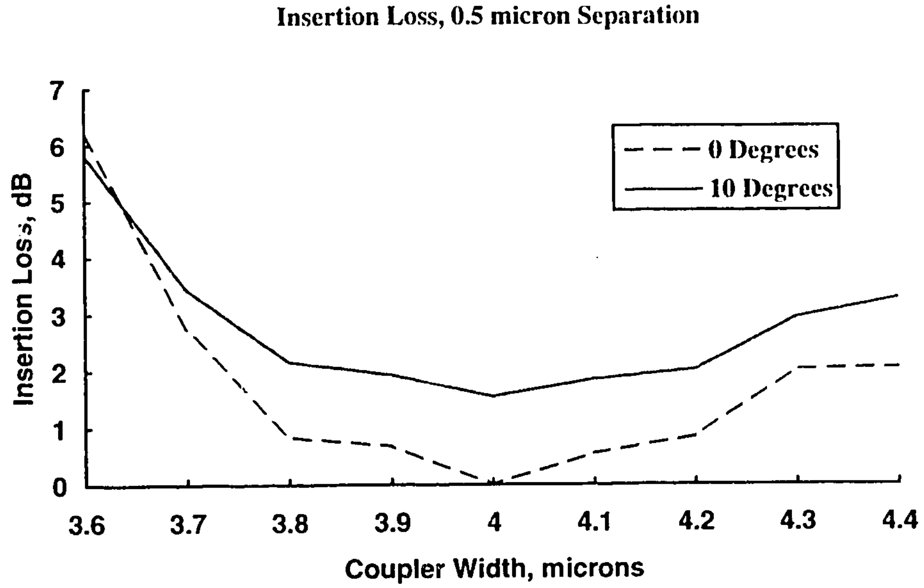


Figure 3.16 Modeled insertion loss as a function of coupler width, for 0 and 10 degree branching angles of access waveguides. The access guides are separated by 0.5 μm in all instances and the coupler lengths are 71 μm and 65 μm respectively for the 0 and 10 degree cases.

We now examine the performance of four-port Mach-Zehnder switches, as shown in Figure 3.7, as a function of the length of the 3-dB couplers at the input and output. We again assume 4 μm-wide couplers, and further assume that identical coupler lengths and junction geometries are used at the input and the output. The first coupler is assumed excited by a guided mode in one input, and the amplitudes and relative phases of the output waveguide modes of the first coupler are calculated using the procedure described in section 3.4. The phase separation between the output waveguide modes is easily determined from their complex amplitudes:

$$\Delta\phi = (\phi_1 - \phi_2) = \arctan\left(\frac{a_1^{im}}{a_1^{re}}\right) - \arctan\left(\frac{a_2^{im}}{a_2^{re}}\right) \qquad \textbf{(3.51)}$$

where the superscripts *re* and *im* denotes real and imaginary parts, respectively. Mode fields with the calculated amplitudes and phase separation (3.51) are then used as the inputs to the second, output coupler. This corresponds to the unbiased state of the switch, while π-phase modulation converts $\Delta\phi \rightarrow \Delta\phi \pm \pi$, which is simply $\Delta\phi \rightarrow -\Delta\phi$ if the couplers are in true 3-dB states so that $\Delta\phi = \pm\pi/2$. The contrast of the device is determined from the amplitudes of the output waveguide modes of the second coupler. The insertion loss of the whole Mach-Zehnder switch, assuming that the constituent waveguides are lossless, is determined using the mode amplitudes throughout the structure, in an extension of (3.49.b).

In Figure 3.17 we show the modeled performance of the four-port Mach-Zehnder switch, where access waveguide separations and branching angles of 0.5 µm and 10 degrees, respectively, are assumed at all junctions. Again, the couplers are 4 µm wide and an identical coupler length is used at the input and the output of the Mach-Zehnder switch.

Contrast peaks corresponding to the restricted and general 3-dB resonances of the couplers are again observed. However, the restricted resonance contrast is now greatly suppressed. This is because a true resonance is not achieved, due to the excitation of the second-order mode of the coupler, and when a second, identical coupler is cascaded, the accumulation of amplitude and phase deviations from the true 3-dB condition severely impairs the total performance. An even smaller contrast peak is observed at ~47 µm coupler length. This is due to restricted resonance cross-coupler operation of both couplers. Again, however, since the correct input condition for restricted resonance operation is not satisfied, the achieved contrast of the total device is very poor. Only the general 3-dB resonance at ~ 65 µm coupler length yields acceptable performance. About 21.5 dB contrast and 2.5 dB insertion loss is expected at this point. Note that in the four-port device, both the restricted and general 3-dB resonances lead to a cross-state of the full Mach-Zehnder switch. In the three-port device, opposite output waveguides were favored depending on which 3-dB resonance was operative.

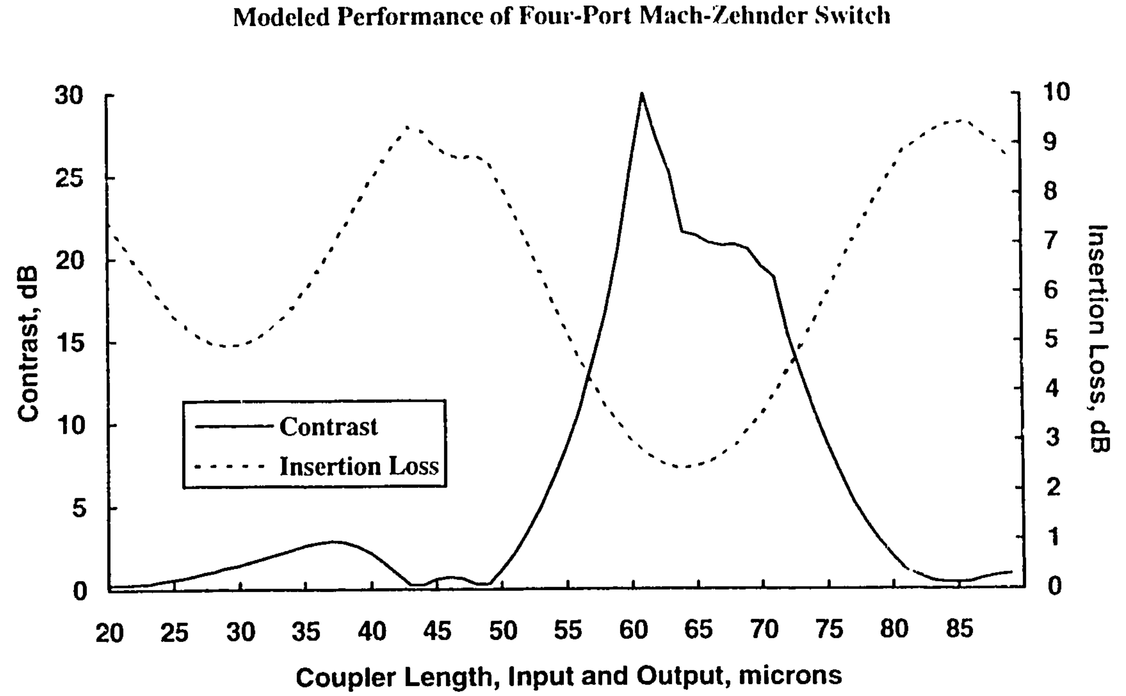


Figure 3.17 Modeled performance of four-port Mach-Zehnder switch in the unbiased state as a function of multi-mode coupler length. Identical input and output couplers are assumed. All junctions are characterized by a 0.5 μm separation and 10 degree branching angle of the access waveguides.

3.7 Mach-Zehnder Switch Design and Fabrication Chrome-on-quartz photomasks were produced by e-beam writing for fabrication of the Mach-Zehnder devices. Since some sidewall slope (~5 degrees from vertical) is produced by reactive ion etching, the bottoms of the deeply etched ribs are wider than the tops. This was taken into account by making the photomask traces defining the waveguides narrower than the desired rib widths, so as to attain the design width at the vertical center of the quantum well stack.

Each individual switch uses 600 μm-long electrodes in the modulator section, leading to expected modulation voltages of 4.5 V (phase=π) and 3 V (phase=π/2). Curved access waveguides are defined according to the following formula:

$$\pm x(z) = \frac{w}{2} + \frac{hz}{l} - \frac{h}{2\pi}\sin\left(\frac{2\pi z}{l}\right) \tag{3.52}$$

where x is the lateral displacement of the center of a single-mode waveguide of (design, not mask trace) width w (2 μm), and h is the final lateral displacement at the total bend length l. This bend geometry reduces mode-mismatch losses at straight-to-bend transitions. In our designs, values of h=10 μm and l=100 μm are used in (3.52) to define the access waveguides of the multi-mode couplers, which are nominally 4 μm wide in all cases, such that the access waveguides converge ($\pm x(0) = w/2$) at the coupler input/output junctions. In fact, the finite resolution of the pattern transfer process leads to blunting of the junction tips, as expected. The realized junction tips were fairly consistently about 0.65 μm wide at the full etch depth. This had the effect of pushing the actual junction location out to z such that $\pm x(z)=w/2+sep/2$, where sep is the total realized separation of the access waveguides. The junction blunting introduces a separation angle of the access waveguides, given effectively by twice the arc tangent of the derivative of (3.52). For sep=0.65 μm, the full junction angle (after accounting for the sidewall slope of the ribs) is calculated to be about 8 degrees. Therefore, the realized junction parameters (0.65 μm separation, 8 degrees branching angle) were very close to the modeled case (0.5 μm separation, 10 degrees branching angle). The actual length extensions of the couplers due to the junction tip blunting is taken into account in the presentation of the experimental results. The unplanned length extension was much less pronounced in the devices processed using the second photomask, where an explicit blunted tip width of 0.6 μm was defined at all the junctions.

A variety of coupler lengths were included in the devices on the masks. For the three-port devices, the input 1X2 coupler targeted the optimized length reported in [52], although this length was not realized in the processing using the first photomask, as a result of the coupler extension due to the junction tip blunting. The four-port devices used equal coupler lengths at the input and output of the Mach Zehnder modulator.

Following photoresist patterning and reactive ion etching to define the waveguide ribs, the wafer was coated with oxide and planarized using polyimide. The oxide was etched to create 1.5 μm-wide openings over the modulator arms for electrical contact. After patterning for metal lift-off, Ti/Pt/Au was evaporated onto the top surface and lifted off to create electrodes and contact pads. The wafer was next reactive ion etched to a depth of 0.3 μm, over patterned regions 10 μm long at the ends of each electrode. This removed the upper p+ layer in these regions and served to enhance the electrical isolation between the arms of the modulators. Finally, the wafer was thinned and back-metalized by Ni/Ge/Au evaporation. The devices were cleaved to a length of 1.5 mm, providing for straight input and output waveguides about 300 μm long. This is believed to be the smallest total size yet reported for an electro-optic Mach-Zehnder switch.

In Figure 3.18, we show an optical photograph of parts of three Mach-Zehnder switches fabricated using the second photomask. Each rib waveguide lies in the centre of a 20 μm-wide etched trench. Straight single-mode waveguides are included adjacent to each switch. The output couplers, contact pads and parts of the modulator arms, as well as some of the alignment marks, are visible. The isolation trench etched through the top p+-layer is visible.

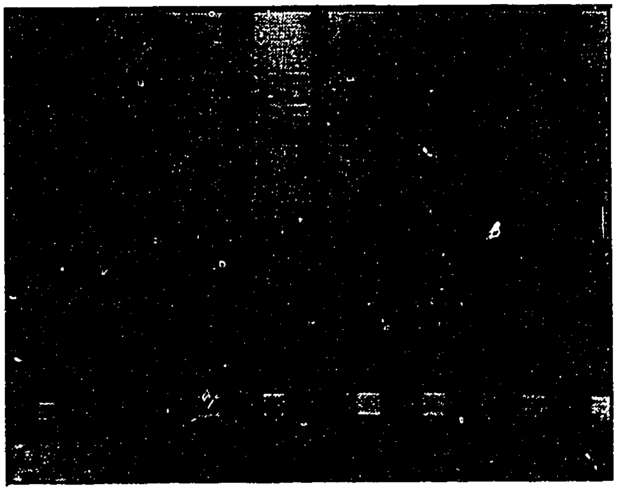


Figure 3.18 Optical photograph of the output portions of three Mach-Zehnder switches, fabricated using the second photomask.

3.8 Experimental Results The fabricated Mach-Zehnder devices were tested using the experimental arrangement shown in Figure 3.19. TE-polarized 1.55 μm light from a DFB laser was collimated by a 10X objective lens, and focused into the devices using a 20X objective. The output light was imaged by a 60X objective into a pair of lenses which expanded the image by four times. A variable beam splitter then directed part of the light into an IR vidicon camera, and the remainder toward a pinholed Ge photodiode which was mounted on a motorized *x-y* positioning stage. This stage was computer-controlled so that two-dimensional true-intensity scans of the near-field output of the device could be obtained. Mechanical probes were used to achieve electrical contact to the contact pads of the devices, applying a reverse bias to one modulator arm, and grounding the other. The metalized bottom contacts were also grounded. In some instances, the devices were mounted on copper holders using conductive epoxy to enhance the mechanical stability of the set-up.

Unfortunately, the fabrication yield was severely limited by certain difficulties. In many cases, the lift-off process removed the metal from the modulator arms of the devices, leaving no way to apply a bias. An attempt was made to repeat the metalization procedure with one batch of devices; however, the correction was only partially successful. An even more serious problem was that the achieved electrical isolation between the modulator arms of individual devices was, in general, poor. For most devices, the measured resistance between the modulator arms was on the order of 2-4 kΩ; for these devices, independent modulation of the two sides, and hence optical switching, could not be achieved. For those few devices where effectively independent modulation of the two sides of the interferometer could be obtained, the measured resistance between opposing contacts was on the order of 80-100 MΩ. Other authors have reported resistances in excess of 900 MΩ for successful operation of integrated Mach-Zehnder interferometers, but such high values are typically achieved either by high-energy (>100 keV) implantation of ions such as hydrogen, helium and fluorine [57], or by etching fully through the upper doped layers, markedly increasing the optical losses of the device [58]. Our approach of etching through only the highly doped layers is technologically simple and leads to only relatively small optical losses, but failed to repeatably produce good electrical isolation between the two sides of the modulators. Finally, the quality of the etched ribs was somewhat disappointing. Width variations and defects were evident that impaired the performance of the devices. Given these difficulties, only limited experimental results were achieved.

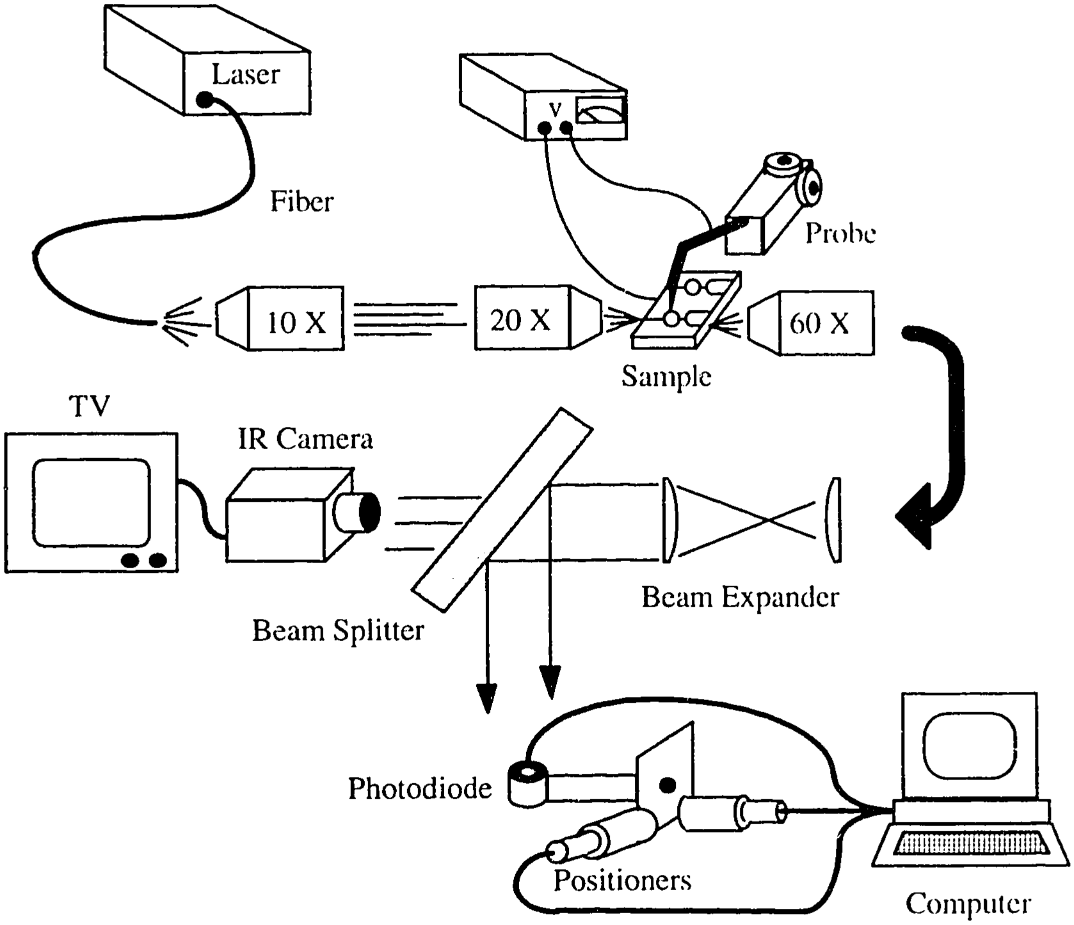


Figure 3.19 Schematic of the experimental arrangement for testing of the Mach-Zehnder switches. Collimated laser light is focused into the devices, and the expanded near-field image is split between an IR camera and a Ge photodiode on a computer-controlled, motorized *x-y* mount. Two mechanical probes (only one shown) were used to reverse bias/ground the opposing electrodes of the devices.

The three-port Mach-Zehnder switches were tested first. In the unbiased state, these devices should direct equal intensities to the two output waveguides. Therefore, individual switches which were more than 1 dB out of balance at zero bias were considered defective and were not tested further. Of the remaining devices, only one (realized) coupler length, 42 μm, was represented in multiple devices with good electrical characteristics. Contrast was determined by comparing the peak powers in the output waveguide mode images, measured using the photodiode, and averaging the measured contrasts of the working devices. The total (on-chip) excess loss of each Mach-Zehnder switch was determined by comparison to a straight single-mode waveguide of equal length. The loss due to the output coupler, for comparison to the model, was determined by comparison to single-output Mach-Zehnder intensity modulators, which employed the 1X2 coupler at both the input and the output and which were included on the first mask.

An average measured contrast of 9.2 dB was measured for the 42 μm coupler devices. The peak contrasts were obtained at a bias voltage of 3 V applied to one electrode - the expected π/2 modulation voltage. The average total on-chip insertion loss was 9.5 dB. The average total insertion loss of the single-output Mach-Zehnder modulators was 10.3 dB. If only the multi-mode couplers are assumed to contribute to the loss, the insertion loss at each 1X2 coupler is therefore 5.15 dB. Subtracting this from 9.5 dB, we deduce the insertion loss of the 42 μm 2X2 coupler to be about 4.3 dB. These performance figures, 9.2 dB contrast and 4.3 dB coupler loss, are in close agreement with the predicted values from Figure 3.13. Furthermore, at 1.5 mm total length, this is the smallest integrated Mach-Zehnder switch yet reported which achieves close to 10 dB contrast [59],[60]. Note that the insertion loss of the 1X2 input coupler was high since the realized junction tip width, ~0.65 μm, was not explicitly defined on the first mask, and hence the length of the 1X2 couplers was extended to a non-optimum value.

In Figure 3.20, we show lateral line scans of the output of one of the devices, taken from a digitized IR-camera image. The dashed lines show the output at 0 V bias, while for the solid line, 3 V reverse bias has been applied to one electrode, illustrating the switching effect. In Figure 3.21, we compare modeled versus experimental characteristics for the optical power in each waveguide, as a function of the reverse bias voltage applied to one electrode. The phase shift induced in one modulator waveguide, as a function of bias voltage, was deduced from Figure 2 in [50], which shows the experimental intensity versus bias voltage characteristic for a single-output Mach-Zehnder intensity modulator. This phase shift was applied to one of the input mode fields in the computer model. The calculated contrast and insertion loss were used to deduce the power in each waveguide, normalized to match the power (in arbitrary units) in the "selected" waveguide for the modeled and experimental curves, at 3 V bias. In Figure 3.21, the modeled curves are shown as dashed lines, while the experimental values are shown as points joined by solid (cubic spline) curves. The variance in the experimental values, determined by measurements on all the working devices, is shown as error bars. The average variance among the bias voltages is used.

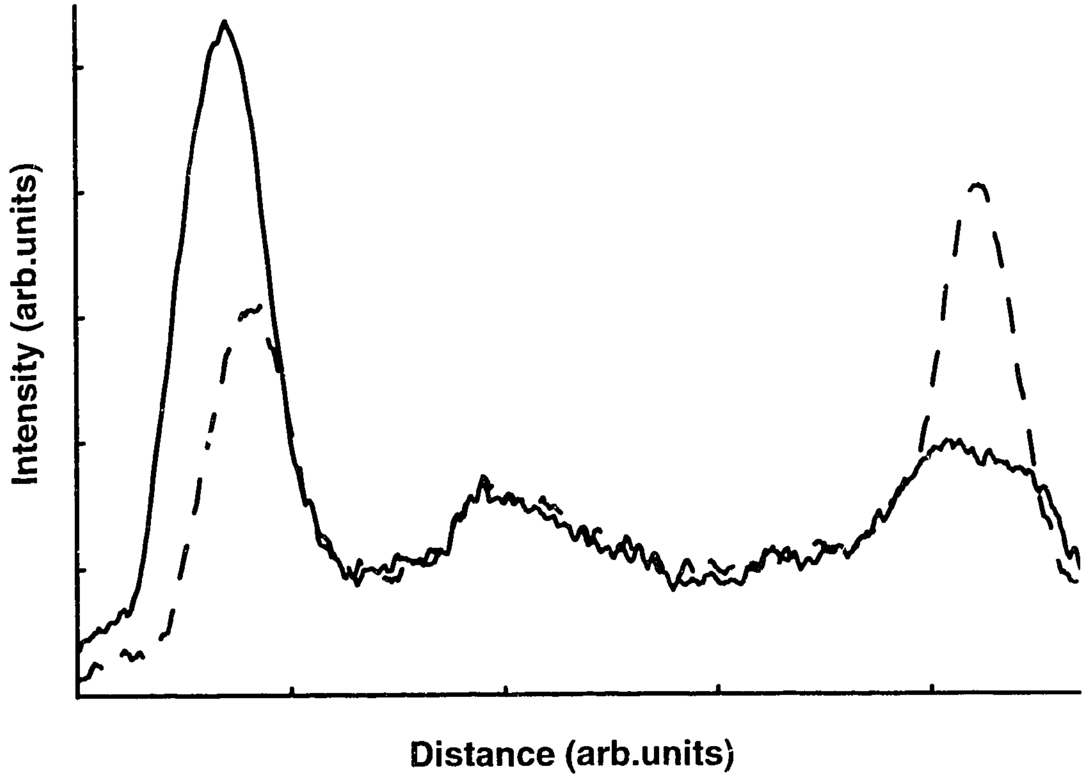


Figure 3.20 Lateral line scans of the output field (from digitized IR-camera image) of a 3-port Mach-Zehnder switch with a 42 μm-long output coupler. The dashed lines show the output at zero bias; for the solid line, 3 V reverse bias has been applied to one electrode.

Of the fabricated four-port devices, none were found with sufficient electrical isolation between opposing electrodes to achieve switching. However, since the performance of the full four-port Mach-Zehnder switch, not just the output coupler, was modeled in Figure 3.17, we can measure the contrast and insertion loss at zero bias for direct comparison to theory. Only devices which showed balanced contrast to within ~1 dB as the input coupling was switched from one input waveguide to the other were considered "defect-free" and construed as valid measurements. The insertion loss is measured relative to adjacent single-mode waveguides. In Figure

3.22.(a)&(b) we show example two-dimensional intensity scans of the near-field image from the output of a four-port Mach-Zehnder switch, in this case with 43 μm-long input and output couplers. The input light is coupled to opposite input ports in

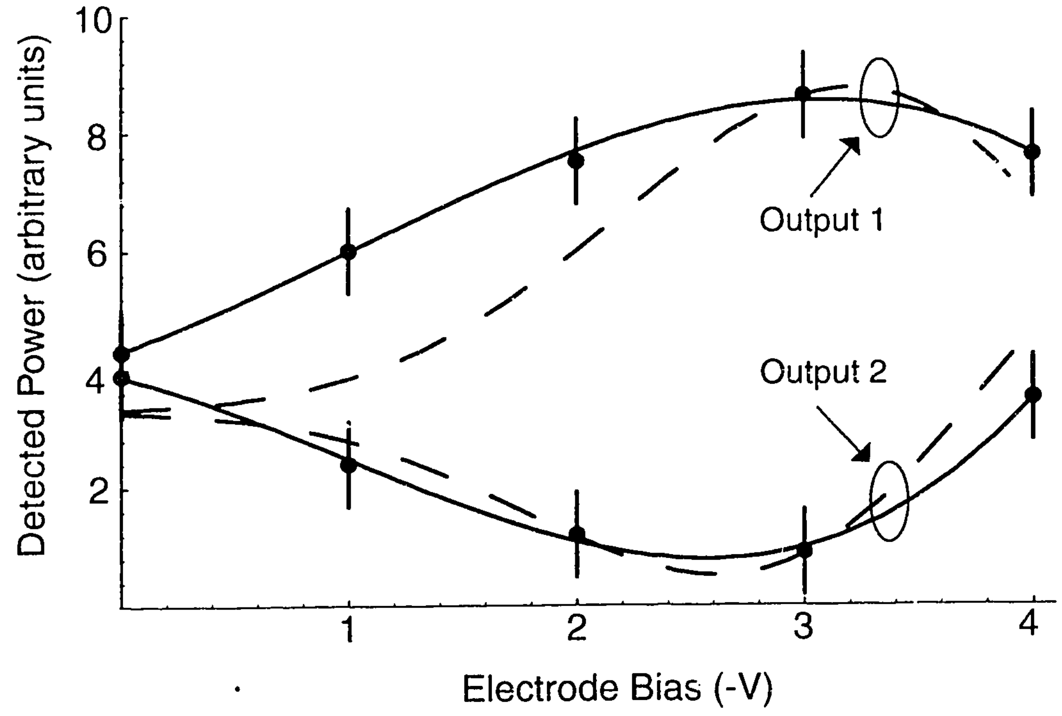


Figure 3.21 Modeled (dashed lines) versus measured (solid lines with experimental points) power from the outputs of a 3-port Mach-Zehnder switch, with a 42 μm-long output coupler, as a function of the reverse bias applied to one electrode. The variance in the experimental values, determined by measurements on several devices, is indicated by error bars.

(a)&(b) respectively. In Figure 3.23 we show the measured contrasts and insertion losses of the fabricated four-port Mach-Zehnder switches, as a function of coupler length (input and output). The experimental values are shown as triangles (contrast) and crosses (insertion loss), and the modeled curves of Figure 3.17 are included for comparison. Some of the experimental points represent averages among several "defect-free" devices.

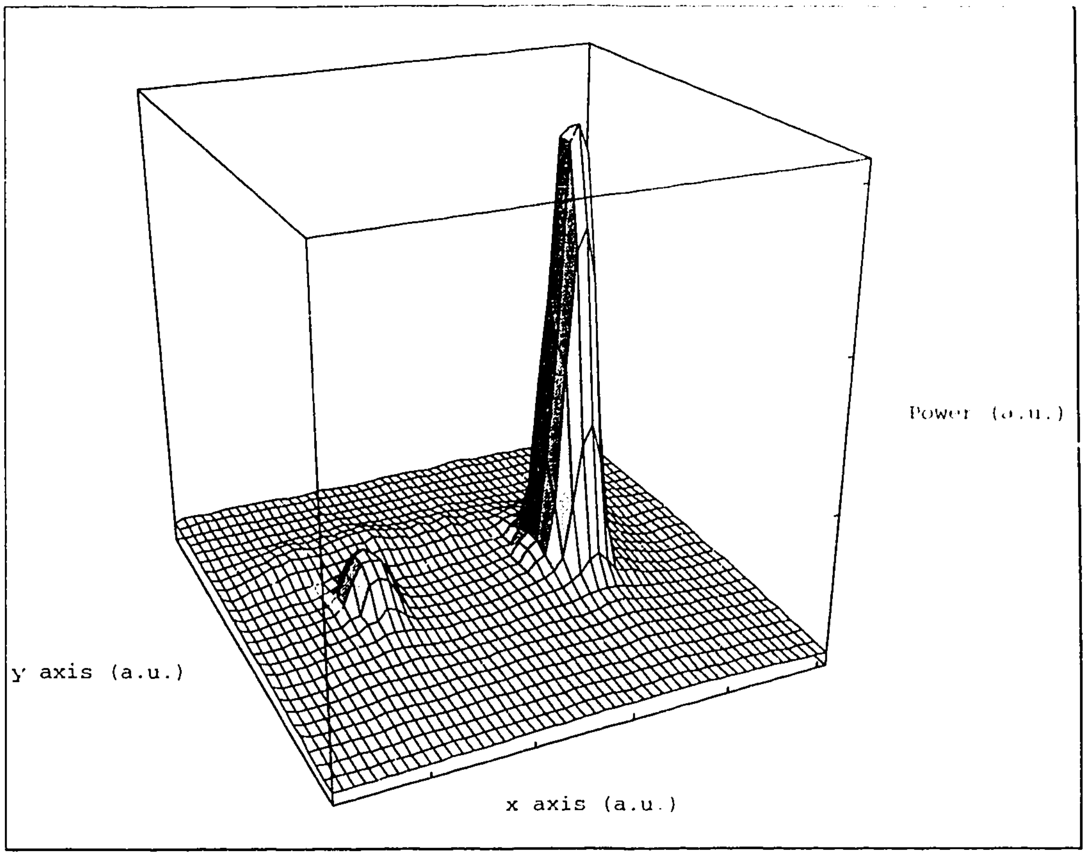


Figure 3.22.(a) Two-dimensional intensity scan of the near-field image at the output of a four-port Mach-Zehnder switch with 43 μm-long input and output couplers.

General agreement between theory and experiment is revealed in Figure 3.23, with the exception that significantly better than predicted performance is observed around the restricted 3-dB coupler resonance. The high contrast observed at 65 μm coupler length, 19.4 dB, confirms the superior performance of the couplers when they are operated on the general 3-dB resonance. Though the total observed on-chip excess loss at this coupler length, 5.6 dB, is higher than the predicted minimum of 2.5 dB, the performance of the Mach-Zehnder switch with 65 μm couplers is nevertheless much better than similar devices reported in the literature [57]. In Figure 3.24, we show line scans of the output intensities, taken from the two-dimensional true intensity scans at the vertical center of the waveguide modes, of one of the four-port switches with 65 μm couplers. The solid and dashed lines are obtained by input coupling to opposite ports.

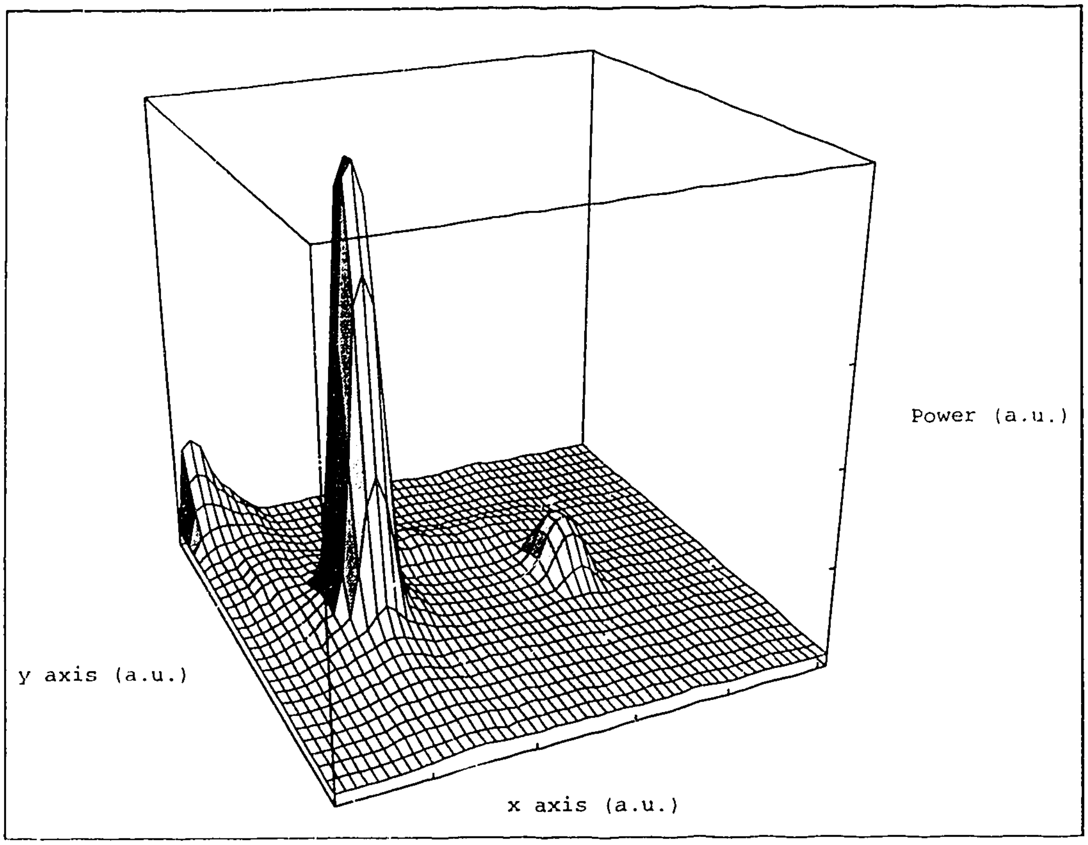


Figure 3.22.(b) Intensity scan of the near-field image at the output of the same four-port Mach-Zehnder switch, but with coupling to the opposite input waveguide.

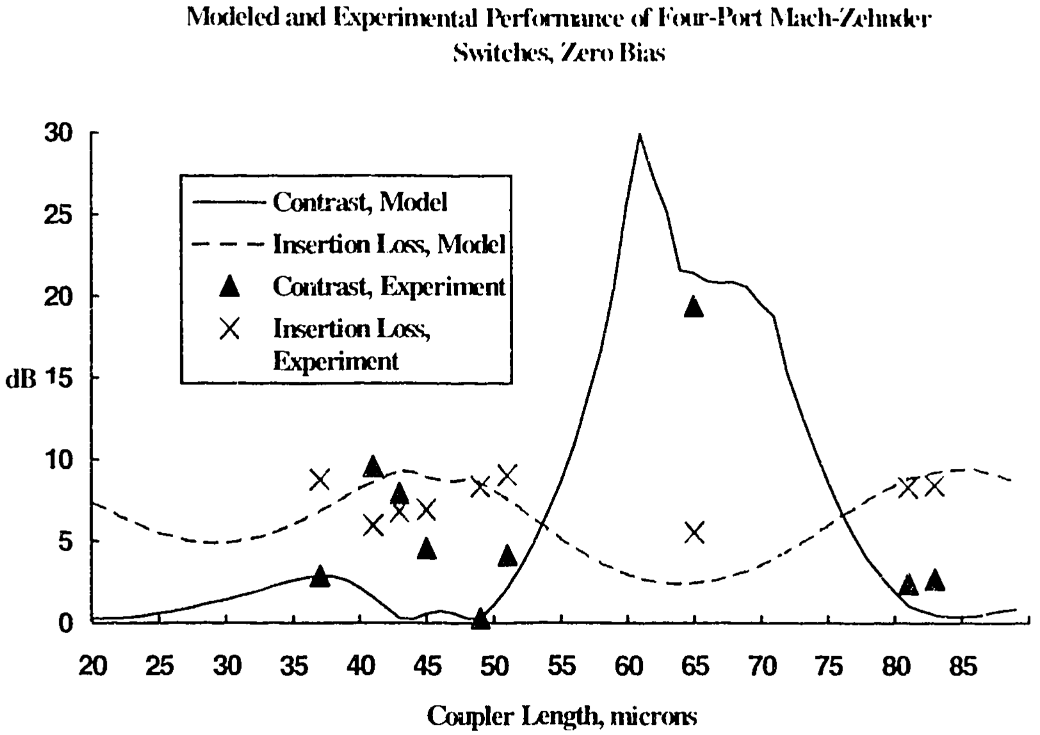


Figure 3.23 Modeled and experimental performance of four-port Mach-Zehnder switches at zero bias. Solid line shows modeled contrast, dashed line shows modeled insertion loss, triangles indicate experimental contrast, crosses indicate experimental insertion loss.

3.9 Conclusions In this chapter, we have discussed the principle of resonant self-imaging in homogeneous multi-mode waveguides, and described a computer model for analyzing waveguide couplers based on this principle. The model, which is well-suited for analyzing couplers implemented using deeply etched rib waveguides, is able to simulate the effect of setting the coupler access (input and output) waveguides at an angle relative to the propagation axis of the coupler. Using this model, we have shown that very compact (<100 μm) 3-dB couplers can be constructed by finding an optimum combination of access waveguide junction angle and coupler width. The calculated performance of such optimized couplers is comparable to that of longer devices reported in the literature. The model was used to design couplers for use in InP-based Mach-Zehnder switches. Three-port switches with couplers operating on the restricted resonance, and four-port switches with couplers operating on both the restricted and general resonances, were experimentally evaluated. We demonstrated that almost 20 dB contrast and less than 6 dB total on-chip insertion loss could be achieved with a Mach-Zehnder switch only 1.5 mm long.

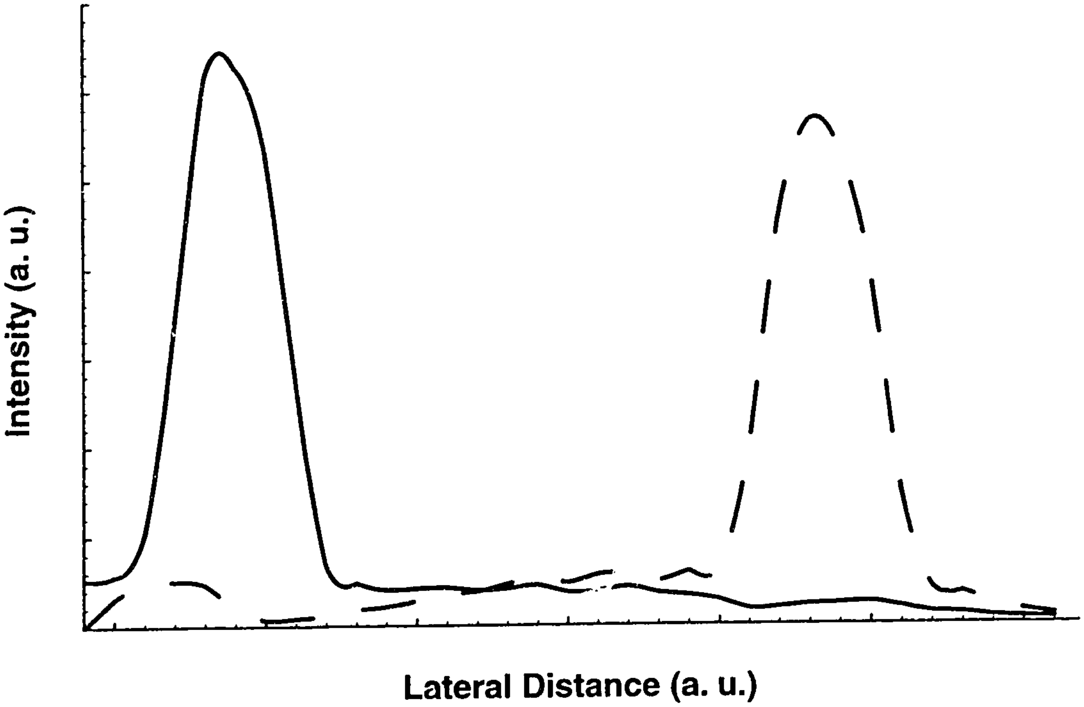


Figure 3.24 Intensity line scans of the imaged output from a four-port Mach-Zehnder switch with 65 μm couplers. The solid and dashed lines are obtained by coupling to opposite input ports.

Chapter 4

Bent Multi-Mode Waveguide Couplers for (De)Multiplexing of Broadly Separated Optical Wavelengths

Integrated devices which achieve two-wavelength optical (de)multiplexing are required for many applications. For example, integrated rare-earth doped waveguide amplifiers and lasers require the combination and separation of signal and pump light. Also, bandwidth augmentation in bidirectional or local network fiber systems could be achieved in the short term by using two channels, or two sets of channels, at widely separated fiber attentuation minima. Such an approach offers a doubling of usable bandwidth using only simple and inexpensive components, with very little inter-channel interference. A number of integrated devices have been demonstrated which achieve this multiplexing function, including conventional directional couplers [61], asymmetric Mach-Zehnder interferometers [62], and y-junction devices employing asymmetry in the refractive index profile of the access waveguides [63]. These approaches are generally characterized by very large device sizes, [61] & [62], or difficult fabrication [63].

Multi-mode waveguide couplers, described in the previous chapter, can also be used for (de)multiplexing. The beat length which determines the state of coupler operation, given by equation (3.10), is wavelength-sensitive. Therefore, a 2X2 coupler can potentially operate in the cross-state at one wavelength, and in the bar-state at the other. It can then be used to combine and/or separate light at those two wavelengths. As we saw in Chapter 3, high contrast (>20 dB) and low insertion loss (<0.5 dB) can be achieved using multi-mode couplers. Furthermore, multi-mode couplers are well-suited for implementation in deeply etched rib waveguide structures, which can be sharply bent without suffering radiation losses are therefore are ideal for ring resonator realization. We found earlier that in the limit of strong lateral guiding, approached with deeply etched rib waveguides, the beat length $\tilde{L}_\pi$ is given by equation (3.40), repeated here:

$$L_\pi \cong \frac{4 n_o^e W^2}{3\lambda_o} \tag{4.1}$$

where n_o^e is the effective index of the region under the rib at wavelength λ_o, and W is the rib width. From (4.1), we see that the ratio of the beat lengths at two different wavelengths is simply:

$$\frac{L_\pi^1}{L_\pi^2} = \frac{n_1^e \lambda_2}{n_2^e \lambda_1} \cong \frac{\lambda_2}{\lambda_1} \tag{4.2}$$

Therefore, in the strong guiding regime, the ratio of coupling lengths at two different wavelengths is essentially fixed at approximately the ratio of the wavelengths themselves. For arbitrary pairs of wavelengths, the relative interaction lengths thus may not be optimized for (de)multiplexer implementation. For example, for wavelengths of 850 nm and 1550 nm, the ratio given by (4.2) is 1.82, whereas a ratio of 1.5 provides for bar-coupling at 850 nm and cross-coupling at 1550 nm when the coupler length is:

$$L_{\text{coupler}} = \begin{cases} 3L_\pi^{1550} = 2L_\pi^{850} & \text{(restricted resonance)} \\ 9L_\pi^{1550} = 6L_\pi^{850} & \text{(general resonance)} \end{cases} \tag{4.3}$$

Good (de)multiplexer performance is therefore often only achievable by resorting to very long devices, i.e. some integer multiple of the beat lengths for which a desired ratio is obtained. Such long devices are undesirable from an integration perspective. In weakly guiding waveguide configurations, adjusting the coupler width can usually modify the ratio of beat lengths at two wavelengths to some degree. However, for strongly guiding structures, a new degree of freedom is needed. Other authors have already demonstrated that the radius of a waveguide bend can sometimes be used as a design parameter [64]. In this chapter, we examine a way to optimize multi-mode waveguide couplers for operation as (de)multiplexers by introducing a circular bend to suitably tailor the properties of the guided mode fields. We begin by developing a method for calculating the modes of circularly bent waveguides.

4.1 Analysis of Bent Waveguides A variety of techniques have been developed for the numerical analysis of curved waveguides. The methods based on conformal transformations [65]-[67] are particularly useful for two reasons. First, they provide a convenient conceptual picture of the physical effects of the bend; second, they yield equations which are well-suited to numerical solution by the same methods applicable to straight waveguides. In this work, we use the mathematical formalism presented in [66].

Consider a rib waveguide as shown in Figure 4.1. The rib has a width of $2a$, and follows a circular path of radius ρ at the center of the rib. Cylindrical coordinates r, φ and y are defined. We also define a second radial coordinate:

$$\xi \equiv r - \rho \tag{4.4}$$

so that $\xi = 0$ at the rib center.

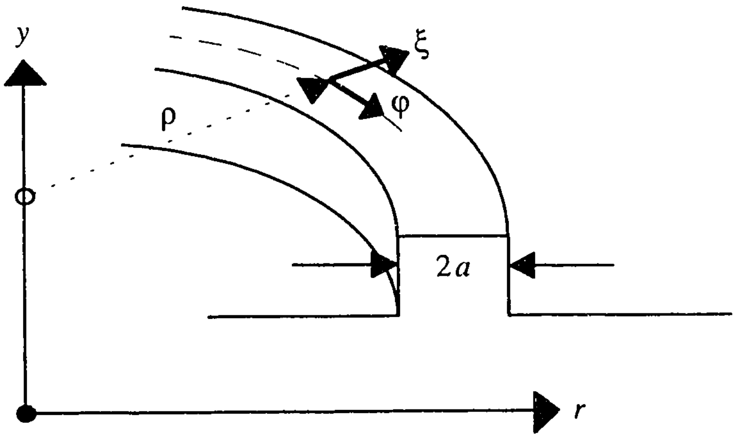


Figure 4.1 Coordinate definitions for analysis of circularly bent optical waveguides The rib width is $2a$, the bend radius (at the center of the rib) is ρ, and the cylindrical coordinates are r, φ and y. A second radial coordinate ξ is also defined.

In the curved geometry of Figure 4.1, the waveguide cross-section is invariant with respect to the angular coordinate φ, so the mode fields of the bent waveguide are defined with respect to an angular phase constant rather than a Cartesian one as in (2.11):

$$\psi(r,y,\varphi)=\psi(r,y)e^{j\beta\rho\varphi} \quad \textbf{(4.5)}$$

where β is a circumferential propagation constant, equal to the angular phase constant divided by ρ. We assume that, according to the arguments of section 2.9, a scalar treatment will provide a satisfactory analysis. Transforming a scalar version of the Helmholtz equations (2.9.a&b) into our new coordinates, we have:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right)+\frac{\partial^2\psi}{\partial y^2}+\frac{1}{r^2}\frac{\partial^2\psi}{\partial\varphi^2}+k_o^2n^2(r,y,\varphi)\psi=0 \quad \textbf{(4.6)}$$

where $\psi=\psi(r,y,\varphi)$. Applying the mode field definition (4.5) in (4.6), we obtain the following Helmholtz (eigen)equation for mode fields:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right)+\frac{\partial^2\psi}{\partial y^2}+k_o^2n^2\psi-\beta^2\left(\frac{\rho}{r}\right)^2\psi=0 \quad \textbf{(4.7)}$$

where now $\psi=\psi(r,y)$ and $n=n(r,y)$. Let us now define a new variable, u, according to the following transformation:

$$u(r,y)\equiv\sqrt{r}\psi(r,y) \quad \textbf{(4.8)}$$

Applying (4.4) and (4.8) in (4.7) we obtain, after some manipulation:

$$\frac{\partial^2u}{\partial\xi^2}+\frac{\partial^2u}{\partial y^2}+\left(\frac{1}{4}-\beta^2\rho^2\right)\frac{u}{(\xi+\rho)^2}-\frac{1}{(\xi+\rho)}\frac{\partial u}{\partial\xi}+k_o^2n^2u=0 \quad \textbf{(4.9)}$$

Next, we make some approximations. We assume that $\rho>>a$ and therefore $\rho>>\xi$ in the region where the bound mode fields are non-negligible. In this case:

$$\frac{1}{(\rho+\xi)^2}=\frac{1}{\rho^2}\frac{1}{(1+\xi/\rho)^2}\cong\frac{1}{\rho^2}\left(1-\frac{2\xi}{\rho}\right) \quad \textbf{(4.10)}$$

and:

$$\frac{1}{(\rho+\xi)}\frac{\partial u}{\partial \xi}\cong\frac{1}{\rho}\frac{\partial u}{\partial \xi}<<\frac{\partial^2 u}{\partial \xi^2},\frac{\partial^2 u}{\partial y^2} \tag{4.11}$$

Applying (4.10) and (4.11) in (4.9) gives us:

$$\frac{\partial^2 u}{\partial \xi^2}+\frac{\partial^2 u}{\partial y^2}+k_o^2\left\{n^2-n_e^2\left(1-\frac{2\xi}{\rho}\right)+\frac{1}{4}\frac{1}{(\rho k_o)^2}\left(1-\frac{2\xi}{\rho}\right)\right\}u=0 \tag{4.12}$$

where n_e is, again, the effective index, given by β / k_o. Defining a new variable:

$$R\equiv\rho k_o \tag{4.13}$$

we can write:

$$n^2-n_e^2\left(1-\frac{2\xi}{\rho}\right)+\frac{1}{4}\frac{1}{(\rho k_o)^2}\left(1-\frac{2\xi}{\rho}\right)=n^2+\frac{1}{4R^2}+\frac{2}{\rho}\left[n_e^2-\frac{1}{4R^2}\right]\xi-n_e^2 \tag{4.14}$$

and therefore (4.12) may be rewritten as:

$$\frac{\partial^2 u}{\partial \xi^2}+\frac{\partial^2 u}{\partial y^2}+k_o^2\left\{N^2(\xi)-n_e^2\right\}u=0 \tag{4.15}$$

where:

$$N^2(\xi)\equiv n^2+\frac{1}{4R^2}+\frac{2}{\rho}\left[n_e^2-\frac{1}{4R^2}\right]\xi \tag{4.16}$$

Again assuming that $\rho>>a$, it is easy to show that in the region of interest, (4.15) is well-approximated by:

$$\frac{\partial^2\psi}{\partial\xi^2}+\frac{\partial^2\psi}{\partial y^2}+k_o^2\{N^2(\xi)-n_e^2\}\psi=0 \tag{4.17}$$

Note that equation (4.17) has the same form as the standard scalar modal Helmholtz equation for straight waveguides (2.93), except that we now have a modified index of refraction given by (4.16). Effectively, bending the waveguide is equivalent to transforming its refractive index profile as shown in Figure 4.2. The transformation becomes more pronounced as ρ is reduced.

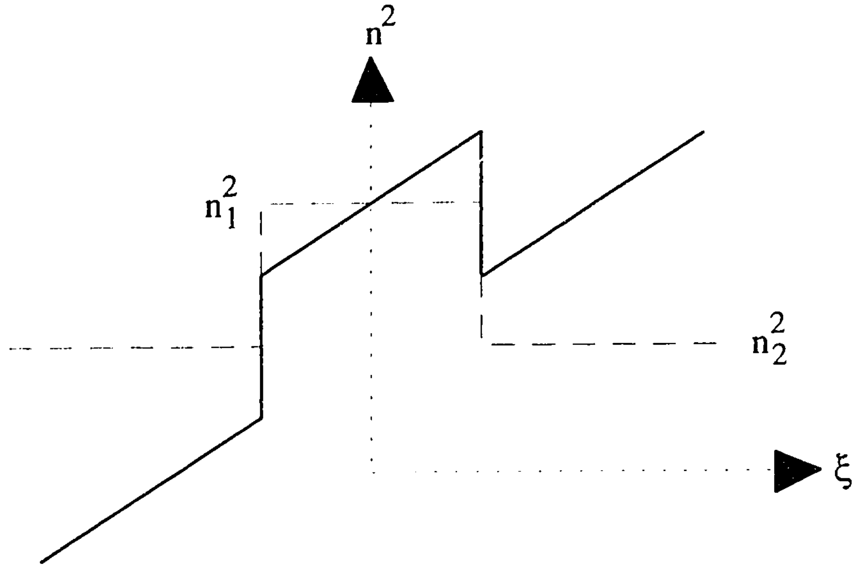


Figure 4.2 Transformation of the lateral (effective) index profile of a rib waveguide by a circular bend. The squared effective index profile of the straight guide is shown by the dashed line; n_1 is the rib effective index, n_2 is the lateral cladding effective index. The solid line shows the effectively modified index profile induced by the bend.

The basic effects of the bend on the guided mode fields are easily understood from Figure 4.2. In general, the mode fields are made asymmetric, narrowed and shifted toward the new high-index region at the outside of the bend, with a corresponding increase in their propagation constants. For these reasons, lateral offsets and width changes are often used to minimize transition losses at junctions between waveguides of different curvatures [68]. Note, further, that if the bend radius is sufficiently small and the slope of the transformed field profile sufficiently large, the inner rib wall may not play any part in guiding the optical fields. Modes guided only by the outer rib wall are called "Whispering Gallery" modes and were first described by Lord Rayleigh in relation to acoustic propagation in a curved gallery [69].

The mechanism by which bent waveguides can radiate light is also apparent from Figure 4.2. For small bend radii, the slope of the transformed index profile may be such that the exponential tails of the mode fields penetrate into the region of increasing index outside the rib. In this case, from the arguments of section 2.8, a slowly decaying sinusoidal transverse field characteristic is associated with the mode in this region. Such a field can transport power away from the rib and therefore leads to radiation losses. Rib waveguides which are deeply etched have very limited exponential tails, and therefore very small bend radii are required to induce radiation losses. This fact has been mentioned previously to justify the use of such waveguides to promote compactness of integrated structures. Radiation loss in waveguide bends has been studied extensively in the literature, and therefore is not specifically examined in this work.

Given the form of equation (4.17) and the arguments of sections 2.9 and 2.10, in strong guiding situations, the effective index treatment described in section (2.10) provides a valid analysis of the bent waveguide modes. We therefore seek to find solutions of a reduced-dimension form of (4.17):

$$\frac{d^2\psi}{d\xi^2} + k_o^2\left\{N^2(\xi) - n_e^2\right\}\psi = 0 \tag{4.18}$$

where a separation of the form of (2.96) has been assumed. Equation (4.18) can be solved using a modification of the finite differences scheme described in section 2.8. We again use the following centered-differences approximation to the second derivative of the field:

$$\frac{d^2\psi(\xi_i)}{d\xi^2} \cong \frac{1}{h^2}\left[\psi(\xi_{i+1}) - 2\psi(\xi_i) + \psi(\xi_{i-1})\right] \tag{4.19}$$

where h is a discretization interval and $\xi_i = ih$. Applying (4.19) to (4.18) we have, using (4.16):

$$\psi_{i-1} + \{k_o^2 h^2 A_i - 2\}\psi_i + \psi_{i+1} = \left(-k_o^2 h^2 n_e^2 B_i\right)\psi_i \tag{4.20}$$

where we have set $\psi_i = \psi(\xi_i)$, and:

$$A_i \equiv n_i^2 + \frac{1}{4R^2} + \frac{\xi_i}{2\rho R^2}, \quad B_i \equiv \frac{2\xi_i}{\rho} - 1 \tag{4.21}$$

where $n_i = n(\xi_i)$. Dividing through by $\bar{h}^2 \equiv k_o^2 h^2$, we obtain:

$$\frac{1}{\bar{h}^2}\psi_{i-1} + A_i'\psi_i + \frac{1}{\bar{h}^2}\psi_{i+1} = -B_i n_e^2 \psi_i \tag{4.22}$$

where:

$$A_i' \equiv A_i - \frac{2}{\bar{h}^2} \tag{4.23}$$

Equation (4.22) is suitably scaled for computation. We assume that the "window" of computation $x_{i=0} \rightarrow x_{i=n+1}$ is sufficiently large to properly contain all of the bound mode fields, and, taking $x_0 = x_{n+1} = 0$, we apply (4.22) at every node i ($i = 1, \ldots, n$) to obtain the matrix eigenequation:

$$[A]\{\psi\} = -n_e^2[B]\{\psi\} \tag{4.24}$$

where $\{\psi\}$ is the vector consisting of the values of ψ_i. The matrix $[A]$ is given by:

$$[A] = \begin{bmatrix} A_1' & \frac{1}{\overline{h}^2} & 0 & \dots & \dots & 0 \\ \frac{1}{\overline{h}^2} & A_2' & \frac{1}{\overline{h}^2} & 0 & \dots & 0 \\ & & & \dots & & \\ 0 & \dots & \dots & 0 & \frac{1}{\overline{h}^2} & A_n' \end{bmatrix} \qquad \textbf{(4.25.a)}$$

and :

$$[B] = \begin{bmatrix} B_1 & 0 & 0 & \dots & \dots & 0 \\ 0 & B_2 & 0 & 0 & \dots & 0 \\ & & & \dots & & \\ 0 & \dots & \dots & 0 & 0 & B_n \end{bmatrix} \qquad \textbf{(4.25.b)}$$

The eigensolutions of (4.24) are found using *LANZ*. The selection of the bound modes from among the various solutions of (4.24) is more complicated than in the case of straight waveguides. Clearly, those solutions with effective index eigenvalues less than the (true) refractive index of the cladding on the outside of the bend ("n_{cl}") cannot represent bound modes and may be rejected. Furthermore, from equation (4.16) we see that:

$$N^2(\xi = a) > n_{cl}^2\left(1 + \frac{2a}{\rho}\right) \qquad \textbf{(4.26)}$$

since $n_e > n_{cl}$ for a bound mode. Therefore, bound mode solutions of (4.24) can only lie in the range of eigenvalues such that:

$$n_e > n_{cl}\sqrt{1 + \frac{2a}{\rho}} \qquad \textbf{(4.27)}$$

This criterion is usually sufficient to eliminate the solutions of (4.24) which do not correspond to bound modes. An inspection of the field solutions themselves can always serve as a check, since, as mentioned earlier, leaky mode solutions of (4.24) have sinusoidal tails in the outer cladding. The outer cladding thickness must be made fairly large, for computational purposes, for this effect to be clear.

Note that a simplification of equation (4.18) is possible which results in an eigenequation of the form:

$$[A]\{\psi\} = -n_e^2\{\psi\} \tag{4.28}$$

If $\rho >> a$, and if we use $n_e \cong n$, then (4.16) may be approximately simplified to:

$$N^2(\xi) \cong n^2\left(1+\frac{2\xi}{\rho}\right) \tag{4.29}$$

This form of the transformed index profile may be advantageous, since the solution of (4.28) is numerically simpler than the solution of (4.24). We have found that both approaches, i.e. equations (4.16) & (4.24) and equations (4.28) & (4.29), yield almost identical results in most cases.

4.2 Modal Characteristics of Bent Multi-Mode Waveguides It is well-known that bending a single-mode waveguide causes the propagation constant/effective index of the mode to increase. This is because the guided intensity is shifted toward the outside of the bend, where the transformed effective refractive index of the guide is higher than the true index of the corresponding straight waveguide. However, the impact of a bend on the higher-order modes of a multi-mode waveguide has only recently begun to be investigated. A recent study by Kumar *et al.* [70] revealed the surprising fact that, in some circumstances, the propagation constant of the first-order mode of a two-mode waveguide may actually be ***reduced*** by introducing a circular bend. They explained this phenomenon as follows. The circular bend actually has two effects on the guided mode fields: they are shifted toward the outside of the bend, and they are deformed (made asymmetric). These two factors have opposite effects on the distribution of the mode power. The shift tends to move the power toward the outside of the bend, while the deformation tends to pull the intensity toward the inside of the bend. The net result of these competing effects depends strongly on the mode profile of the corresponding straight waveguide. In Figure 4.3, we show the field profiles of the fundamental and first-order modes at 630 nm wavelength, calculated using (4.16) and (4.18)-(4.27), of a 7 μm-wide planar waveguide of index 1.4623 and (symmetric) cladding index 1.457, bent to a radius of 7 mm. This is very close to the example studied by Kumar *et al.* The cladding thickness used in the calculations was 5 μm; therefore the waveguide center is at 8.5 μm. The shifting and asymmetry of both mode profiles is clearly seen. For the fundamental mode, the shift to the outside of the bend is the dominant effect, and hence the greater part of the mode intensity is located toward the outside of the bend; see Figure 4.4, which shows the intensity

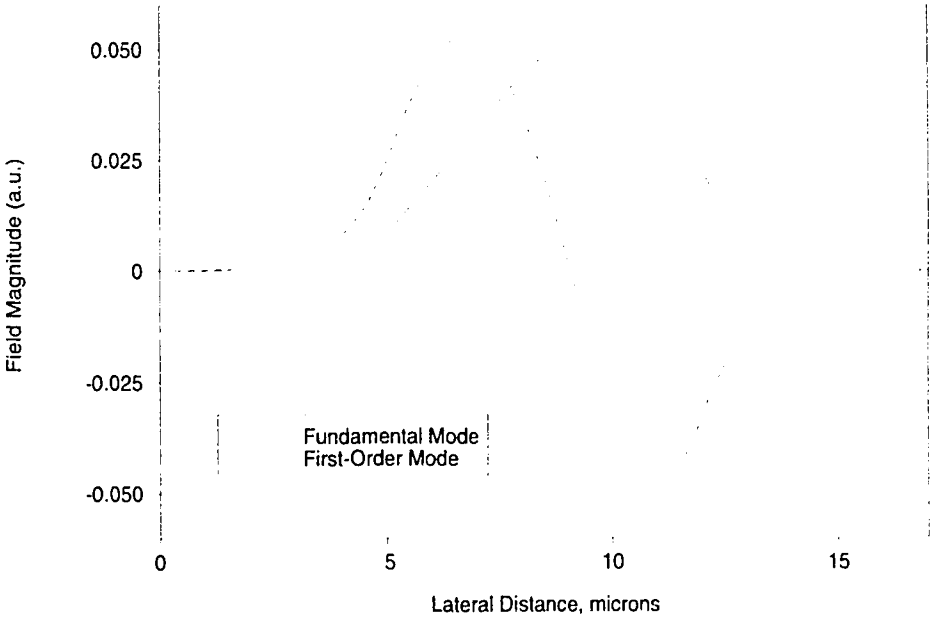


Figure 4.3 Mode field profiles of a planar waveguide, bent to a radius of 7 mm. The core is 7 μm wide and has index 1.4623, while both claddings have index 1.457. The core center is located at 8.5 μm.

profiles of the modes. Since most of the mode intensity is located in a region of increased (transformed) refractive index, the propagation constant of the fundamental mode is higher than in the straight waveguide. However, in the case of the first-order mode, the deformation of the field profile dominates, and the greater part of the mode intensity is located toward the inside of the bend, in a region of reduced effective refractive index. Therefore, the propagation constant of the first-order mode is reduced.

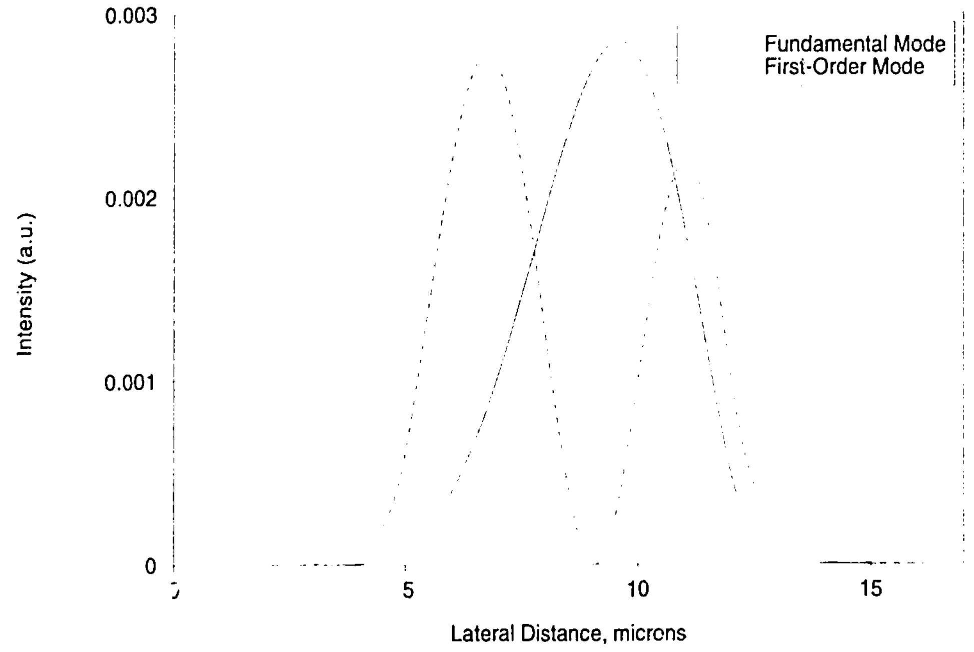


Figure 4.4 Intensity profiles of the mode fields shown in Figure 4.3.

Kumar *et al.* found that this effect, i.e. opposing changes in the propagation constants of fundamental and first-order modes with the introduction of a bend, is restricted to waveguides wherein the guided mode intensity in the claddings is small. Where significant evanescent field penetration into the outer cladding exists, the intensity in this region of high transformed index may have a dominant effect on the change in the propagation constant of the guided mode. Since high-order modes have more pronounced evanescent fields than low-order modes, the propagation constant of the first-order mode of such a waveguide may increase, rather than decrease, with bending. Therefore, the phenomenon of opposing bend-induced changes in the propagation constants of fundamental and first-order modes is restricted to waveguides operating in the strong-guiding regime.

Kumar *et al.* examined only two-mode waveguides. Our investigations show that, for strongly guiding structures, the propagation constants of second- and higher-order modes also tend to be reduced with bending, but to a smaller degree with increasing mode order. In general, high-order modes are less perturbed by a bend than are modes of low order.

Note that if the propagation constant of the fundamental mode of a waveguide is increased by a bend, while that of the first-order mode is reduced, then the beat length defined by equation (3.10) is reduced. Given this fact, it is tempting to consider the possibility that bending a multi-mode waveguide can reduce the device lengths required to achieve the self-imaging resonance effects described in Chapter 3. However, two factors prevent the general application of this concept. The introduction of a bend removes the quadratic dependence of the mode propagation constants on the lateral mode order. Since this dependence is the basis of resonant self-imaging, bending a multi-mode waveguide destroys the self-imaging effect, and hence only a two-mode device could operate effectively as a coupler. Second, a bend renders the mode fields of a waveguide asymmetric. While, by itself, this has no effect on bar-state resonances, it means that true cross-coupler states could not be achieved even if the quadratic phase relationships of the mode propagation constants were preserved. This can be seen be referring to equation (3.13). At the cross-coupling resonance lengths, the even modes are out of phase with the odd modes. Where the mode fields are symmetric as in equation (3.1), this situation effectively produces a superposition of lateral mirror images of the modes, as shown by equations (3.14)-(3.19). If the mode profiles are asymmetric, this is not the case and an efficient transfer of power, from one side of the multi-mode waveguide to the other, is precluded. Given these arguments, *in general* only a two-mode bent waveguide could be used for self-imaging, and it could only be effectively operated in the bar-state.

We have noted that the high-order modes of a waveguide are less perturbed by the introduction of a circular bend than are the low-order modes. The high-order modes tend to be of broader lateral extent, and more effectively "fill" the waveguide, than the modes of low order. The competing bend-induced changes to the modes, shift and deformity, are more balanced and the net alterations of the mode fields, and their propagation constants, are reduced. A similar effect is expected with respect to wavelength. The mode fields at long wavelengths are broader than those at shorter wavelengths. Therefore, we anticipate greater shifts in the propagation constants, and greater bend-induced asymmetry in the lateral mode profiles, at short wavelengths than at longer wavelengths, at a fixed bend radius. Accordingly, while bending a strongly guiding waveguide reduces the beat length defined by equation (3.10), we expect this reduction to be more pronounced at a short wavelength than at a longer one. The ratio of the beat lengths at two wavelengths can therefore be altered by introducing a bend, and adjusted to a desired value, e.g. 1.5. The strongly affected short wavelength can be operated in the bar state, in accordance with the limitations described earlier, while the longer wavelength may be sufficiently unaffected by the bend to be operated in the cross state. Therefore, compact couplers effectively based on two-mode interference might be optimized for operation as two-wavelength (de)multiplexers by bending them to an appropraite radius. It is this concept that we examine in this chapter.

4.3 Material Parameters and Waveguide Design

The modeled results presented in the next section, which validate the concept described above and lead to a specific design, were obtained for a particular waveguide configuration and material system, SiO_2/SiON glass, which was subsequently used to fabricate devices. Therefore, we first briefly describe these materials and how they were fabricated, characterized and processed.

Silica-based glass films were produced by Plasma Enhanced Chemical Vapour Deposition (PECVD). In this process, vapour-phase reactants are delivered to the surface of a substrate wafer, and react chemically to produce solid films of a desired composition. An RF glow discharge plasma provides energy to activate the reaction. The chamber temperature and pressure, RF power, and gas flow rates can all affect the reaction and the physical and chemical characteristics of the resulting film. Our process used 29 W of RF power, and a chamber temperature and pressure of 349 degrees Celsius and 1000 mtorr, respectively. The following vapour constituents were used: silane (7.5 sccm), helium (700 sccm), nitrogen (100 sccm), nitrous oxide (100 sccm) and ammonia (variable). The flow rate parameter sccm refers to cubic centimeters per minute, at standard temperature and pressure. The ammonia flow rate determines the fractional incorporation of nitrogen in the films; higher nitrogen levels were found to increase the refractive index. We used ammonia flow rates varying from 15-30 sccm. The optical characteristics of the resulting films were determined by broadband (200-1700 nm) ellipsometry, and modeled using a six-term Cauchy relation for the refractive index:

$$n(\lambda) = A + \frac{B}{\lambda} + \frac{C}{\lambda^2} + \frac{D}{\lambda^3} + \frac{F}{\lambda^4} + \frac{G}{\lambda^5} \qquad (4.30)$$

The terms for the Cauchy formula (4.30), obtained at different ammonia flow rates, are shown in Table 4.1. In some cases, the coefficients are averages of the fitted values of different film samples. The dispersion curves given by (4.30) using these values are shown in Figure 4.5. In all cases, the curves are relatively flat beyond ~850 nm: the refractive index dispersion is less than -0.007 from 850 nm to 1550 nm. In addition, none of the films showed any measurable absorption over the entire measured spectral range. These films are therefore ideal for implementing waveguides for operation at all of the standard fiber optic wavelength bands.

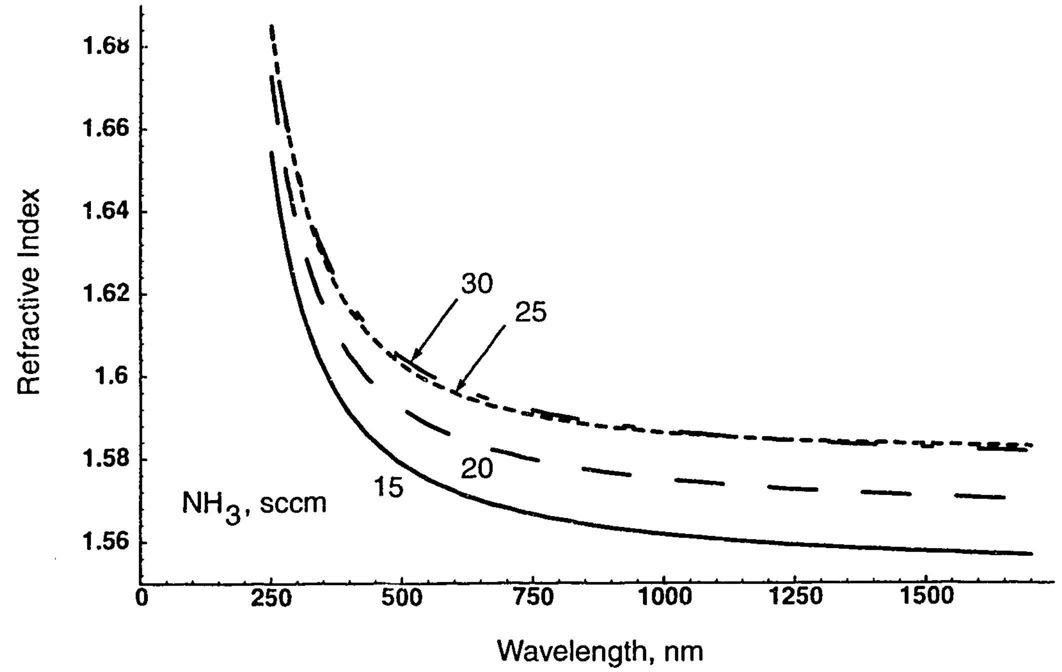
NH_3 Flow, sccm	**Cauchy *A***	**Cauchy *B***	**Cauchy *C***	**Cauchy *D***	**Cauchy *F***	**Cauchy *G***
15	1.551	8.156E-3	1.64E-3	0	1.31E-4	0
20	1.564	1.055E-3	-4.70E-4	1.13E-3	8.00E-6	0
25	1.581	0	5.11E-3	0	8.70E-5	0
30	1.576	8.85E-2	-1.20E-3	0	0	-4.60E-5

Table 4.1 Coefficients for Cauchy refractive index formula, equation (4.30), of SiON films with various NH_3 flow rates. The coefficients are fitted from ellipsometric measurements.

The refractive index at all wavelengths increases as the ammonia flow rate is raised from 15 to 25 sccm; beyond this point, no further increase is evident. Higher flow rates of the other nitrogen-bearing gases may achieve further increases in the refractive index. At four wavelengths (630 nm, 850 nm, 980 nm & 1550 nm), the refractive indices were fitted to a quadratic function of the ammonia flow rate (x):

$$n^{\lambda}(x) = a + bx + cx^2 \quad \textbf{(4.31)}$$

using Newton's method (for nonlinear systems) to match $n(x)$ to the values from equation (4.30) at x=15, 20 & 25 sccm. The fitted coefficients a, b, and c at the four wavelengths above are listed in Table 4.2. The resulting functions (4.31) at each wavelength are plotted in Figure 4.6. The function (4.31), with the coefficients of Table 4.2, should be taken as a reasonably accurate model only in the range of ammonia flow rates 15-25 sccm. The index values at 30 sccm have been excluded from the fit, as a quadratic function would be inadequate to model the index over the whole measured range of flow rates.



<u>Figure 4.5</u> Dispersion curves of SiON films grown using NH_3 flow rates of 15, 20, 25 and 30 sccm, plotted using equation (4.30) with the coefficient values listed in Table 4.1.

Wavelength, nm	*a*	*b*	*c*
630	1.514	4.6E-3	-5.6E-5
850	1.510	4.3E-3	-4.7E-5
980	1.510	4.1E-3	-4.2E-5
1550	1.507	3.9E-3	-3.4E-5

Table 4.2 Coefficients of equation (4.31), determining the index of refraction as a function of NH_3 flow rate in sccm, at four different wavelengths.

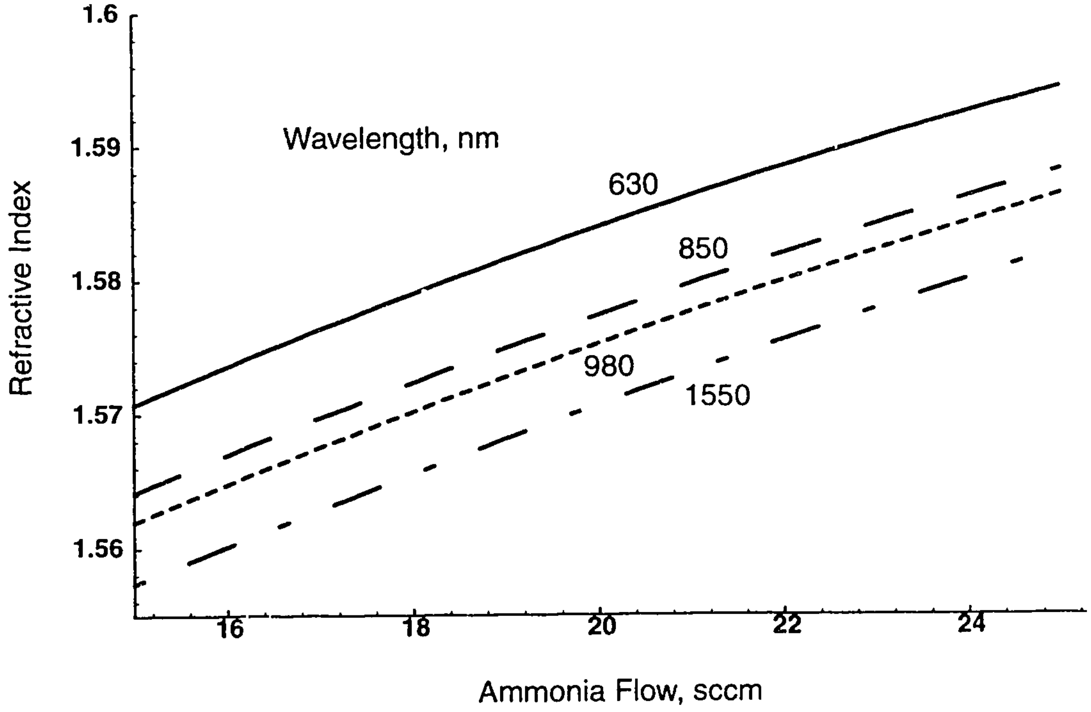


Figure 4.6 Approximate values of refractive index versus ammonia flow in sccm, at four different wavelengths. The curves are generated using equation (4.31) and the fitted coefficients of Table 4.2.

From equation (4.31) and Table 4.2, we can select an ammonia flow rate to yield a chosen index of refraction, at a particular wavelength. We choose a nominal index of refraction of 1.565, and seek the flow rate which most closely produces this index at both 850 nm and 1550 nm wavelength; these are the operating wavelengths of the devices considered in this chapter. An ammonia flow rate of 16.5 sccm is selected to yield a predicted film index of 1.568 at 850 nm, and 1.562 at 1550 nm, a good compromise. The index repeatability of the PECVD process, from run to run, is observed to be about ±0.005; therefore, at both wavelengths, the film index should be repeatably 1.565 ± 0.01.

The following waveguide configuration is used for device design and implementation. The SiON film described above, with nominal refractive index 1.565, is grown to a thickness of 1.0 μm, above a 5.0 μm-thick layer of SiO_2, produced by a PECVD process with the following parameters: 5.5 sccm silane, 600 sccm helium, 500 sccm nitrous oxide and 100 sccm nitrogen, at 29 W RF power and 1000 mtorr pressure. The index of this film, which serves as a cladding, is about 1.465 at 850 nm, and 1.459 at 1550 nm. Again, a repeatability of about ±0.005 is assumed. Strongly guiding rib waveguides are then defined by reactive ion etching to a depth of 1.0 μm, i.e. fully through the high-index SiON layer. See Figure 4.7. The reactive ion etching process used 80% CHF_3 and 20% CF_4, at a total pressure of 40 mtorr with 100 W of RF power. The etch rate was ~1 μm/27 minutes. Good control of the photolithographic process was achieved using positive photoresist (#504), spun-on at 500 rpm for 10 s followed by 6000 rpm for 30 s, to achieve a thickness of about 1.1 μm, then exposed for 1.6 s and developed for 20 s. An accuracy in the dimensions of photoresist-patterned features of ±0.1 μm, with respect to the photomask feature dimensions, was repeatably achieved.

4.4 Concept Validation and Device Design

We first verified the validity of the separation relation (2.96) for the (straight) rib waveguide above, at both 850 nm and 1550 nm, and for various rib widths, using equation (3.48). The (quasi-TE) modes of the bent waveguides were then calculated using (4.16) and (4.18)-(4.27), in conjunction with the effective index method of section 2.10, for application in a modal analysis similar to that described in Chapter 3. The ratio of beat lengths at the two wavelengths 850 nm and 1550 nm is, according to equation (4.2), naturally far from the optimum value of 1.5, and thus this choice of wavelengths provides a vivid demonstration of our concept.

Access waveguides 3 μm wide were used in all cases; this width supports three modes at 850 nm and a single mode at 1550 nm. A coupler width of 6 μm, supporting six modes at 850 nm and two modes at 1550 nm, was also chosen. The ratio of beat lengths at the two wavelengths in a 6 μm-wide straight waveguide is

1.76, close to the prediction of equation (4.2). In Figure 4.8, we show the calculated beat lengths defined by equation (3.10), at each wavelength, and their ratio, as a function of radius of curvature. The desired synchronization, i.e. a ratio of 1.5, is predicted at a radius of 0.925 mm. This radius is assumed, for both the coupler and the access waveguides, in the following results.

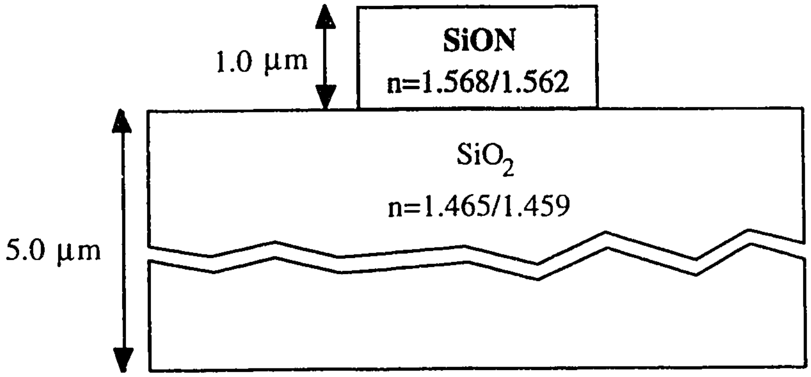


Figure 4.7 Silica-based glass waveguide system for design and implementation of bent multi-mode (de)multiplexing coupler devices. The two films are grown on Si wafers by PECVD. The shown values of refractive index are for 850 nm and 1550 nm wavelength, respectively.

The performance of the coupler was simulated in the demultiplexer configuration. A single input waveguide, bent at a radius of 0.925 mm, was assumed centered on the outside half of the coupler, i.e. the input waveguide and the coupler form a continuous sidewall on the outside of the bend. See Figure 4.9, which shows a scanning electron micrograph of a fabricated device. Two output waveguides are located at the end of the coupler, each bent at 0.925 mm radius. A blunted junction tip of width 0.7 μm is assumed to separate the output waveguides at the coupler output junction. The following parameters were varied in the simulations: the (circumferential) length of the coupler, and the junction angle of each of the two output waveguides with respect to the local propagation axis of the coupler at the output junction plane. Unlike the coupler model in Chapter 3, the total junction angle is not necessarily assumed equally divided between the two output waveguides.

A computer model, based on the modal analysis method and similar to that described in section 3.4, was used to calculate the performance of the device in the demultiplexer configuration. As only two modes are sustained in the coupler at 1550 nm, only restricted resonance phenomena need to be considered at this wavelength. Furthermore, the mode field profiles at this wavelength are not significantly perturbed by the 0.925 mm radius bend. Therefore, cross-state operation can be effectively achieved at 1550 nm. At 850 nm, six coupler modes are sustained, and the low-order

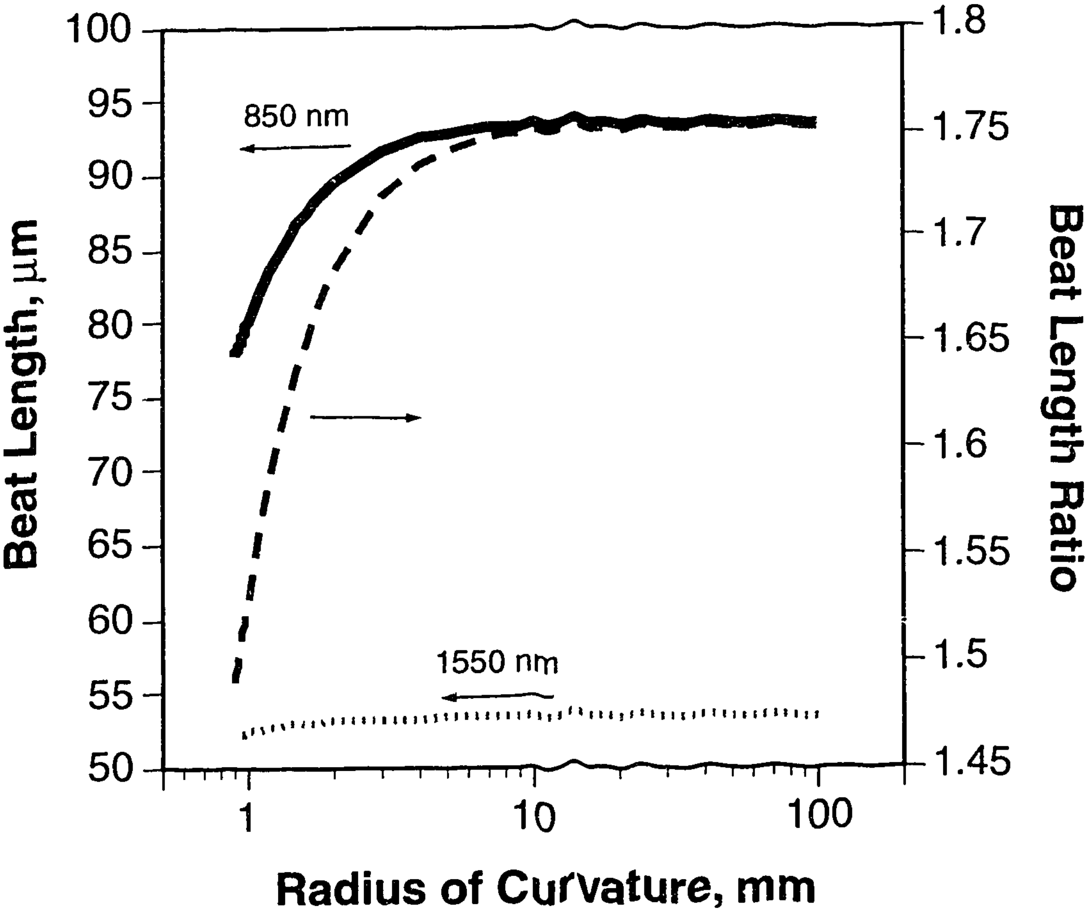


Figure 4.8 Calculated beat lengths L_π of the 6 μm-wide coupler, at 850 nm and 1550 nm, and their ratio, as a function of radius of curvature. A ratio of 1.5 is predicted at a radius of 0.925 mm.

modes are quite asymmetric at the design radius. Therefore, given our previous arguments, the device must operate in the bar-state at 850 nm, and if the excitation of the second- and higher-order modes is small so that only the fundamental and first-order modes are significantly excited, a restricted resonance situation effectively exists at 850 nm as well. The (minimum) optimum circumferential length of the bent coupler is then:

$$L_{\text{coupler}} = 3L_{\pi}^{1550} = 2L_{\pi}^{850} \quad (\text{restricted resonance}) \qquad \textbf{(4.32)}$$

or, from the calculations of Figure 4.8, ~157 μm. The performance parameters at each wavelength are contrast and insertion loss. Refer to Figure 3.2; P_1 and P_2 are taken such that the device is in the bar-state at 850 nm, and in the cross-state at 1550 nm.

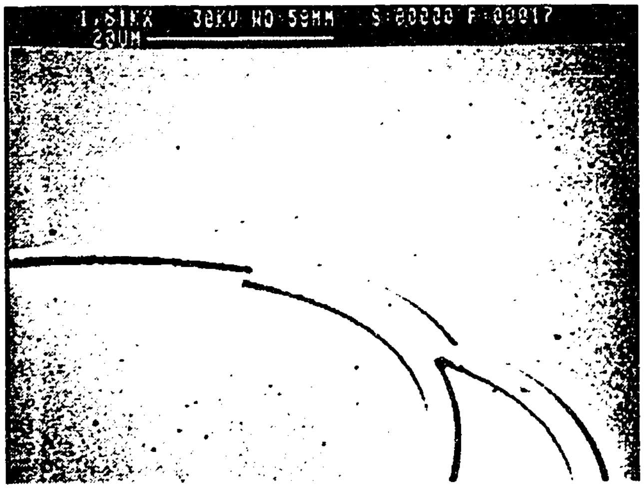


Figure 4.9 Scanning electron micrograph of a fabricated bent multi-mode waveguide (de)multiplexing coupler. All waveguides are bent at a radius of 0.925 mm. There is no junction angle between the 3 μm-wide input waveguide and the 6 μm-wide coupler. The 3 μm-wide output waveguides are separated by a 0.7 μm-wide blunted junction tip. Various angular offsets of the two output waveguides, relative to the local propagation axis of the coupler at the output junction, are examined in the model; here, offsets of 7 degrees (bar-state output) and zero (cross-state output) are used.

We first consider the operation of the device at 850 nm. In this case, the input waveguide is multi-moded, and the distribution of power among these modes is expected to have an impact on the performance of the coupler. We consider this by assuming that the bent input waveguide is excited by a (planar) free-space gaussian beam with a field profile defined by:

$$\Psi(x) = \exp\left[-4x^2/w^2\right] \tag{4.33}$$

where w is the $1/e$ diameter of the field, and the beam center is at x=0. In Figure 4.10, we show the calculated values of the amplitude coefficients of the three modes of the bent input waveguide at 850 nm, as a function of w. The beam and rib centers are assumed coincident.

In all cases the power coupled into the first-order mode is negligible. The amplitudes of the fundamental and second-order modes vary considerably as the input beam diameter changes. At a beam field diameter of ~3 μm, the excited intensities of the fundamental and second-order modes reach a maximum and minimum, respectively. At 10 μm field diameter, the fundamental mode amplitude is considerably reduced, while that in the second-order mode is increased. Both of these cases are considered in the simulations to follow. The input field to the coupler in every case is constructed using the fundamental and second-order modes only, and they are taken as both in phase and out of phase (0 and π). It is reasonable to neglect the first-order mode, even when w=3 μm since the intensity in the second-order mode in this case is more than four times that in the first-order mode. The phase difference between the fundamental and second-order modes is determined by a beat length given by equation (3.10) with β_1 replaced by β_2, and which is ~10 μm here. The relative phase of the input waveguide modes at the coupler input is therefore a highly sensitive function of the length of the input waveguide. In fact, this relative phase can be adjusted by a small wavelength tuning if the input waveguide is long; this is examined in more detail later.

The amplitude coefficients of the six coupler modes at 850 nm are shown in Table 4.3 for the four different input conditions discussed above: a 3 μm and a 10 μm diameter gaussian field excitation of the input waveguide, with the fundamental and second-order input guide modes both in phase and out of phase for each case. The fraction of the power in the input waveguide which is coupled into the fundamental and first-order modes, and all other modes, is shown in Table 4.4.

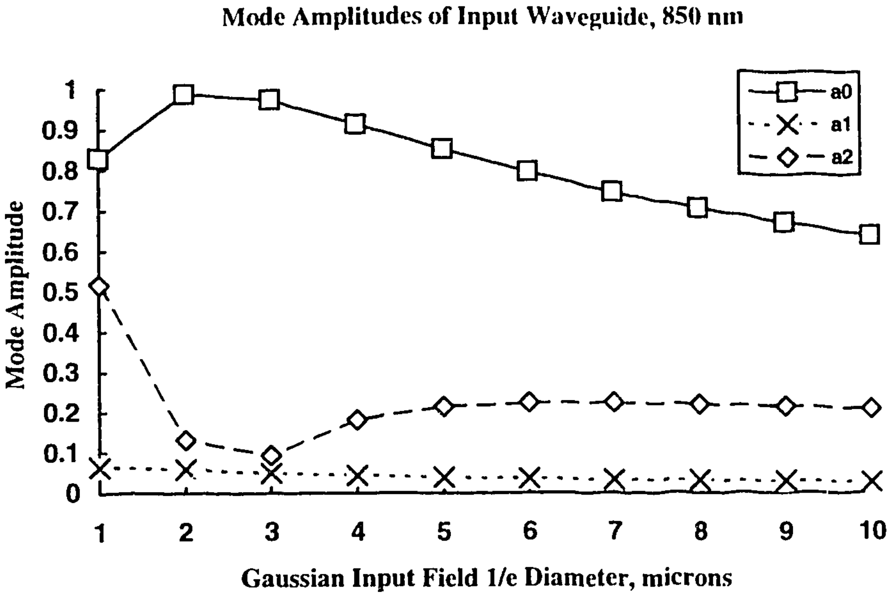


Figure 4.10 Calculated amplitude coefficients, a_i, of the 850 nm modes of the 3 μm input waveguide, bent at 0.925 mm radius, excited by a gaussian beam of 1/*e* field diameter *w* (*x*-axis). The beam and the waveguide rib centers are assumed coincident. Solid line with squares - fundamental mode; dotted line with crosses - first-order mode; dashed line with diamonds - second-order mode.

Due to the 0.925 mm radius bend, modes 2-5 at 850 nm are not phase resonant with the fundamental and first-order modes, so ideally we want no power coupled into these high-order modes. The performance of the coupler at 850 nm is expected to suffer as the amplitudes of modes 2-5 increases. From Table 4.4, the best performance is expected when the gaussian input has a 1/*e* diameter of 3 μm, and modes 0 and 2 of the input waveguide are out of phase. In fact, for both beam diameters, the best performance is anticipated when these modes are out of phase and the resultant field distribution at the coupler input is narrow.

Mode Order	3 μm, 0 phase	3 μm, π phase	10 μm, 0 phase	10 μm, π phase
0	0.836	0.863	0.765	0.854
1	0.501	0.476	0.509	0.428
2	0.156	0.121	0.194	0.081
3	0.106	0.072	0.135	0.024
4	0.128	0.007	0.260	0.135
5	0.037	0.086	0.179	0.224

Table 4.3 Amplitude coefficients of the six coupler modes, where the input waveguide is excited by a gaussian beam of the given $1/e$ diameter. The phase difference between the fundamental and second-order modes of the input waveguide at the coupler input junction is 0 or π as indicated.

Group of Modes	3 μm, 0 phase	3 μm, π phase	10 μm, 0 phase	10 μm, π phase
0, 1	0.948	0.971	0.844	0.912
2-5	0.052	0.027	0.156	0.075

Table 4.4 Intensities in the coupler modes 0 and 1, and 2-5, as a fraction of the total power in the input waveguide, for the four cases of the field diameter of the gaussian beam exciting the input waveguide, and the relative phase of the fundamental and second-order modes of the input waveguide at the coupler input.

Since there is only one access waveguide at the coupler input there is no need to introduce a junction angle. At the output, where we have two access waveguides, repeatable fabrication dictates that we introduce some angular separation of these guides, as discussed in Chapter 3. We choose a total angle, between the two output waveguides, of 7 degrees, in conjunction with a lateral separation of 0.7 μm at the junction tip. We next examine how this total angle should be divided between the bar- and cross-outputs. In Table 4.5, we show the calculated performance parameters of the coupler, contrast and insertion loss, at 850 nm, as a function of the angular offset of each output waveguide (θ_1: bar-output, θ_2: cross-output) in degrees, where a gaussian input beam of 3 μm diameter has been assumed, and modes 0 and 2 of the input waveguide are out of phase. The coupler length is the predicted optimum of 157 μm. Because the output waveguides sustain three modes at 850 nm, the performance at this wavelength is quite insensitive to the angular offsets between these guides. However, at 1550 nm, the access waveguides are mono-mode, and hence greater dependence with respect to angle is observed. See Table 4.6, which shows the cross-state contrast and insertion loss of the 157 μm-long bent coupler, excited by the single guided mode in the input waveguide, for the same three distributions of the total junction angle. The insertion loss increases markedly as the cross-output angular offset rises. Therefore, for all the following simulations and for the implemented devices, we choose angular offsets of 7 degrees for the bar-output and 0 degrees for the cross-output, as shown in Figure 4.9. This has the added effect of increasing the contrast at 1550 nm, since the coupling to the angled, mono-mode bar-output is degraded.

The modeled performance of the coupler at 1550 nm, as a function of circumferential coupler length, is shown in Figure 4.11. Since the coupler only sustains two modes at this wavelength, the performance parameters are roughly symmetric about the optimum length. A peak contrast in excess of 30 dB and a minimum insertion loss of 2.2 dB are predicted. In Figures 4.12 and 4.13, we show the calculated performance of the device at 850 nm. The input gaussian beam

θ_1 (degrees)	**θ_2 (degrees)**	**Contrast, dB**	**Insertion Loss, dB**
0	7	25.06	0.63
3.5	3.5	24.31	0.63
7	0	24.15	0.63

Table 4.5 Calculated performance of 157 μm-long coupler at 850 nm, as a function of the angular offsets of the bar-output (θ_1) and the cross-output (θ_2). The input waveguide is assumed excited by a 3 μm-diameter gaussian beam, and the fundamental and second-order modes of the input waveguide are assumed to be out of phase at the coupler input.

θ_1 (degrees)	θ_2 (degrees)	Contrast, dB	Insertion Loss, dB
0	7	29.66	4.36
3.5	3.5	31.28	2.73
7	0	31.80	2.20

Table 4.6 Calculated performance of 157 μm-long coupler at 1550 nm, as a function of the angular offsets of the bar-output (θ_1) and the cross-output (θ_2).

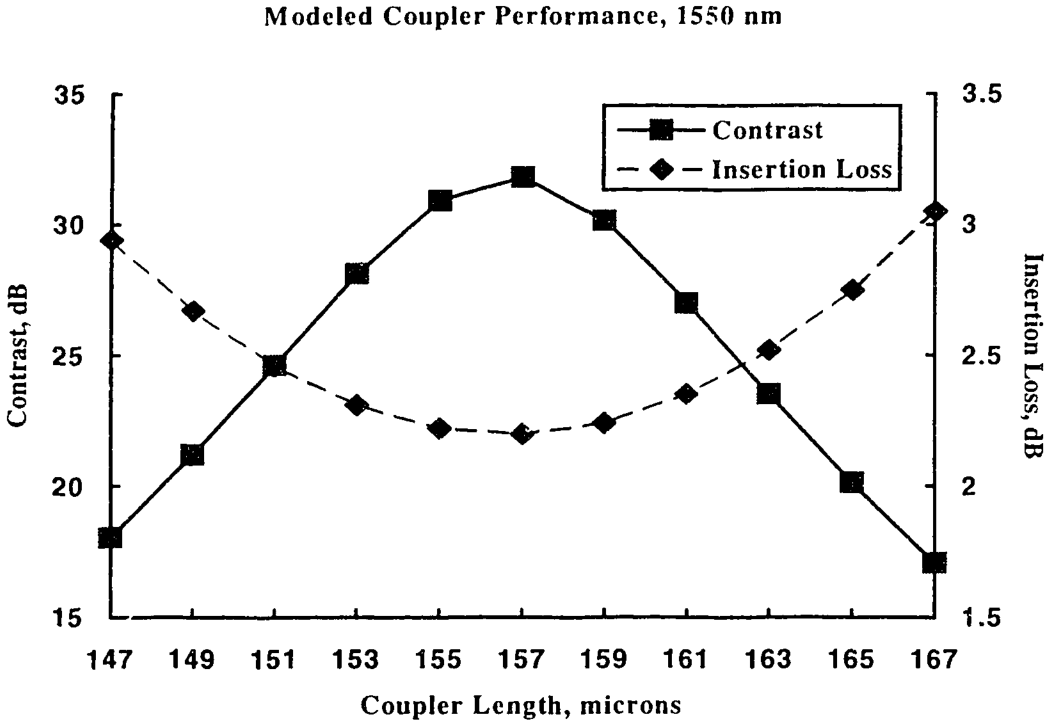


Figure 4.11 Modeled contrast (solid line, squares) and insertion loss (dashed line, diamonds) versus coupler length at 1550 nm.

diameter is 3 μm in Figure 4.12 and 10 μm in Figure 4.13; in both figures, the solid lines show the calculated characteristics when the fundamental and second-order modes of the input waveguide are in phase, while for the dashed lines they are taken out of phase. The optimum performance occurs at 155 μm coupler length. In general, the 850 nm characteristics are flatter than those at 1550 nm, since the beat length is larger at 850 nm, and the curves are no longer symmetric. The performance variations with respect to length are more pronounced in Figure 4.13 than in Figure 4.12. This may be attributed to the greater power resident in the non-resonant second- and higher-order modes. Finally, the performance variations with respect to the relative phases of the input waveguide modes are less pronounced in Figure 4.12. This is easily understood from Figure 4.10: when the gaussian beam radius is 3 μm, almost all the incident power is coupled into the fundamental mode of the input

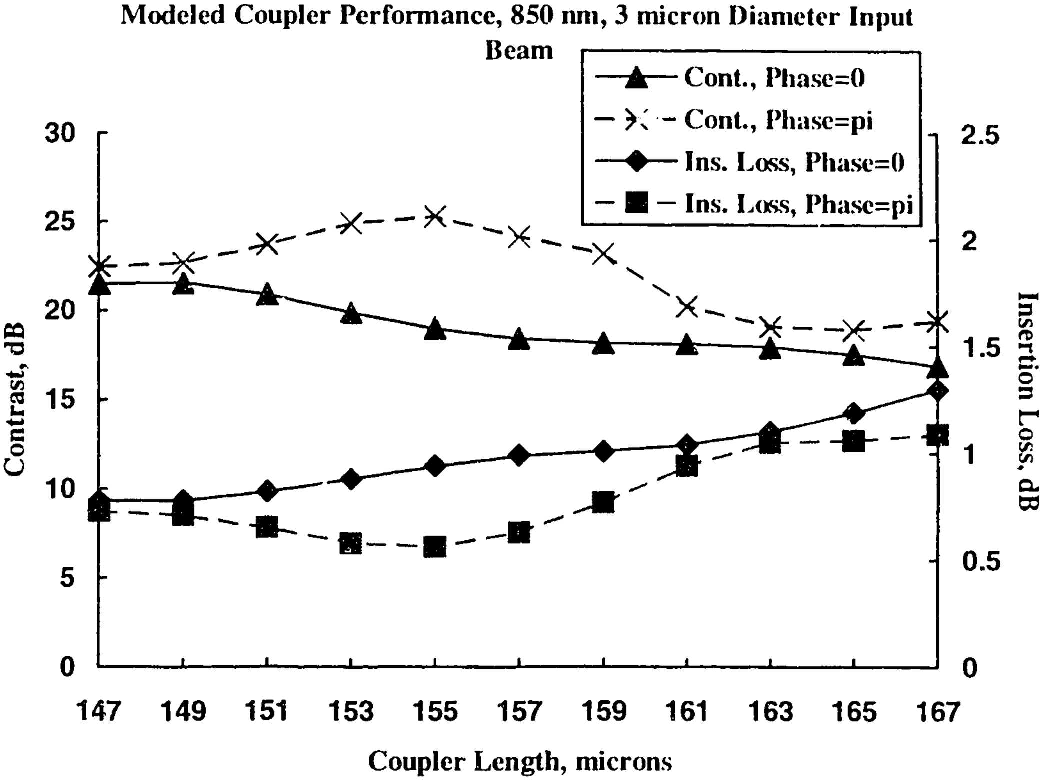


Figure 4.12 Modeled contrast (solid line, triangles - input modes in phase, dashed line, crosses - input modes out of phase) and insertion loss (solid line, diamonds - input modes in phase, dashed line, squares - input modes out of phase) versus coupler length at 850 nm. The input waveguide is assumed excited by a gaussian beam with a 1/*e* field diameter of 3 μm.

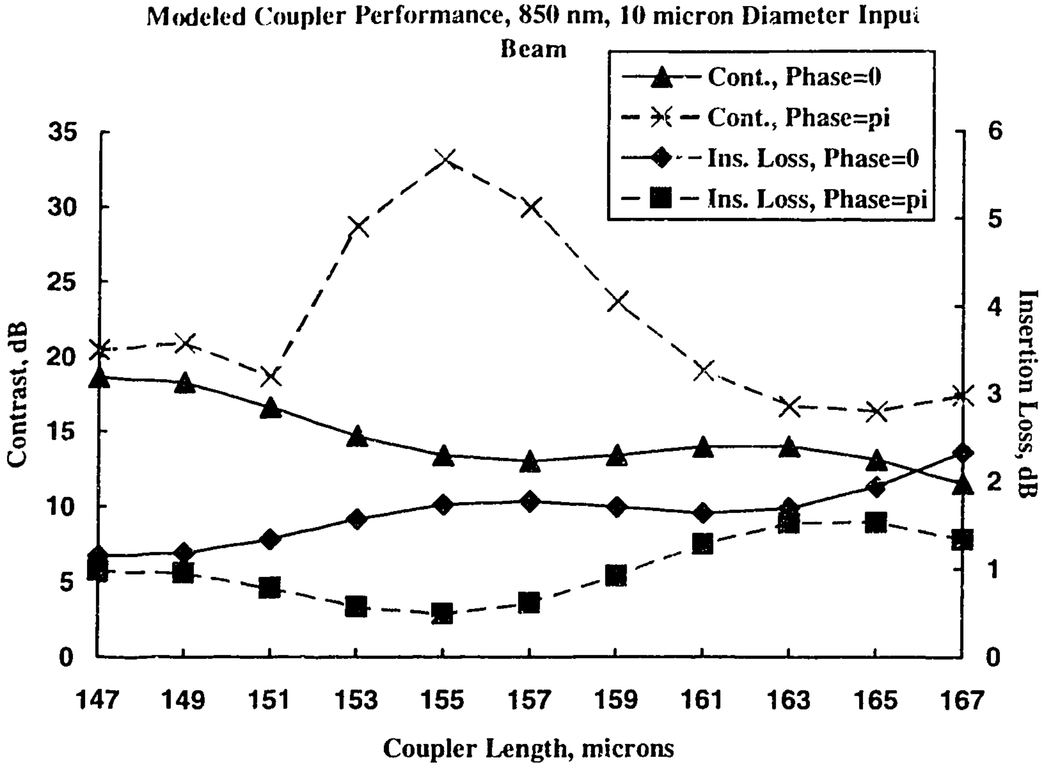


Figure 4.13 Modeled contrast (solid line, triangles - input modes in phase, dashed line, crosses - input modes out of phase) and insertion loss (solid line, diamonds - input modes in phase, dashed line, squares - input modes out of phase) versus coupler length at 850 nm. The input waveguide is assumed excited by a gaussian beam with a 1/*e* field diameter of 10 μm.

waveguide. Therefore, the phases of the input guide modes have little impact on the intensity distribution at the coupler input plane. Excellent peak performance is predicted at 850 nm: for both input beam diameters, a peak contrast in excess of 25 dB and a minimum insertion loss of less than 0.7 dB are expected.

The performance advantage and size reduction achieved by bending the coupler can be appreciated by examining the performance of a coupler with the same configuration as described above (i.e. same waveguide widths and junctions geometries), but composed of straight waveguides. Simulations show that a straight 6 μm-wide coupler cannot achieve greater than 20 dB (calculated) contrast, at both wavelengths, at a coupler length less than ~1.1 mm. In this particular case, therefore, bending the coupler to an optimum radius achieves an effective length reduction of more than seven times. The length reduction achieved with other pairs of wavelengths and using other waveguides will vary, depending on the exact ratio of the beat lengths, at the chosen wavelengths, for a straight coupler. However, optimizing the ratio of beat lengths by bending a multi-mode coupler can always yield high performance (de)multiplexing operation using a compact device (<~200 μm), if the operating wavelengths λ_1 and λ_2 are such that $\lambda_2/\lambda_1>1.5$.

4.5 Experimental Results Devices were fabricated on silicon substrates, by the processes described in section 4.3, using an e-beam-written chrome-on-quartz photomask. The couplers, including input and output waveguides, were cleaved to lengths of ~5 mm, and tested, in the demultiplexer configuration, using the experimental arrangement shown in Figure 4.14. A Ti:Sapphire laser, pumped by an argon laser, and wavelength-tuned by computer control, provided light around 850 nm wavelength. A polarizing beam-splitter followed by a half-wave plate produced a normally TE-polarized beam from the Ti:Sapphire laser, while rotating the half-wave plate by 45 degrees could produce TM light. A fiber collected the light reflected from the beam-splitter for direction to a wavemeter, which measured the wavelength of the Ti:Sapphire output. The main beam was focused into a single-mode fiber using a 10X objective lens, and combined with TE-polarized 1550 nm light from a DFB laser using a fiber multiplexer. The combined beam was collimated in free space using a 10X objective and focused into the samples using a 20X objective. The samples were mounted on a three-axis micropositioning stage. A butt-coupled multi-mode fiber collected the output light from the samples, for measurement using optical power meters. A He-Ne laser beam coupled into the 1550 nm port of the fiber multiplexer was used for the initial alignment of each sample.

Contrast was determined by measuring the output optical powers from the output waveguides of each sample, with one of the two laser sources illuminating the sample. Insertion loss was measured by comparison with 3 μm-wide waveguides which followed the same curved path, over the same length, as the input and output waveguides of the couplers. A number of samples at each coupler length were tested, and the average values are presented in what follows. In Figure 4.15, we show the measured values of contrast and insertion loss at 1550 nm, and the best values for contrast at ~850 nm. In accordance with the previous arguments, the best contrast at

850 nm is obtained when the fundamental and second-order modes of the coupler are out of phase. This condition was obtained, in all cases, by a slight wavelength tuning (~5 nm) around 850 nm. We shall comment further on this later. The 1550 nm characteristics are obtained by a least-squares quadratic fit to the experimental points, while the 850 nm contrast values are joined by a cubic splines fit. The optimum contrast at 850 nm is observed at a coupler length of 145 μm, while that at 1550 nm occurs at about 150 μm. This slight detuning indicates that a beat length ratio of ~1.47 was achieved. The peak contrasts are reduced from the predicted values by about 10 dB in both cases; this is typical for couplers implemented in deeply etched ribs. Though the input beams are focused to an estimated diameter of ~3 μm, the 850 nm contrast characteristic is more sharply peaked than the predicted curve in Figure 4.12, bearing a greater resemblance to the predicted peak contrast characteristic for a 10 μm diameter input beam (Figure 4.13). This indicates that the power in the non-resonant second- and higher-order coupler modes, at the coupler output, was higher than expected. This could be explained by inter-mode power scattering caused by sidewall roughness. Fabrication methods which reduce sidewall roughness, e.g. in [6], could therefore be expected to push the device performance closer to the

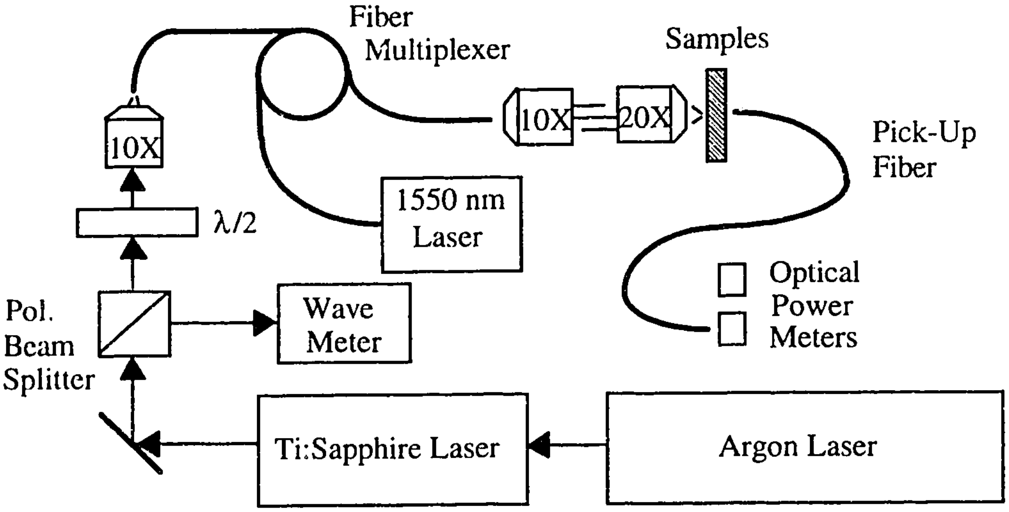


Figure 4.14 Schematic of the experimental arrangement used for testing the bent multi-mode coupler (de)multiplexer devices.

calculated values. At 145 μm coupler length, almost 20 dB and 14 dB contrast are observed at 1550 nm and 850 nm, respectively. The measured insertion losses at 1550 nm were somewhat lower than predicted: 1.4 dB at 145 μm coupler length.

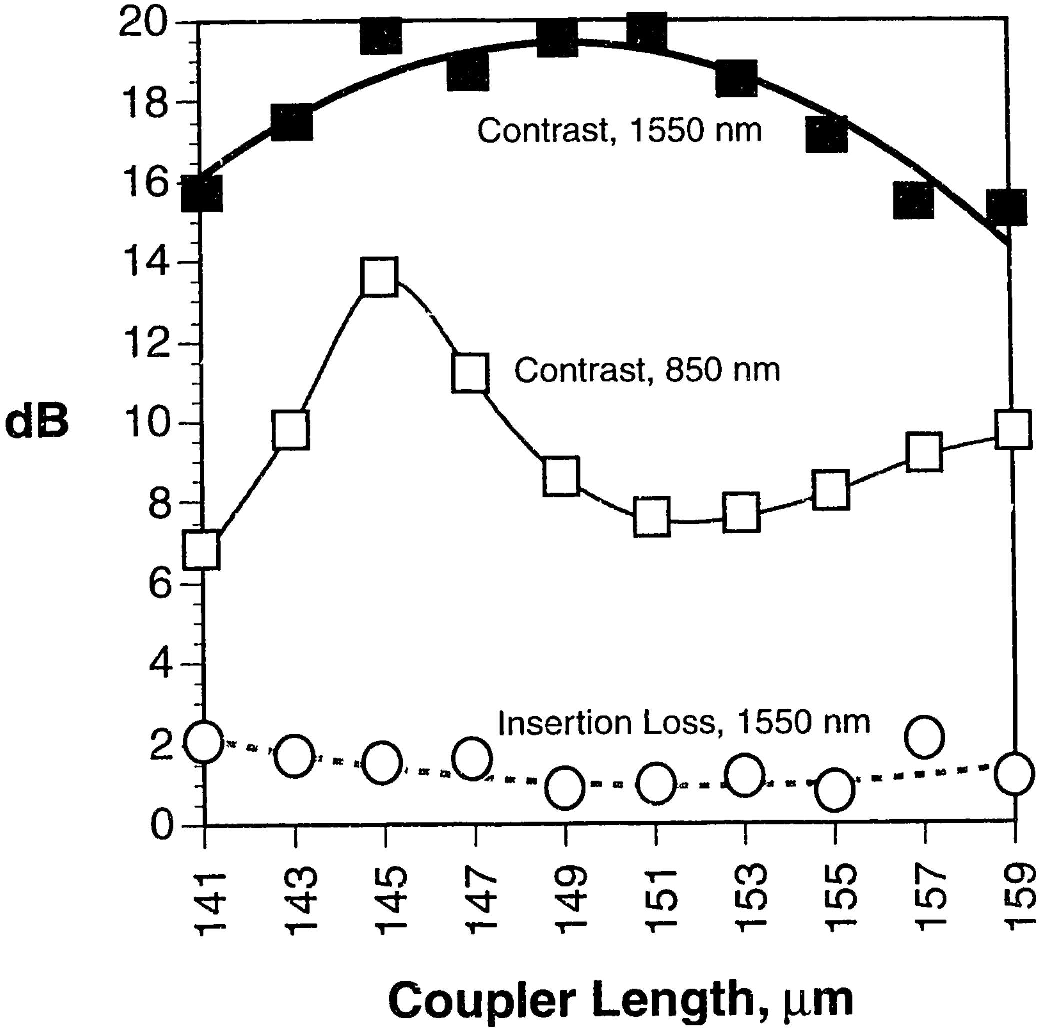


Figure 4.15 Measured performance of the fabricated devices, in the demultiplexer configuration, as a function of coupler length. Contrast and insertion loss are shown (filled and open circles, respectively - experimental values, bold and dashed lines - quadratic curve fits.) Optimum contrast at ~850 nm (open squares - measured values, tin line - cubic splines fit) is also shown.

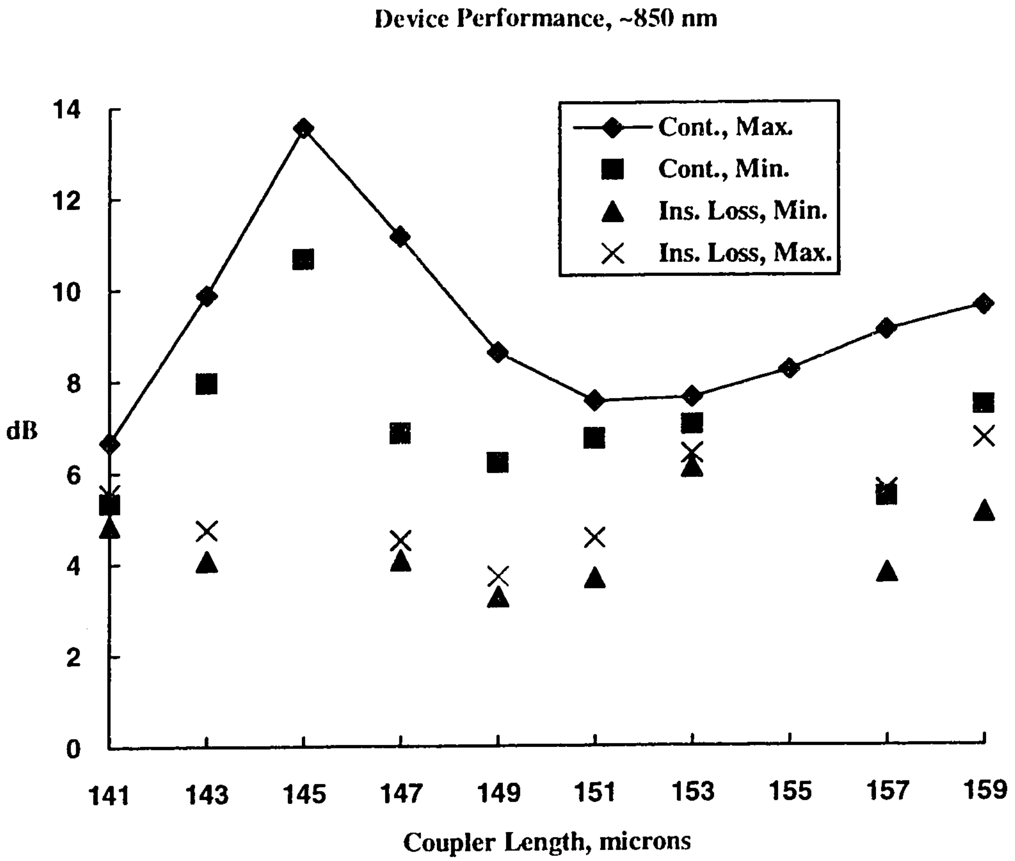


Figure 4.16 (Average) measured performance parameters at ~850 nm wavelength, as a function of coupler length in microns. The maximum contrast (diamonds, solid line) and minimum insertion loss (triangles) were obtained together. The minimum contrast (squares) and maximum insertion loss (crosses) were also measured together. The two sets of measurements were obtained at slightly different wavelengths (~5 nm separation.)

All the measured parameters at ~850 nm wavelength are shown in Figure 4.16; i.e. maximum and minimum contrast and insertion loss, each obtained by a small wavelength tuning around 850 nm. Some data points are omitted as the measured values, among different samples, varied considerably, making their "average" value suspect. At the peak contrast, the spread between maximum and minimum values is about 3 dB, slightly less than predicted in Figure 4.12. The insertion losses are higher than predicted, reaching a minimum of just over 3 dB.

In Figure 4.17, we show the contrast observed with a 145 μm-long coupler as the wavelength is tuned from 840 nm to 854 nm. The wavelength tuning and data collection was computer-controlled, and the optical power from each of the two outputs of the coupler was measured on successive scans. The contrast appears roughly periodic with respect to wavelength, reaching a peak value at about 846 nm, and a minimum at about 851 nm. This behaviour can be explained as due to the change of the phase separation between the fundamental and second-order modes of the input waveguide, at the coupler input, from π to zero. Let us define a beat length between these two modes as follows:

$$L_\pi^2 \equiv \frac{\pi}{\beta_0 - \beta_2} \tag{4.34}$$

This parameter is calculated to have values of 9.24 μm and 9.29 μm at 851 nm and 846 nm, respectively. Defining the input waveguide length as L, at the coupler input, we must have:

$$\frac{L}{L_\pi^2\big|_{851\text{ nm}}} - \frac{L}{L_\pi^2\big|_{846\text{ nm}}} = 1 \tag{4.35}$$

if modes 0 and 2 of the input waveguide are in phase at 851 nm and out of phase at 846 nm, at the coupler input junction. Rewriting (4.35), we have:

$$\frac{L_\pi^2\big|_{846\text{ nm}} - L_\pi^2\big|_{851\text{ nm}}}{\left(L_\pi^2\big|_{846\text{ nm}}\right)\left(L_\pi^2\big|_{851\text{ nm}}\right)} = \frac{1}{L} \tag{4.36}$$

From (4.36) and the calculated values of the beat lengths (4.34), L has a calculated value of 1.72 mm. The actual length of the input waveguide of the measured device is ~2 mm: a reasonable agreement.

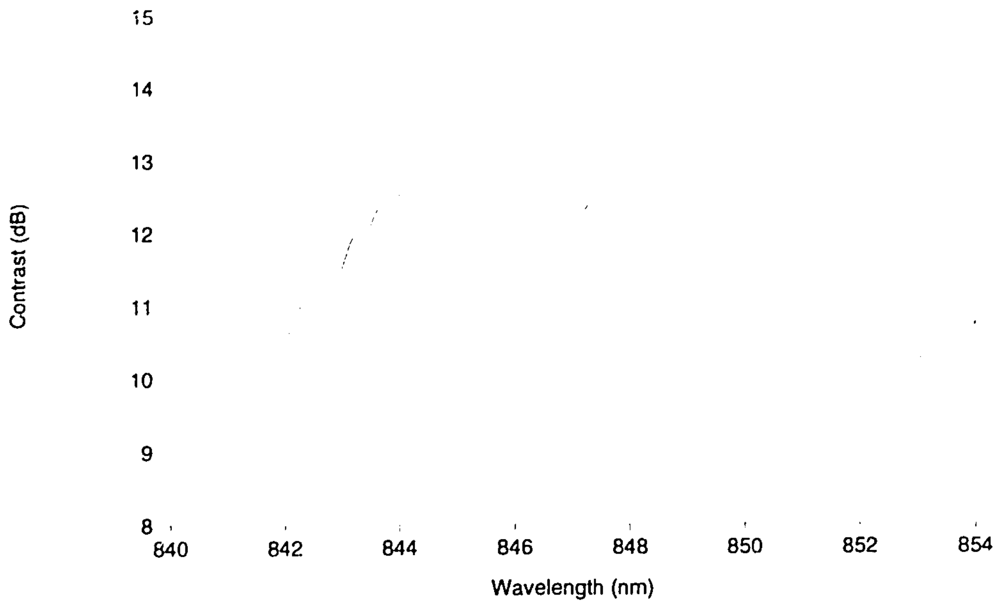


Figure 4.17 Measured contrast of a 145 μm-long coupler, in the demultiplexer configuration, as a function of the input wavelength.

The polarization dependence of the device was not specifically modeled; however, in accordance with the discussion of section 3.2, only a small difference in performance, with respect to polarization, is expected. Contrast and insertion loss measurements at ~850 nm were conducted using TM-polarized light, obtained by rotating the half-wave plate by 45 degrees. In general, the observed difference in contrast between TE- and TM-polarizations was less than 1 dB. The measured polarization dependence of the insertion loss was generally less than 0.5 dB. Also, although the devices were not modeled or tested in the multiplexing configuration, reciprocal operation is expected and insertion losses comparable to those obtained in demultiplexing operation are anticipated.

4.6 Conclusions In this chapter, we have described a novel method of realizing compact multi-mode waveguide couplers to serve as two-wavelength (de)multiplexers. The ratio of beat lengths at two semi-arbitrary wavelengths can be reduced to an optimum value of 1.5 by bending the coupler to an appropriate radius. High contrasts and low insertion losses can be achieved at both wavelengths using this approach. We have designed, fabricated and experimentally tested devices based on this concept using deeply etched silicon oxynitride rib waveguides. The observed performance of the devices was in reasonable agreement with the theoretical predictions. Contrasts of almost 20 dB and 14 dB were obtained at 1550 nm and 850 nm, respectively, with a bent coupler only 145 μm long. This is the smallest device yet reported which can achieve such high performance. The design concept is applicable to many combinations of wavelengths and is ideally suited for implementation in deeply etched rib waveguides.

Chapter 5

Integration of Laser and Electroabsorption Functions on One Semiconductor Substrate: LAMDA

Truly monolithic integration of photonic integrated circuits requires the implementation of optoelectronic devices with various functions on a single substrate. For example, lasers and optical amplifiers, transparent waveguides, detectors and electroabsorption/electrorefraction modulators, and possibly electronic components such as field effect transistors, may all be required to implement a complex photonic circuit. Implementing these various devices on a single semiconductor substrate is difficult, since each optoelectronic component type requires a different characteristic band-gap energy for operation at a given wavelength. The work described in this chapter addresses one way of overcoming this problem. An AlGaAs device is designed, fabricated and tested to demonstrate monolithic integration of laser and electroabsorption modulation functions. This device is called LAMDA, for LAser/Modulator/Detector/Amplifier, since all of these functions can potentially be realized.

Achieving monolithic integration of optoelectronic devices of different functions is an important goal, and considerable research continues to be devoted to it. Some impressive results have been reported in the literature to date. In [71] a monolithically integrated cascaded DFB laser and electroabsorption modulator in InGaAsP was reported which used quantum wells as the active regions of both sections, by growing the materials over patterned substrates. The effective band-gap energy in the modulator section was enhanced by growing the quantum wells over a patterned region of large width; the resulting large growth area seems to reduce the rate of semiconductor growth, leading to thinner well layers and states of increased quantized energy. A similar structure was reported in [72]. In [3] and [73], integration was achieved by including two sets of quantum wells in a single vertical structure. These quantum wells had different widths, to offset their characteristic energies. In both cases, the thicker quantum wells were etched away and new semiconductor material regrown to fabricate modulator sections. In [3] an integrated amplifier/Mach-Zehnder modulator was demonstrated, while [73] described an integrated DFB laser/electroabsorption modulator. In [74] it was shown that an integrated DFB laser and electroabsorption modulator could be fabricated using a single quantum well stack by choosing the period of the distributed feedback grating to offset the resonant lasing wavelength from the effective band-gap of the quantum wells. Still another method, based on disordering of the quantum wells, is reported in [75]. In general, patterned epitaxy and selective-area regrowth are the most widely applied approaches for multi-functional integration, and indeed have been applied in the fabrication of highly complex photonic integrated circuits [4], [5].

Semiconductor quantum wells, which are periodic heterostructures consisting of very thin layers (~10 nm) of alternating band-gap composition, offer the possibility of using quantum size effects to tailor the energy properties of the combined layer stack. This phenomenon has seen useful application in lasers, modulators and detectors. Quantum well devices are particularly suited to monolithic multi-functional integration since not only the material parameters, but also certain physical dimensions, i.e. the well widths, can be used to influence the energy properties - there is an extra degree of freedom. The examples listed above use this property to advantage, and for the same reason, we also make use of quantum wells in LAMDA. Quantum well fabrication requires the use of advanced epitaxial growth processes, such as Molecular Beam Epitaxy (MBE). MBE is essentially an evaporation technique, where an ultra-high vacuum ($<10^{-9}$ torr) chamber and precision source control, coupled with sophisticated *in-situ* monitoring techniques, are used to achieve high-quality crystal growth with extremely precise control of layer compositions and thicknesses. Heterostructures with layer thicknesses down to a single atom can be fabricated using MBE. Similar control can be achieved using Metal-Organic Chemical Vapour Deposition (MOCVD) processes.

Our goal is to demonstrate that multi-functional integration can be achieved without either patterned epitaxy or selective-area regrowth, since both processes lead to expensive, low-yield fabrication. To achieve this, we include two sets of quantum wells in a single vertical layer growth. The well widths in each stack are made different to achieve an effective band-gap separation. Unlike [3] and [73], the LAMDA structure is configured as a p-i-n-i-p-doped three-terminal device with electrical connections to the upper cladding layer, the guiding layer, and the substrate; see Figure 5.1. This feature eliminates the need for etching and regrowth and allows us to "stack" the laser and modulator sections vertically rather than cascade them. Intensity modulation can therefore be achieved through variable intracavity absorption in a single laser cavity. In cases where separate modulator or detector sections are required, the gain layer need not be removed since the three-contact arrangement allows it to be biased to transparency. The guided mode interacts with both active layers in all sections and hence local function is determined exclusively by biasing. This greatly simplifies the device processing and results in optimum optical coupling between segments since the layer structure is identical throughout.

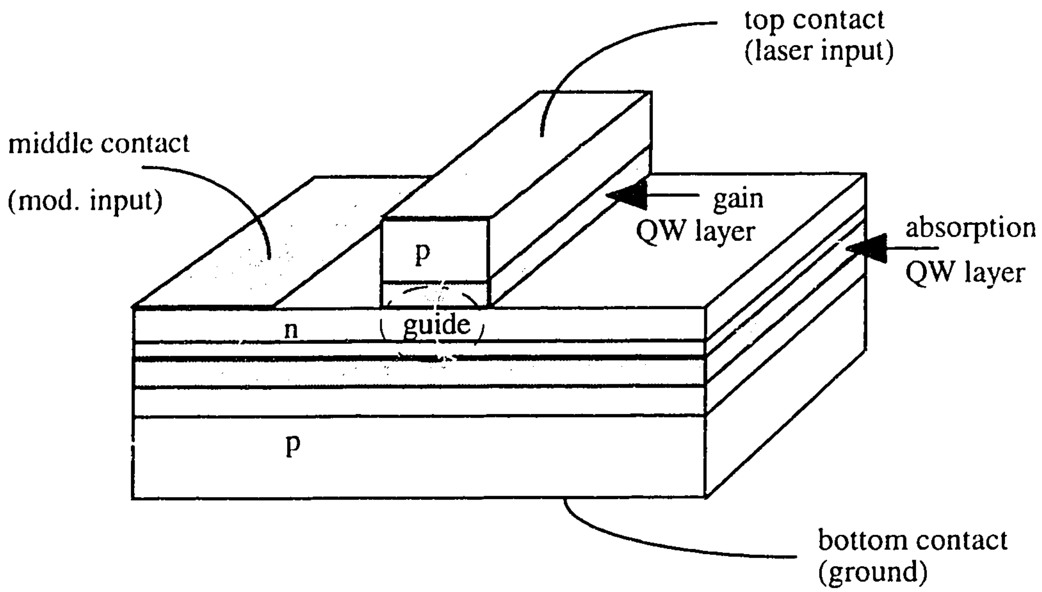


Figure 5.1 Schematic of LAMDA structure, processed for ridge waveguide configuration. Note the p-n-p doping profile; the quantum well stacks are intrinsic. The guided optical field overlaps with both active regions, which can be separately biased.

5.1 Optical Gain in Semiconductor Quantum Wells To design the LAMDA structure, it was necessary to develop models for optical gain and absorption processes in quantum well semiconductor structures. In this section, we consider the physics of optical gain in quantum wells.

Quantum well (QW) lasers offer several advantages over conventional, double heterostructure (DH) lasers [76]. Lasing wavelengths can be controlled by the energy level quantization in the wells, which is a function of well width and depth. The high current confinement offered by the thin active region leads to a substantial reduction of the threshold current, and a high differential gain (change of gain with respect to injected carrier density), resulting from the quasi-two dimensional character of the well, leads to an increased relaxation oscillation frequency. Moreover, a large difference in gain between the TE and TM modes in QW structures results in enhanced polarization stability.

As with any laser diode, gain is obtained in QW structures by current injection under forward bias. Conduction band electrons, injected from the n-material, combine with valence band holes, injected from the p-material, and since the resulting bound electron is less energetic than the free electron-hole pair, energy is released, sometimes as a photon. This photon release can occur as spontaneous emission, with a certain probability per unit time, or as stimulated emission, with a probability proportional to the population of photons in a particular optical mode. The reverse process can also occur: a photon may be absorbed by a bound electron, creating a free electron-hole pair. Stimulated emission and absorption, for a given optical mode, have equal inherent probabilities of occurrence. Therefore, if the population of free carriers exceeds the remaining number of conduction states available at a particular energy, i.e. if a population inversion is achieved, stimulated emission exceeds absorption and optical gain results.

In forward bias, quantum size effects are manifested in two important respects. The first is the partial quantization of energy states. The energy of a particle confined in a quantum well is given by, for small transverse momenta:

$$E(N,k_x,k_y) = E_N + \frac{\hbar^2}{2m^*}(k_x^2 + k_y^2) \tag{5.1}$$

where $\hbar$ is Planck's constant divided by 2π, m^* is the effective mass of the particle, k_x and k_y are the particle wave numbers along the unconfined axes, and E_N are the quantized energy states resulting from the small well thickness, obtained as solutions of the time-independent Schrodinger equation for a finite potential well. Therefore, in contrast to conventional structures where the values of E_N form a continuum, the minimum particle energies in the active region are displaced from the band edges by E_0. Hence, the minimum energy of emitted photons is augmented by $E_0^{cb} + E_0^{vb}$, the fundamental energy eigenvalues in the conduction and valence bands, respectively. According to the Schrodinger equation, the sub-band energies E_N are determined by the well's physical thickness and energy depth. These parameters can be controlled during epitaxial growth, and hence the peak photon emission energies are easily tailored, since radiative transitions decrease at energies exceeding the quantized sub-band transition energies due to fast intraband relaxations along the unconfined axes. The well depth is a function of the relative stoichiometric composition of the well and barrier layers. A GaAs/AlGaAs quantum well has an energy depth determined by the concentration of Al in the barrier layers. Approximately 57% of the band-gap difference between the well and barriers contributes to the conduction band well, and 43% comprises the valence band well. See Figure 5.2.

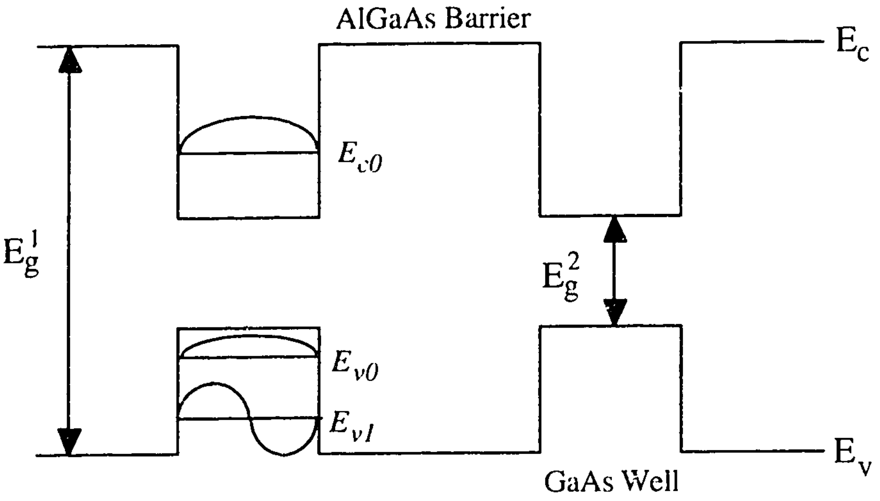


Figure 5.2 Energy schematic of an AlGaAs multiple quantum well system. GaAs has a smaller band-gap energy E_g than AlGaAs, establishing potential "wells" where the total band-gap difference is divided between the conduction and valence bands. If the well layers are sufficiently thin, quantized carrier states are created with respect to the axis of epitaxial growth.

The second important quantum size effect in forward bias is the alteration of the density of states functions, which are important in calculating optical gain. If we label the axis of epitaxial growth z then the carriers in the wells are unconfined along the x and y axes. The carrier wave numbers along these axes are then:

$$k_x = \frac{m\pi}{L_x}, \quad k_y = \frac{n\pi}{L_y} \tag{5.2}$$

where m and n are integers, and L_X and L_Y are the transverse dimensions of the well. The area per eigenstate, in wave number space, is therefore:

$$A = \frac{\pi^2}{L_x L_y} \tag{5.3}$$

The number of eigenstates that have transverse momenta less than $\hbar k$, where:

$$k = \sqrt{k_x^2 + k_y^2} \tag{5.4}$$

is therefore given by:

$$N(k) = \frac{\pi k^2}{4}\frac{2}{A} = \frac{k^2 L_x L_y}{2\pi} \tag{5.5}$$

since the momentum value $\hbar k$ defines an area of $\pi k^2/4$ in k-space, and two spin orientations of each carrier are possible. The number of states per unit area with energies between E and $E+dE$ is then:

$$\frac{1}{L_x L_y}\frac{dN(E)}{dk}\frac{dk}{dE}dE = \frac{k}{\pi}\frac{dk}{dE}dk \tag{5.6}$$

The carrier energy and wave number, along the unconfined axes, are related by the well-known equation:

$$k = \sqrt{\frac{2m^*}{\hbar^2}}\left(E - E_{\min}\right)^{1/2} \tag{5.7}$$

where $E_{\min}$ is the smallest allowed energy. In quantum wells, from equation (5.1), a function of the form of (5.7) is repeated for every quantized energy level resulting from the vertical confinement. Therefore, the density of states function is:

$$\rho(E) = \frac{1}{L_z}\frac{k}{\pi}\frac{dk}{dE} = \sum_{n=1}^{\max}\frac{m^*}{\pi\hbar^2 L_z}H\left(E - E_g - E_{qn}\right) \tag{5.8}$$

where $H(x)$ is the unitary Heaviside function (i.e. $H(x)=1$, $x>0$, $H(x)=0$, $x<0$) and $q=c$ for electrons in the conduction band, and $q=v$ for holes in the valence band. E_g is the band-gap energy, and the quantity L_z is the thickness of the quantum well and has been included in (5.8) to define $\rho(E)$ per unit volume, rather than area. The sum index limit "max" corresponds to the highest-order quantized field along the growth axis. The density of states functions are thus step-like in QW structures rather than parabolic as in bulk semiconductors; see Figure 5.3. This has a number of important effects. First, it dictates that the wavelength of peak gain will shift discretely with increased injection, rather than continuously as in unconfined structures. Second, it allows all injected carriers to contribute to the gain at its wavelength peak, resulting in a very large value of differential gain, $\partial g/\partial n$, where g is the gain coefficient and n is the injected carrier density. This both reduces the threshold current and increases the frequency of relaxation oscillations. Finally, the altered character of the density of states functions leads to a saturation of the gain achievable via any given sub-band transition, a phenomenon which does not occur in unconfined structures. These phenomena are illustrated in Figure 5.4.

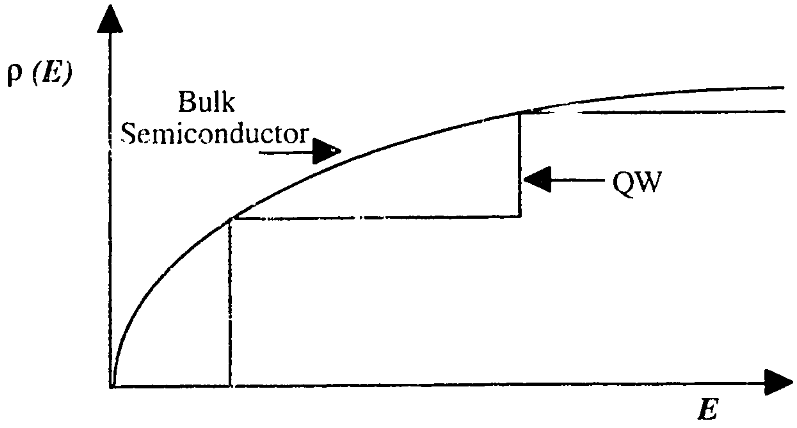


Figure 5.3 Density of states function in a quantum well, compared with that in a bulk semiconductor layer.

QW lasers exhibit enhanced polarization stability due to a large difference in gain between TE and TM modes; this is another result of the quasi-two-dimensionality of the active region. Since the (quasi-)TE modes have their principal electric field component in the plane of the layers, greater coupling exists between the radiation electric field and the carrier dipole moments than for the TM modes, because the (electron-heavy hole) dipoles tend to lie more in the plane of the layers due to the vertical quantum confinement. The result is an enhanced probability of stimulated emission, which in turn is directly manifested as a greater gain for the TE modes.

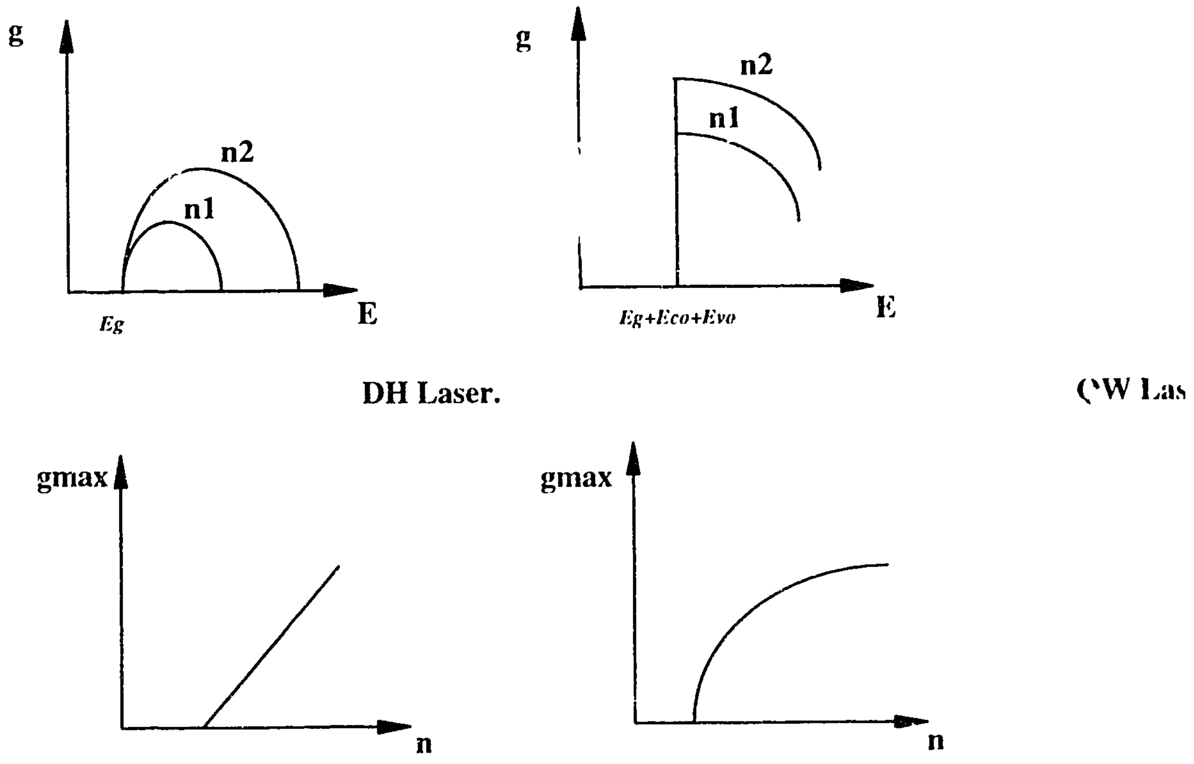


Figure 5.4 Comparison of the gain characteristics of DH and QW laser structures. The spectral characteristic (gain g vs. photon energy E) of the QW laser is step-like, providing excellent control of the lasing wavelength. The peak gain (*gmax*) vs. injected carrier density (n, $n2>n1$) of the QW exhibits an initially higher slope (larger differential gain), and saturates at high injection levels. Higher sub-band transitions, which appear at large n for QW lasers, are not shown. E_g is the band-gap energy.

5.2 Calculation of Gain in Quantum Wells The expressions which determine the rates of optical absorption, and spontaneous and stimulated emission are derived from Fermi's Golden Rule:

$$W_l = \frac{2\pi}{\hbar}\left|\langle 1, p_l + 1|H_l|2, p_l\rangle\right|^2 \delta(E_2 - E_1 - \hbar\omega_l) \tag{5.9}$$

taken in the limit of harmonic emission (δ is the Dirac delta). In (5.9), W_l is the (probability) rate of emission into the lth optical mode, p_l is the number of photons in that mode, ω_l is its angular frequency, E_1 and E_2 are the final and initial energies of the emitter, and H_l is the perturbation Hamiltonian causing the transition. In semiconductors, the considered transition is the collapse of a free electron-hole pair. The Hamiltonian is normally considered in the dipole approximation:

$$H_l = -e\bar{E}_l \bullet \bar{r} \tag{5.10}$$

where $\bar{E}_l$ is the time-independent electric field of the lth optical mode, and $\bar{r}$ is the position operator. By writing expressions for the quantized optical fields, it can be shown (see Chapter 6) that (5.9) reduces to:

$$W_l = \left(\frac{2\pi}{\hbar}\right)\left(\frac{e^2\hbar\omega}{2n_{cav}^2\varepsilon_o V}\right)\left|\langle 1|\bar{E}_l(\bar{r})\bullet\bar{r}|2\rangle\right|^2 (p_l + 1)\delta(E_2 - E_1 - \hbar\omega) \tag{5.11.a}$$

$$= \left(\frac{\pi e^2\hbar}{n_{cav}^2\varepsilon_o V m_o^2 E}\right)|M|^2 (p_l + 1)\delta(E_2 - E_1 - E) \tag{5.11.b}$$

where $|M|^2$ is called the momentum matrix element and is given by:

$$|M|^2 = m_o^2\omega^2\left|\langle 1|\bar{E}_l(\bar{r})\bullet\bar{r}|2\rangle\right|^2 \tag{5.12}$$

where m_o is the electronic rest mass and:

$$\left|\langle 1|\bar{E}_l(\bar{r})\bullet\bar{r}|2\rangle\right|^2 \tag{5.13}$$

is the field-dipole matrix element, with dimensions of length^2. The derivation of equation (5.11) assumes that the optical mode l is one of a complete set of plane waves of a cavity of uniform refractive index n_{cav} and volume V, and the mode field has been energy-normalized on this basis. The field-dipole matrix element (5.13) therefore does not depend on the field magnitude since this is the same everywhere for all modes. The momentum matrix element determines the strength of the transition for a given magnitude and orientation of the dipole and polarization (i.e. orientation of the electric field vector) of the optical mode. The determination of this parameter for quantum wells has lead to some confusion in the literature. An excellent discussion is presented in [77] which gives, for $\Delta k = 0$ inter-band (i.e. conduction band-valence band) transitions in GaAs, and TE modes:

$$|M_l|^2 = \frac{1}{2}|M_{avg}|^2 \qquad \textbf{(5.14.a)}$$

$$|M_h|^2 = \frac{3}{2}|M_{avg}|^2 \qquad \textbf{(5.14.b)}$$

where:

$$|M_{avg}|^2 = 3.38 m_o E_g \qquad \textbf{(5.15.c)}$$

where E_g is the band-gap energy of GaAs. The subscripts l and h refer to light and heavy holes, respectively, and denote the effective hole masses in the two direct-gap valence bands of GaAs. The momentum matrix element value for TM modes and heavy holes is zero and hence TM modes are not considered further. In fact, (5.14) is strictly valid only for transitions between the quantized sub-band energy levels but is a reasonable approximation for small k_x, k_y. Equations (5.14) and (5.15) account for spin degeneracy and spin conservation in inter-band transitions.

The gain in a single optical mode can be determined from equation (5.11.b) by considering only the term proportional to p_l which corresponds to stimulated emission or absorption. Again assuming momentum-conserving inter-band transitions, the rate of photon absorption is:

$$W_l \rho_{red}^{j,N}(E) V f_{vo}^{j,N}(E)\left(1 - f_{co}^{j,N}(E)\right) \qquad \textbf{(5.16)}$$

where W_l now refers only to the term of (5.11.b) which is proportional to p_l, and we have used a reduced density of states function given by:

$$\rho_{red}^{j,N}(E)=\left(\frac{m_e^* m_{jh}^*}{m_e^*+m_{jh}^*}\right)\left(\frac{1}{\pi\hbar^2 L_z}\right)H(E-E_g-E_{cN}-E_{jN}) \tag{5.17}$$

since $\Delta k=0$ and the electron and hole densities of states are considered together. In (5.17), the subscript e denotes electrons, $j=l,h$ denotes light and heavy holes, and N denotes sub-band energy quantum numbers. Note that we assume the same value of N for electrons and holes; this is because the electron and hole wavefunctions (along the vertical axis) are approximately orthogonal in deep potential wells, leading to a negligible field-dipole matrix element value for most transitions where $\Delta N\neq 0$. The functions f_c and f_v are Fermi-Dirac distribution functions which give the probability of *electron* occupancy of an energy state in the conduction and valence bands, respectively. These functions are given by:

$$f_{co}^{j,N}(E)=\frac{1}{1+\exp\left(\dfrac{E_{cN}+\left(\dfrac{m_{jh}^*}{m_e^*+m_{jh}^*}\right)(E-E_g-E_{cN}-E_{jN})-E_{fc}}{kT}\right)}$$

(5.18.a,b)

$$f_{vo}^{j,N}(E)=\frac{1}{1+\exp\left(\dfrac{-E_{jN}+\left(\dfrac{m_e^*}{m_e^*+m_{jh}^*}\right)(E-E_g-E_{cN}-E_{jN})+E_{fv}}{kT}\right)}$$

where k is Boltzmann's constant, T is the temperature, and E_{fc} and E_{fv} are the conduction and valence band quasi-Fermi levels, respectively. These are measured from the respective well bottoms, with increasing magnitudes at higher levels in the wells; the sub-band levels E_{cN} and E_{jN} are measured the same way.

The number of stimulated emission transitions per unit time is:

$$W_l \rho_{red}^{j,N}(E) V f_{co}^{j,N}(E)\left(1 - f_{vo}^{j,N}(E)\right) \qquad \textbf{(5.19)}$$

The gain coefficient is obtained by subtracting (5.16) from (5.19), integrating with respect to energy to eliminate the Dirac delta functions, and multiplying by n_{cav}/c_o, where c_o is the speed of light in free space, to convert the emission rate per unit time into a rate per unit length. This yields:

$$g(E) = \sum_N \sum_j \left(\frac{\pi e^2 \hbar}{\varepsilon_o n_{cav} c_o m_o^2 E} \right) \left| M_j \right|^2 \rho_{red}^{j,N}(E) \left[f_{co}^{j,N}(E) - f_{vo}^{j,N}(E) \right] \qquad \textbf{(5.20)}$$

if gain-broadening mechanisms are not considered [78],[79]. Since the number of quantized levels in the conduction and valence bands are not generally equal, the summation limit of N is the smaller of the two for each j. Equation (5.20) includes a sum over all the inter-band transitions, since more than one transition may produce a photon of energy E. Since the optical mode l in (5.16) and (5.19) is a plane wave extending over all space, (5.20) must be taken as a bulk gain expression. In practical cases, an approximate modal gain coefficient can be obtained by multiplying (5.20) by the optical power confinement factor of the mode being considered. Equation (5.20) assumes infinite lifetimes for the carriers in the well eigenstates. Of course, intraband relaxation does occur and hence some energy broadening is expected. Usually, this is treated by a convolution of (5.20) with a Lorentzian function, with an energy-dependent lifetime on the order of 10^{-13} s. We ignore gain broadening in our treatment, since we seek only the *peak* gain value and wavelength, which determines the threshold of a laser.

The rate of spontaneous emission, into a specific optical mode l, is given by integrating (5.19) with respect to energy where now W_l is taken as the term of (5.11.b) which is *not* proportional to p_l. This yields:

$$r_i^{sp}(E) = \left(\frac{2\pi}{\hbar}\right)\left(\frac{e^2\hbar\omega}{2n_{cav}^2\varepsilon_o m_o^2\omega^2}\right)\left|M_j\right|^2 \rho_{red}^{j,N}(E) f_{co}^{j,N}(E)\left(1 - f_{vo}^{j,N}(E)\right) \tag{5.21}$$

The *total* rate of spontaneous emission, into all optical modes, may be obtained by integrating (5.21) over all energies, taking account of the mode density with respect to energy:

$$R_{sp} = \sum_N \sum_j \int_E \left(\frac{2\pi}{\hbar}\right)\left(\frac{e^2\hbar\omega}{2n_{cav}^2\varepsilon_o m_o^2\omega^2}\right)\left|M_j\right|^2 D(E)\rho_{red}^{j,N}(E) f_{co}^{j,N}(E)\left(1 - f_{vo}^{j,N}(E)\right)dE \tag{5.22.a}$$

$$= \sum_N \sum_j \int_E \frac{n_{cav}e^2 E}{\pi\hbar^2 m_o^2 c_o^3 \varepsilon_o}\left|M_j\right|^2 \rho_{red}^{j,N}(E) f_{co}^{j,N}(E)\left(1 - f_{vo}^{j,N}(E)\right)dE \tag{5.22.b}$$

where:

$$D(E) = \frac{n_{cav}^3 E^2}{\pi^2\hbar^3 c_o^3} \tag{5.23}$$

is the optical mode density, per unit volume per unit energy, in a large cavity of uniform refractive index n_{cav}. The total spontaneous emission rate (5.22) has units of (volume×time)$^{-1}$. In situations where the complete set of optical modes is not simply a set of uniform plane waves, e.g. in the vicinity of an optical waveguide, the presentation given in this section is, to some, degree, incorrect. This issue is examined in some detail in Chapter 6 with respect to spontaneous emission. The evaluation of (5.22) is important in determining the threshold current of a laser, since spontaneous emission is the dominant radiative recombination mechanism below threshold.

5.3 Determination of Quasi-Fermi Levels The quasi-Fermi levels, required for the evaluation of (5.20) and (5.22), are determined by the densities of free carriers resulting from current injection. The carriers occupy energy states according to the

Fermi-Dirac distribution function. In general, carrier density is calculated by integrating the product of the density of states and the Fermi function, over all energies. Analytical expressions can be obtained for the densities of carriers which occupy the quantum well energy states given by equation (5.1). For the conduction band, we have:

$$n = \sum_N \int_{E_{cN}}^{\infty} \rho_c(E) f_c(E) dE \tag{5.24.a}$$

$$= \sum_N \int_{E_{cN}}^{\infty} \frac{m_e^*}{\pi \hbar^2 L_z} \left[\frac{1}{1+\exp\left(\frac{E - E_{fc}}{kT} \right)} \right] dE \tag{5.24.b}$$

$$= \frac{m_e^*}{\pi \hbar^2 L_z} \sum_N \ln\left[1+\exp\left(\frac{E_{fc} - E_{cN}}{kT} \right) \right] \tag{5.24.c}$$

where n is the volume electron density. Similarly, for the valence band we obtain:

$$p = \frac{1}{\pi \hbar^2 L_z} \sum_j \sum_N m_{jh}^* \ln\left[1+\exp\left(\frac{E_{fv} - E_{jN}}{kT} \right) \right] \tag{5.25}$$

where p is the total hole density. Since the quantum wells are undoped, we make the standard assumption of space charge neutrality and hence require that $p=n$. If all of the carriers were confined to the well energy states of equation (5.1), we could calculate the quasi-Fermi levels directly by numerical inversions of (5.24) and (5.25).

It has been indicated by some authors [80] that, under certain circumstances, e.g. with narrow quantum wells, small potential barriers and/or at high injection current levels, various carrier overflow mechanisms can influence the quasi-Fermi levels. Some fraction of the carriers may spill into the continuum of energy states above the barrier potential, either physically in the well layers or in the surrounding barriers or other semiconductor layers. Also, some electrons may occupy states in the

L conduction band. Neglecting these mechanisms can lead to overly optimistic estimates of laser threshold currents and therefore we include them here.

We refer to the electrons with energies exceeding the conduction band barrier potential, but which nevertheless are physically located in the well layers, as unconfined electrons, with volume density n_{un}. These carriers are governed by a standard parabolic density of states function, and hence they influence the quasi-Fermi level according to:

$$n_{un} = 4\pi\left(\frac{2m_e^*}{h^2}\right)^{\frac{3}{2}} \int_{E_{bar}}^{\infty} (E - E_{bar})^{\frac{1}{2}} \left[\frac{1}{1+\exp\left(\frac{E-E_{fc}}{kT}\right)}\right] dE \qquad (5.26)$$

where E_{bar} is the barrier potential (i.e. the bottom of the Γ-valley in the barrier layers.) The density of electrons which penetrate into the barrier layers is similarly given by:

$$n_{ba} = 4\pi\left(\frac{2m_{eba}^*}{h^2}\right)^{\frac{3}{2}} \int_{E_{bar}}^{\infty} (E - E_{bar})^{\frac{1}{2}} \left[\frac{1}{1+\exp\left(\frac{E-E_{fc}}{kT}\right)}\right] dE \qquad (5.27)$$

where n_{ba} is the density of electrons in the barriers, and m_{eba}^* is the effective mass of electrons in the barrier layer Γ-valley. We assume that each quantum well is surrounded by a barrier layer and therefore exclude the possibility of carrier overflow into other semiconductor layers. Furthermore, we neglect *L*-valley population in both the wells and the barriers. This is justified as long as the barrier aluminum composition is less than about 37%, since the *L*-valley minimum in both the wells and the barriers then lies above the (Γ-valley) barrier potential. The total electron concentration is thus taken as:

$$n = n_{well} + n_{un} + n_{ba} \qquad (5.28)$$

where n_{well} is the density of confined electrons in the quantum wells, given by (5.24). Given a value for the total electron concentration, the conduction band quasi-Fermi level can then be found by self-consistent solution of equations (5.24) and (5.26)-(5.28).

Similar overflow mechanisms are not significant for the valence band, since the heavy hole effective mass is much greater than the electron Γ-valley effective mass, resulting in quantized sub-band levels lying much lower in the wells. Therefore, we assume the valence band quasi-Fermi level to be strictly related to the total hole population according to (5.25).

5.4 Calculation of Carrier Concentrations

In bulk GaAs lasers biased below threshold, the steady-state rate of carrier recombination, G, is well-described by [81]:

$$G = An + Bn^2 + Cn^3 \tag{5.29}$$

where A results from recombination at Shockley-Read-Hall traps, B determines spontaneous (radiative) recombination, and C is the Auger (multi-body) recombination coefficient. While different Auger processes are possible, corresponding to different carrier interactions, in most cases they can be adequately characterized collectively.

The description of spontaneous emission recombination using the bimolecular coefficient B is valid only where the radiative transitions are not k-selective, i.e. where $\Delta k \neq 0$ generally. It is therefore not an accurate description of radiative recombination in quantum wells, where k-selection rules are operative and the electron and hole populations cannot be considered independently. The term Bn^2 in (5.29) must therefore be replaced by the quantum mechanical expression (5.22). This considerably complicates the evaluation of the gain coefficient, since (5.22) makes the carrier concentration a function of the quasi-Fermi levels, and the system of equations (5.22), (5.24) and (5.26)-(5.28) must be evaluated self-consistently. Therefore, in this work, we retain the bimolecular expression for the spontaneous recombination rate, but seek a value of B which brings the values of Bn^2 close to the calculated values of (5.22) for the quantum well system, and over the current density range, of interest. Some validation of this method is provided by Figure 7 in [82], where the radiative recombination rate for GaAs quantum wells of various widths, calculated using a version of (5.22), is compared with that in bulk GaAs, as a function of temperature.

This figure shows that decreasing the well width may be effectively considered by increasing B. For 10 nm GaAs wells, the spontaneous emission rate increases by about a factor of three, versus bulk GaAs, at room temperature, corresponding to a value of $B \sim 1.2 \times 10^{-15}$ m^3s^{-1}. Anticipating that we will be modeling 10 nm GaAs wells with 32% Al barriers, we can check the accuracy of this value when applied in (5.29). At various current densities, we obtain n using (5.29), and then apply it in (5.24) and (5.25) to estimate the quasi-Fermi levels. We next apply these quasi-Fermi levels in (5.22), using an appropriate value of n_{cav}, to calculate the electron concentration using the correct quantum mechanical method. The values of n calculated by (5.22) and (5.29) are found to agree to within ~10% over a wide range of injection current densities when a temperature of 300 K is assumed.

The contribution of S-R-H trap recombination is widely neglected in QW laser modeling, although in fact trap recombination may be somewhat enhanced in QW structures, versus bulk GaAs, due to the large hetero-interface area inherent in a QW stack. The evidence suggests that these interfaces are densely populated with S-R-H traps. Studies of carrier lifetimes in MBE-grown GaAs quantum wells are available [83],[84] which give numerical values for the parameter A; in our model we scale this parameter by assuming that the trap recombination rate is a linear function of the hetero-interface surface area in the QW stack. Values for the Auger recombination coefficient, C, for GaAs QW structures are fairly well agreed upon in the literature; we take the value 4.2×10^{-42} $m^{-6}s^{-1}$. We find, in agreement with other studies, that the non-radiative contributions to recombination are almost negligible, so that in fact:

$$G \cong R_{sp} \cong Bn^2 \qquad \textbf{(5.30)}$$

where B is appropriately scaled as discussed above.

5.5 Modeling of Electroabsorption in Quantum Wells We next examine the physics governing the operation of semiconductor quantum wells as modulators. In recent years, the development of QW electroabsorption modulators has reached a relatively mature state. QW devices are particularly useful as modulators due to the existence of strong excitonic features in the absorption spectra [85]. Excitons are hydrogenic electron-hole entities not seen in bulk semiconductors except at very low temperatures, due to their low binding energy and ease of ionization. In quantum wells, however, the confining effect of the narrow potential wells increases the exciton binding energy by reducing the effective Bohr radius of the exciton. This enables the excitons to survive long enough, even at room temperature, to give rise to sharp absorption resonances at the exciton characteristic energies. These large magnitude resonances produce an extraordinarily sharp spectral transition from a low absorption to a high absorption state. Furthermore, when an electric field is applied perpendicular to the layers, the excitons survive, due again to the confining action of the potential barriers, while the characteristic energies are red-shifted according to the well-known Stark Effect. This Quantum Confined Stark Effect (QCSE) is a particularly powerful electroabsorption phenomenon and is the basis of operation for QW modulators. Since the QCSE does not involve particle injection or recombination, large modulation bandwidth is possible, limited essentially only by the RC-constant of the capacitance of the QW modulator in series with a drive impedance.

The energy states of electrons and holes in quantum wells are given by equation (5.1), where E_N are the discrete states resulting from confinement along the axis of layer growth (the *z*-axis.) Absorption of a photon creates an electron-hole pair, and the absorbed photon energy, *E*, is therefore determined from (5.1), and is given by, in the general case:

$$E = E_g + E_{cN} + E_{jN} + \frac{\hbar^2}{2}\left\{\frac{1}{m_e^*}\left(k_x^2 + k_y^2\right)_e - \frac{1}{m_{jh}^*}\left(k_x^2 + k_y^2\right)_{jh}\right\} \qquad \textbf{(5.31)}$$

Where the generated carriers have similar transverse momenta (i.e. k selection), Coulomb forces pull newly created electron-hole pairs into bound states: excitons. These excitons are generally in a (compressed) 1S orbital state with some characteristic binding energy, E_{jb}; higher energy orbitals are not significantly populated. With the selection rule $\Delta k = 0$ the photon energies giving rise to excitonic absorption resonances are therefore, approximately:

$$E^j = E_g + E_{cN} + E_{jN} - E_{jb} \qquad \textbf{(5.32)}$$

where the superscript on the photon energy denotes the hole type involved. Note that, as for the forward bias case, a $\Delta N = 0$ selection rule is implicit in (5.31) and

(5.32); this will be justified (and qualified) shortly. Equation (5.32) determines the exciton resonance energies even when an electric field is applied along the z-axis, and E_{cN} and E_{jN} are reduced by the Stark Effect. The exciton binding energies are not significantly perturbed by an applied field. Although the QCSE is directly an absorption spectrum modulation effect, it can also be used to achieve electrorefraction, such as in the Mach-Zehnder switch of Chapter 3, since the absorption and refractive index spectra of a material are intimately related via the well-known Kramers-Kronig relations.

We note here that excitons play no part in optical gain processes, at least in III-V semiconductors, since the large population of free carriers in quantum wells under forward bias screens the Coulombic interaction that binds excitons, greatly shortening their lifetimes.

The magnitudes of the excitonic resonances can be determined from first principles (Fermi's Golden Rule); however, this is a detailed quantum mechanical calculation. A semi-empirical approach, widely used in the literature, models the resonance strength with an applied electric field using experimentally determined values at zero field. The decay in resonance strengths with an electric field is well-approximated using the square of the overlap integral of electron and hole wave functions. This method is easily justified by looking at the Golden Rule [86]. The field-dipole matrix element (5.13) which determines the strength of absorption transitions is, for inter-band absorption transitions:

$$\left|\int_V \Psi_{jN}^{*}(\bar{r})\vec{E}_l(\bar{r})\Psi_{cN}(\bar{r})dV\right|^2 \cong \left|\vec{E}_l(\bar{r})\int_V \Psi_{jN}^{*}(\bar{r})\Psi_{cN}(\bar{r})dV\right|^2 \qquad (5.33)$$

where Ψ are the particle wave functions of the holes and electrons. The optical mode electric field is removed from the integral since it is approximately constant over the dimensions of the quantum well. Physically, (5.33) means that the transition probability depends on the extent to which the electron and hole are created in the same location. As an electric field is applied to a quantum well, the wave functions of the discrete particle states are pulled in opposite directions along the z-axis. Therefore, if we assume that the wave functions are separable into real functions of z and the transverse coordinates x and y, and that there is no change in the particle wave functions in the plane of the layers when a field is applied, we can write:

$$\alpha_{ex} \cong \alpha_{ex}^{0} \left| \int_{-\infty}^{\infty} \psi_{hN}(z)\psi_{eN}(z)dz \right|^{2} \tag{5.34}$$

provided that the wave functions are normalized; α_{ex} is the magnitude of the exciton absorption resonance, and α_{ex}^{0} is its magnitude at zero field. Note that the selection rule $\Delta N = 0$ is explained by this relation; under zero field, the electron wave functions are approximately orthogonal to the hole wave functions of different quantum number. This orthogonality is progressively destroyed as a field is applied and the wave functions become skewed in opposite directions; the selection rule is therefore relaxed with the application of a field. Resonances corresponding to $\Delta N \neq 0$ start to be observed; in fact, it can be shown that the collective strength of these new peaks exactly compensates for the decay in the strength of the peaks allowed at zero field [87].

The shape and width of the exciton resonances are determined primarily by the lifetime-limiting processes of the excitons. The three most significant homogeneous processes are phonon interaction, tunneling through the barriers, and electron-hole recombination. It has been shown experimentally [88] that phonon ionization at room temperature dominates over recombination, and for normal barrier thicknesses and potentials, tunneling is not significant, even at high fields. We therefore expect Lorentzian lineshapes for the exciton peaks, with a width determined by the phonon ionization lifetime. This width is approximately field-independent, since phonon energies at room temperature are much larger than the exciton binding energy, even at zero field, and any collision with a phonon will likely ionize the exciton. In any case, the binding energy is approximately independent of field. The phonon linewidth is proportional to the density of phonons, which is governed by Bose-Einstein statistics.

Two inhomogeneous processes also broaden the exciton peaks: variations in well width, and field variations from well to well due to residual doping in the QW stack. The latter is quite generally encountered, since, to effectively apply an electric field, the QW stack is usually made the intrinsic part of a p-i-n diode, and therefore residual doping gases in the growth chamber and/or diffusion lead to some dopant contamination of the wells. These mechanisms, which are field dependent, yield Gaussian lineshapes which properly must be convolved with the dominant Lorentzian function. A simpler approach is that taken by Lengyel *et al.* [89], where a single

Lorentzian line function is retained, but with a linewidth augmented by the inhomogeneous mechanisms. Lengyel fitted a well width and field dependent function to experimental data to obtain a linewidth function for GaAs-AlGaAs QW's:

$$\Gamma(L_z,\varepsilon) = 7.374 - 0.511L_z + 0.0182L_z^2 - 0.054\varepsilon + 0.0161\varepsilon^2 \tag{5.35}$$

where Γ is the HWHM linewidth in meV, the well width is in nm, and ε is the field in mV/nm. This width, while fitted for the $N = 1$ electron to heavy hole resonance peak, is also a good approximation for the other exciton absorption resonances.

The full absorption spectrum can be obtained by convolving the exciton lineshapes with each other and with a function representing the continuum of transitions generating free electron-hole pairs. The latter can be modeled by a broadened Sommerfeld factor, which accounts for transitions between generated particle pairs of varying in-plane momenta. As an approximate alternative to convolution, the various absorption contributions can simply be summed. We therefore obtain, incorporating only the first two excitonic resonances, the following expression for the absorption coefficient of the QW:

$$\alpha(E) = \frac{\alpha_{hh}}{1+\left[\dfrac{(E-E^h)}{\Gamma}\right]^2} + \frac{\alpha_{lh}}{1+\left[\dfrac{(E-E^l)}{\Gamma}\right]^2} + \frac{2\alpha_c}{\left[1+\exp\left(\dfrac{E-E^c}{\Gamma_c}\right)\right]\left[1+\exp\left(\dfrac{-2\pi}{\left[(E-E^c)/R_y\right]^{1/2}}\right)\right]} \tag{5.36}$$

where E^h and E^l are given by (5.32) with N=1, α_{hh} and α_{lh} by (5.34), and Γ by (5.35). E^c is the energy of the continuum edge and is taken as equal to E^l at zero field, α_c and Γ_c are the magnitude and edge width, respectively, of the continuum, and are taken from experimental data, and R_y is the Rydberg constant of the material. Equation (5.36) has been demonstrated to yield accurate approximations to the absorption spectra of GaAs-AlGaAs QW's near the absorption edge [88].

5.6 Determination of the Carrier Wavefunctions The modeling of both optical gain and the QCSE in quantum wells requires a knowledge of the quantized sub-band energy levels, E_{cN} and E_{jN}. These are obtained by solving the time-independent Schrodinger equation:

$$\frac{-\hbar^2}{2m_i^*}\frac{\partial^2\psi_{iN}(z)}{\partial z^2}+V(z)\psi_{iN}(z)=E_{iN}\psi_{iN}(z) \tag{5.37}$$

where $i=c, j$ ($j=l,h$) denotes electrons or light or heavy holes, and V is the potential energy function. In our case, V is a finite potential well, rectangular in the case of zero applied electric field, but with sloping potential otherwise. In the literature, (5.37) has been solved using variational methods [90], Airy functions [91], and finite element methods [92]. Note that (5.37) has the same form as equation (4.18); we could therefore apply the numerical method described in section 4.1 for the analysis of circularly-bent planar waveguides. Here, however, we have applied another method which avoids the requirement to solve a matrix eigenequation of the form of (4.24). This is the so-called resonant tunneling technique [93]. To derive this method, we begin by writing the general solution for (5.37) in a region of constant potential:

$$\psi_{iN}(z)=A\exp(jkz)+B\exp(-jkz) \tag{5.38}$$

where A and B are constants, and:

$$k^2\equiv\frac{2m_i^*(E_{iN}-V(z))}{\hbar^2} \tag{5.39}$$

An equivalent solution which avoids complex exponentials is given by:

$$\psi_{iN}(z)=\frac{A}{k}\sin(kz)+B\cos(kz)\ ;\ k^2\geq 0 \tag{5.40.a}$$

$$\psi_{iN}(z)=\frac{A}{|k|}\sinh(|k|z)+\text{[illegible]}h(|k|z)\ ;\ k^2<0 \tag{5.40.b}$$

If k^2 is uniform over an interval $z_0 \leq z \leq z_0 + \delta z$, and if $z_0 = 0$, then (5.40) is the solution of (5.37) in the interval if:

$$A = \frac{\partial \psi_{iN}(z_0 = 0)}{\partial z} \; ; \quad B = \psi_{iN}(z_0 = 0) \qquad \textbf{(5.41)}$$

We can therefore write:

$$\begin{pmatrix} \psi_{iN}(z_0 + \delta z) \\ \dfrac{\partial \psi_{iN}(z_0 + \delta z)}{\partial z} \end{pmatrix} = \begin{pmatrix} \cos(k\delta z) & \dfrac{1}{k}\sin(k\delta z) \\ -k\sin(k\delta z) & \cos(k\delta z) \end{pmatrix} \begin{pmatrix} \psi_{iN}(z_0) \\ \dfrac{\partial \psi_{iN}(z_0)}{\partial z} \end{pmatrix} \; ; \; k^2 \geq 0 \quad \textbf{(5.42.a)}$$

$$\begin{pmatrix} \psi_{iN}(z_0 + \delta z) \\ \dfrac{\partial \psi_{iN}(z_0 + \delta z)}{\partial z} \end{pmatrix} = \begin{pmatrix} \cosh(|k|\delta z) & \dfrac{1}{|k|}\sinh(|k|\delta z) \\ |k|\sinh(|k|\delta z) & \cosh(|k|\delta z) \end{pmatrix} \begin{pmatrix} \psi_{iN}(z_0) \\ \dfrac{\partial \psi_{iN}(z_0)}{\partial z} \end{pmatrix} \; ; \; k^2 < 0 \textbf{(5.42.b)}$$

which is true for general z_0 as long as k^2 is uniform in the interval. Equations (5.42.a&b) therefore give us a transfer matrix formalism, by which we can relate the wave function and its derivative at the left side of an interval to that at the right side. The extension of this formalism to numerical computation of the energy eigenvalues and wave functions is straightforward. The problem region (barrier-well-barrier) is broken up into sections; in each section, the local potential is taken as fixed, thus providing a chain of intervals of uniform k^2. The sloped potential characteristic of a QW in an electric field, therefore, is approximated as a staircase; see Figure 5.5. Initial values of the wavefunction and its derivative (the wave function value can be arbitrary but the derivative value must be zero) are then set at the left side of the first interval, and, guessing a value of E_{iN}, the trial wave function is propagated from left to right through the problem region. For incorrect guesses of the energy eigenvalues, the trial wave function will grow exponentially in the far barrier, whereas eigenvalues will generate wave functions which decay exponentially in both barriers. Therefore, the eigenvalues E_{iN} and the corresponding wave functions are found simultaneously by bracketing and refinement.

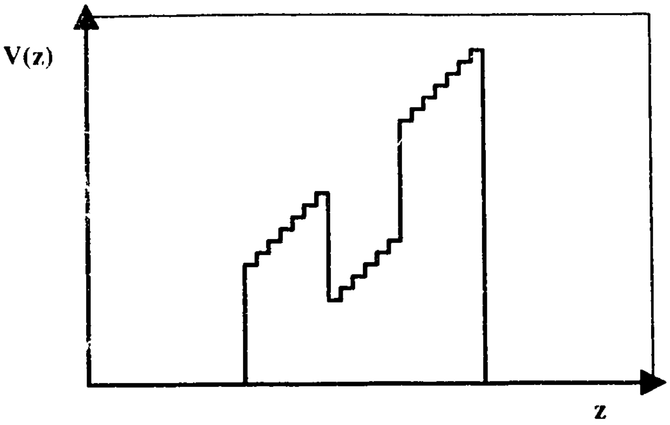


Figure 5.5 Resonant tunneling model approximation to the potential energy characteristic of a barrier-well-barrier system in an electric field.

5.7 LAMDA Model Algorithm and Simulation Results The theory described in the preceding sections was used as the basis of a computer program to model the behaviour of the LAMDA structure. The program assumes a single-cavity ridge waveguide structure wherein the upper (gain) QW stack is completely etched through (Figure 5.1). Therefore, a current passing from the top contact to the middle contact, without any shunt reductions, is taken as the injection current for the gain calculations. The total current is assumed evenly divided among the quantum wells in the upper stack for calculation of the resulting carrier concentrations using equation (5.29). The conduction band quasi-Fermi level is then calculated by solving equations (5.24) and (5.26)-(5.28) self-consistently using a bisection-type algorithm, where (5.26) and (5.27) are evaluated using Simpson's Rule. The valence band quasi-Fermi level is obtained from (5.25) using bisection, and the bulk gain coefficient can then be calculated using (5.20). The peak (net) mode gain is taken as:

$$g_m = \Gamma_g g - \Gamma_a \alpha_a - \left(1 - \Gamma_g - \Gamma_c\right)\alpha_r - \frac{1}{l}\ln\left(\frac{1}{\text{[illegible]}}\right) \qquad (5.43)$$

where g is the bulk gain coefficient from equation (5.20) and Γ_g and Γ_a are the mode intensity confinement factors in the gain and absorption quantum wells, respectively. These are calculated using the finite element method discussed in section 2.11, considering only the fundamental (TE) transverse mode if more than one mode exists. Since it uses only real refractive indices, the method of section 2.11 cannot account for gain-guiding effects; however, these are not expected to be very important near the lasing threshold. The absorption coefficients α_a and α_r denote the absorptions assumed in the bottom quantum wells, and in all other layers of the structure, respectively. The value of α_a is determined from equation (5.36) and depends on the configuration of the bottom QW stack and the assumed applied bias. The wavelength at which α_a is chosen is determined by the expected peak gain wavelength from (5.20). Equation (5.43) is evaluated over a range of injection current values to determine the laser threshold, taken as the current where the peak net mode gain g_m=0. The various parameters not so far listed explicitly, e.g. effective masses, band-gap energies and exciton binding energies, are widely available in the literature and therefore are not given here.

This procedure should provide a good estimate of the threshold current. A single-cavity laser/modulator can be modeled by varying the assumed electric field in the evaluation of (5.36). In practice, this is achieved by varying a reverse voltage bias applied across the middle and bottom contact terminals. The indicated lasing wavelength from (5.20) must be considered approximate, since we have neglected the wavelength shifting effects of line broadening. However, this effect is countered by a reduction of the band-gap energy which occurs under current injection. Note that the model determines α_a from equation (5.36) but supplies only a single value to the gain model. Therefore, we cannot account for wavelength chirp which may occur if the bottom quantum wells are biased such that the heavy hole exciton resonance is red-shifted past the lasing wavelength. This is of no consequence as, at that point, we have already reached the point of maximum intensity modulation, while below that point, no chirp is expected since no shift of the lasing wavelength can increase the mode gain.

The first step in the design of LAMDA was to develop a heterostructure profile which would produce a single-mode waveguide when etched to the bottom of the upper quantum wells, for a reasonable ridge width, while giving a high intensity confinement (>10%) in the top (gain) quantum wells and a low confinement (<2%) in the bottom quantum wells. The peak gain is expected at the N=1 electron to heavy hole transition; the calculated wavelength for this transition is 847.5 nm for 10 nm quantum wells. The material refractive index as a function of aluminum content at this wavelength was calculated using the model of Adachi [94] for application in the finite element algorithm of section 2.11. On this basis we arrived at the structure shown in Figure 5.6. A 0.2 μm-thick, n-doped $Al_{.18}Ga_{.82}As$ waveguide is clad from above by an intrinsic $GaAs/Al_{.32}Ga_{.68}As$ MQW active layer with six 10 nm wells and seven 10 nm barriers. A 2 μm-thick, p-doped $Al_{.51}Ga_{.49}As$ layer constitutes the remainder of the upper cladding, and a thin, heavily-doped GaAs cap forms an ohmic contact with the top bias terminal. A 3 μm-wide ridge etched to the top of the waveguide layer supports a single TE mode, with an intensity confinement factor in the top QW layer of approximately 10.5% (i.e. Γ_g=5.25%). A 10 nm AlAs layer between the waveguide layer and the top quantum well stack was included as an etch-stop layer.

A lower cladding is formed by two layers of $Al_{.28}Ga_{.72}As$ separated by a single GaAs quantum well. The upper AlGaAs layer is n-doped and 0.2 μm thick, the lower is p-doped and 3 μm thick. The SQW layer is operated under reverse bias and used for intensity modulation. A well width of 8 nm was chosen to offset the absorption peak from the lasing wavelength of the upper active layer and reduce residual, zero-bias losses. Splitting the lower cladding vertically serves a dual purpose. By varying the position of the lower quantum well we have tailored our design to have an acceptable confinement factor in that layer: the calculated value is approximately (Γ_a=) 0.5%. In addition, the split cladding results in a thicker n-doped, middle contact layer than would be available if the SQW were adjacent to the core. This is advantageous since the middle contact of the device must support the lasing current. Peak doping levels were $5X10^{17}$ cm^{-3} (n) and $5X10^{18}$ cm^{-3} (p). Cleaved, uncoated facets were assumed and hence R was taken as 0.3 in equation (5.43).

In Figure 5.7, we show the spectral absorption coefficient of the lower, 8 nm quantum well, modeled using (5.32)-(5.42). The electric field is calculated from the reverse bias voltage assuming a 0.1 μm-thick intrinsic region around the quantum well. At zero bias, the heavy hole absorption peak is at about 843 nm wavelength. At the predicted operating wavelength of 847.5 nm, the absorption coefficient is about 7000 cm^{-1}, corresponding to a modal absorption of $\Gamma_a\alpha_a$=35 cm^{-1}. At a reverse bias of 2 V, the absorption peak is already red-shifted past the operating wavelength. A maximum increase in $\Gamma_a\alpha_a$ of about 55 cm^{-1} is predicted at approximately -1 V bias. In practice, the magnitude of the spectral shift of the exciton resonances with reverse bias will depend strongly on the true thickness of the intrinsic region around the quantum well.

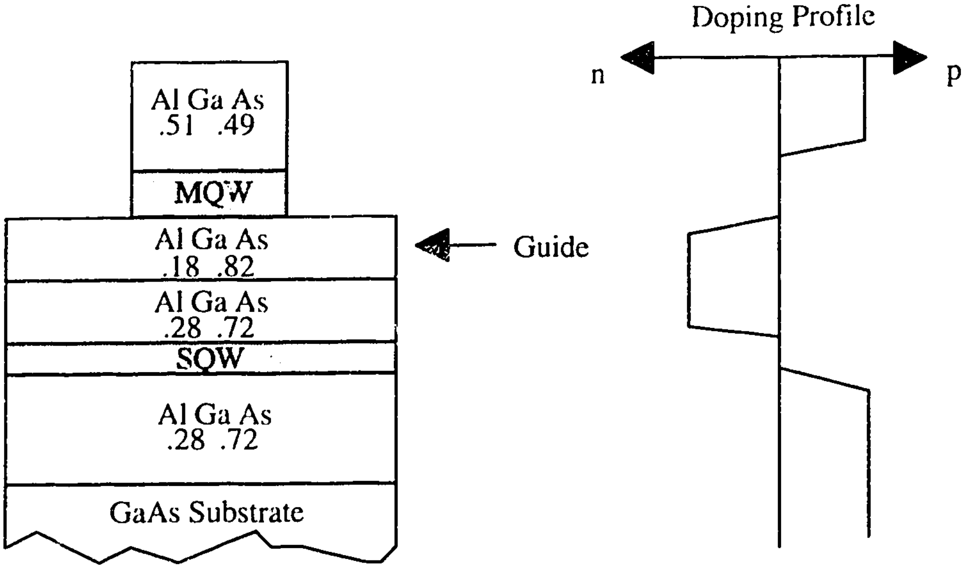


Figure 5.6 Schematic of the LAMDA structure design; layer compositions and doping profile. Layer thicknesses are not to scale. A thin GaAs cap at the top of the structure is not shown.

In Figure 5.8, we show the calculated values of the peak net mode gain, given by equation (5.43), for three cleaved cavity lengths: 250 μm, 300 μm and 500 μm. We have taken the residual absorption coefficient value α_r=20 cm^{-1}, and the modulator quantum well is assumed unbiased. The predicted threshold currents are given by the points where the gain curves intersect the g_m=0 line; for all three cavity lengths, the lasing wavelength is predicted to be 847.5 nm, i.e. the peak gain is attained on the N=1 electron to heavy hole transition. Anticipating the experimental results discussed in the next section, the actual measured thresholds for the three cavity lengths are indicated by asterisks. Agreement between theory and experiment is excellent. The lowest threshold current, for a 250 μm cavity, is 57 mA. Note that for all three cavity lengths, at the maximum predicted absorption coefficient of the lower quantum well, threshold currents in excess of 100 mA are predicted. Therefore, with a steady-state current applied to the upper quantum wells, we predict that very dramatic intensity modulation can be obtained in a single cavity by modulating the bias applied to the bottom quantum well, from 0 V to -1 V.

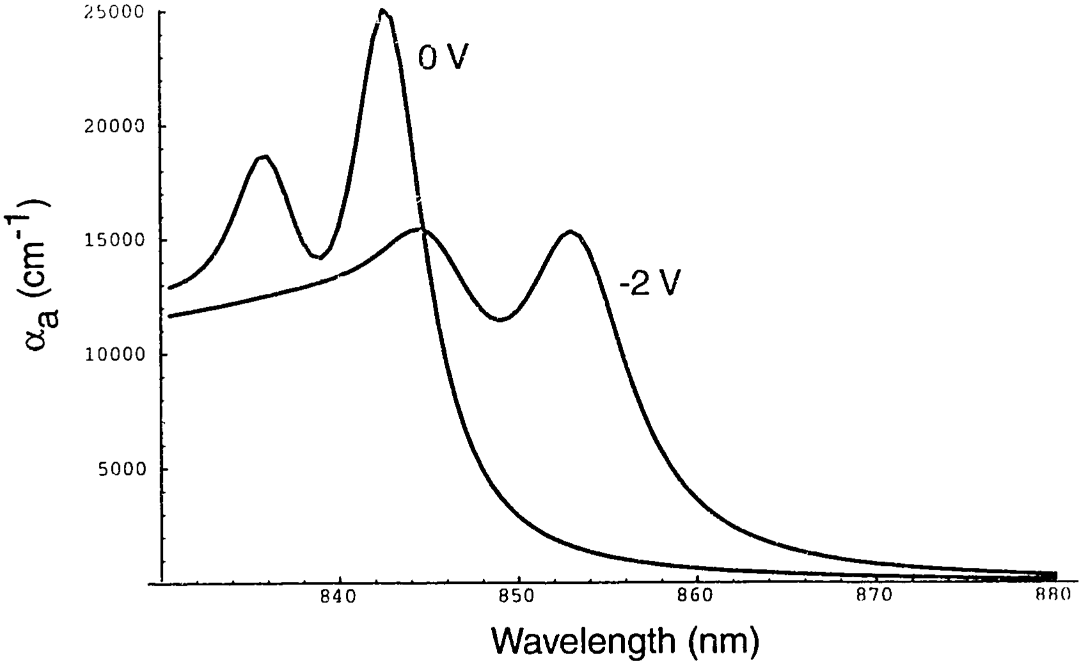


Figure 5.7 Modeled absorption spectra of the 8 nm modulator quantum well of the LAMDA structure, under 0 V and -2 V bias.

For any given value of the total modal absorption loss in the cavity there is a specific number of (gain) quantum wells which provides the optimum mode gain and the lowest threshold current. This optimum number of wells results from the competition of two effects. Increasing the number of wells reduces the recombination current in each well, thereby reducing the carrier concentrations, the quasi-Fermi levels and the bulk gain per well. However, the modal gain is the product of the bulk gain and the total optical confinement in the wells, and therefore increases with the number of quantum wells. A good discussion of this problem is given in [95]. In our case, with the predicted values of $\Gamma_a\alpha_a$ and $(1-\Gamma_g-\Gamma_a)\alpha_r$, the calculated optimum number of gain quantum wells is six.

It is interesting to consider the effects of the carrier overflow mechanisms, discussed in section 5.3, on the performance of the device. In Figures 5.9-5.11, we consider the LAMDA structure as designed, but with only a *single* quantum well providing gain. Figure 5.9 shows the unconfined electron concentration in the quantum well, n_{un}, as a fraction of the total electron concentration n, for two cases: a 7 nm quantum well, and a 10 nm quantum well, as a function of the total injected current. This parameter increases, roughly linearly, with current for both quantum well widths. However, the fraction of unconfined electrons is much higher in the 7 nm quantum well-filling of the confined energy states occurs faster in the narrower quantum well.

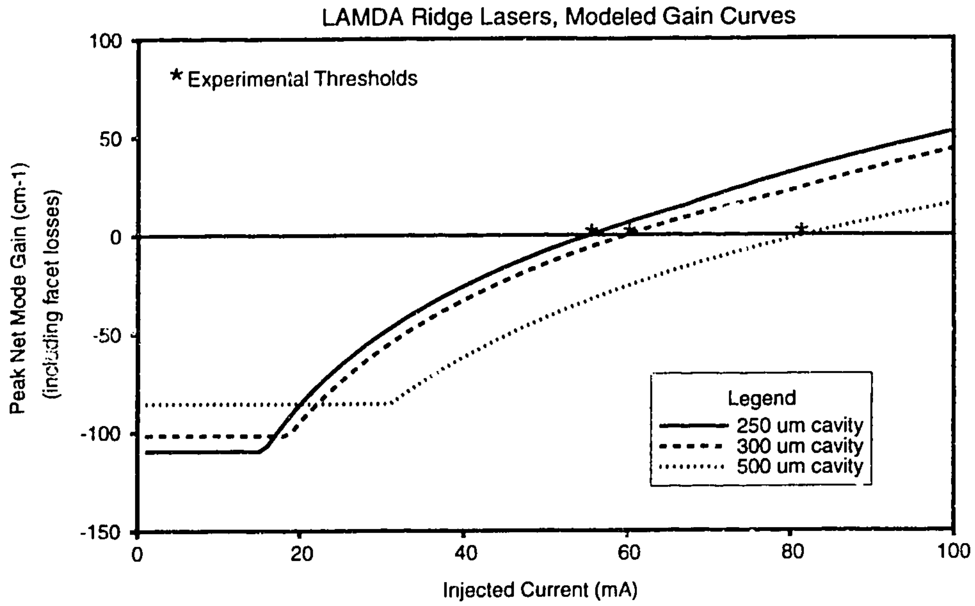


__Figure 5.8__ Calculated peak net mode gain curves of the LAMDA structure for three different cavity lengths. The predicted threshold currents are given by the intersections of the curves with the mode gain = 0 line; the experimentally measured threshold currents are shown for comparison. Setting the residual absorption coefficient to the standard value α_r=20 cm^{-1} leads to an excellent match between theory and experiment.

The total calculated overflow current, into unconfined well states and into the barriers, is shown as a fraction of the injected current, for the same two cases, in Figure 5.10. At 100 mA current, the overflow current in the 7 nm well reaches more than 15% of the total, while that in the 10 nm well reaches about 3%. When *six* 10

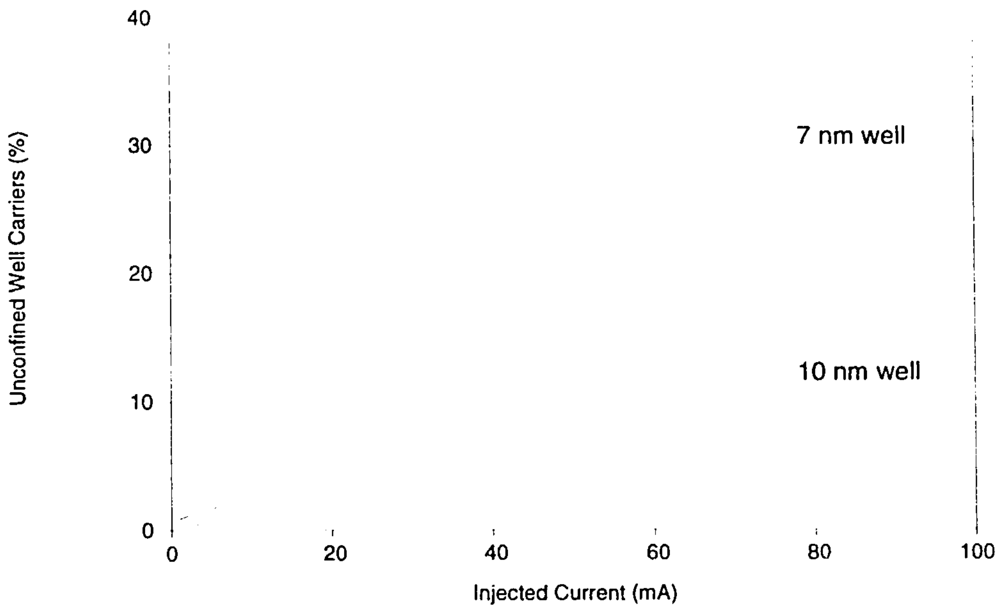


Figure 5.9 Unconfined electron concentration in the quantum well, n_{un}, as a fraction of the total electron concentration n, as a function of the total injected current, for two cases: a 7 nm quantum well (solid line), and a 10 nm quantum well (dashed line).

nm quantum wells are considered (not shown), the calculated overflow current drops to less than 0.05% at 100 mA total current. Clearly, the width of the quantum wells, and their number, have a dramatic influence on the impact of the current overflow mechanisms. This is confirmed by considering the calculated conduction band quasi-Fermi levels as a function of the injection current. In Figure 5.11, this is shown for both a single 7 nm well and a single 10 nm well, with (solid lines) and without (dashed lines) considering the carrier overflow mechanisms in the calculations. In both cases, band-filling effects are evident only at high injection current levels, and they are markedly less important for the wider quantum well. Again, when six quantum wells are used in the calculations, the carrier overflow mechanisms are even less important; in this case, the resulting quasi-Fermi level reductions are negligible. This is a new finding since the impact of carrier overflow on quantum well gain has only been assessed in the literature for the case of narrow, single quantum wells.

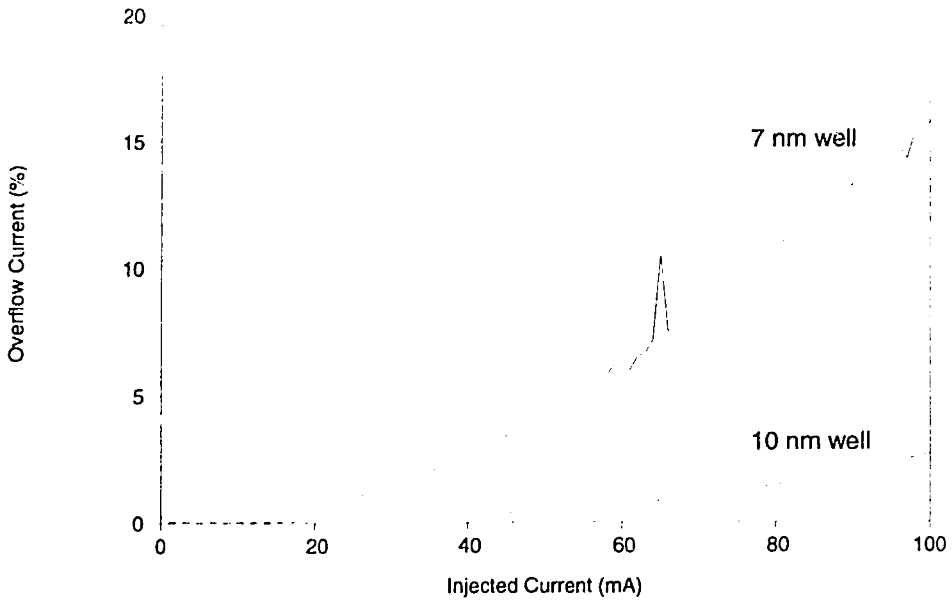


Figure 5.10 Total calculated electron overflow as a fraction of the total injected current for a 7 nm quantum well (solid line), and a 10 nm quantum well (dashed line).

5.8 LAMDA Structure Fabrication and Experimental Results

The MBE-grown LAMDA structure described above was photoresist-patterned with 3 μm lines to define the waveguide ridges. Following reactive ion etching to a depth of 1.8 μm, wet etching using 95:5 H_2O_2:NH_4OH completed the ridge definition by etching to the etch-stop depth. In later samples, a second reactive ion etch was done over a larger area surrounding the rib, to a depth of at least 0.5 μm to isolate the bottom quantum well between adjacent devices and hence [illegible] capacitance. Oxide and polyimide were deposited to planarize the etched wafer. Areas for the middle (n) and top (p) contacts were patterned for removal of the polyimide using reactive ion etching, and removal of the oxide using wet etching. Patterning, evaporation and lift-off of Ni/Ge/Au on the n-level, and Ti/Pt/Au on the p-level, followed. Finally, the wafer was lapped to a thickness of 100 μm and Ti/Pt/Au evaporated onto the back (p) surface. The metal surfaces were rapidly thermally annealed for 15 s in flowing H_2/N_2 at 415 C prior to cleaving of the samples at lengths of 250 μm, 300 μm and 500 μm.

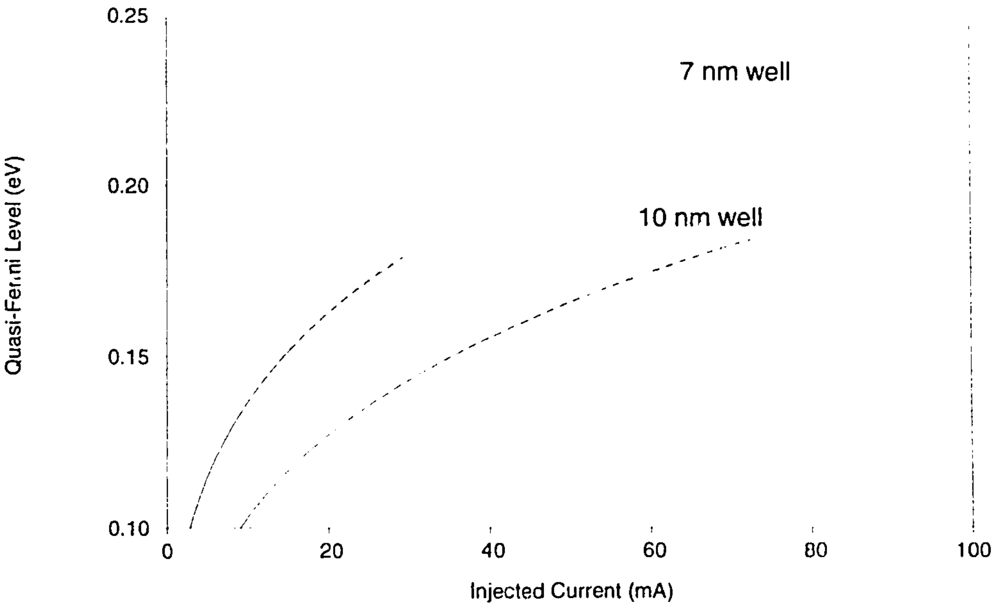


Figure 5.11 Calculated conduction band quasi-Fermi levels, as a function of the injection current for a single 7 nm well and a single 10 nm well, with (solid lines) and without (dashed lines) considering the carrier overflow mechanisms in the calculations.

The devices were tested as single-cavity laser/modulators, as modeled, using a custom-designed pulsed-current driver for the laser. The devices were mounted on a floating potential copper block. The middle (n) contacts of the devices were grounded using a probe. A forward bias was supplied to the top contact from the current-controlled driver, also using a probe, while a negative potential was applied to the copper block to reverse bias the lower, modulator diode. Unfortunately, the devices were only operable in a very low duty cycle (<2.5%) pulsed mode, since larger duty cycles lead to electrical destruction of the devices, even when the devices were mounted on an actively-cooled heat sink. The reasons for this are unknown, although we suspect that a large contact resistance existed at the middle (n) level which may have lead to high and concentrated resistive heating at that location. This suspicion is supported by the unusually high observed forward bias resistance of the upper (gain) diode: ~80 Ω in most cases, and also by the very low device yield (fraction of total devices which were operable) and poor appearance of the n-contacts under scanning electron microscope examination. With such a low duty cycle and correspondingly low average optical power output from the device, the emission wavelength could unfortunately not be measured. Furthermore, the modulation bandwidth, with the laser emitting CW and the modulator quantum well driven under reverse bias, also could not be evaluated.

Peak optical power versus bias current characteristics were measured for all cavity lengths. Threshold currents of 57, 60 and 81 mA were observed for cavity length[illegible]0, 300 and 500 µm, respectively, in excellent agreement with the theo[illegible]tions shown in Figure 5.8. Threshold modulation by applying a reve[illegible]e bottom quantum well was also demonstrated. In Figure 5.12, we show [illegible]er versus current characteristic of a 300 µm-long device, with 0 V (solid line) and -4 V (dashed line) applied to the modulator. A threshold current increase of about 5 mA, and a maximum optical power modulation of 6 dB, is obtained. No further threshold increase is observed beyond -4 V bias. According to the model, the observed modulation corresponds to an absorption coefficient increase, in the lower quantum well, of only about 2400 cm^{-1}. This is well below the expected value of about 11000 cm^{-1} at -1 V bias, indicating that gain-guiding effects may be significant above threshold. The observed modulation of the threshold current with reverse bias was independent of the laser duty cycle, which was varied between 0.025% and 2%. Also, the "leakage" current from the bottom (modulator) diode under -4 V bias was less than 10 µA, through a 1 kΩ resistor.

An attempt was made to measure the absorption spectra of the upper and lower quantum wells in the devices, and thereby determine their quality and characteristics by comparing the observed exciton resonances with the theoretical predictions. This was accomplished by focusing light from a tunable Ti:Sapphire laser (see Figure 4.14) into the devices with the aid of an IR vidicon camera, and measuring the collected photocurrent, a direct indication of the absorption magnitude. To measure the absorption spectrum of the top quantum well stack, the light was focused into the waveguide ridges and the current collected between the top and middle contacts. For the bottom quantum well spectrum, the light was focused into the slab waveguide layer between broadly-separated ridges, and

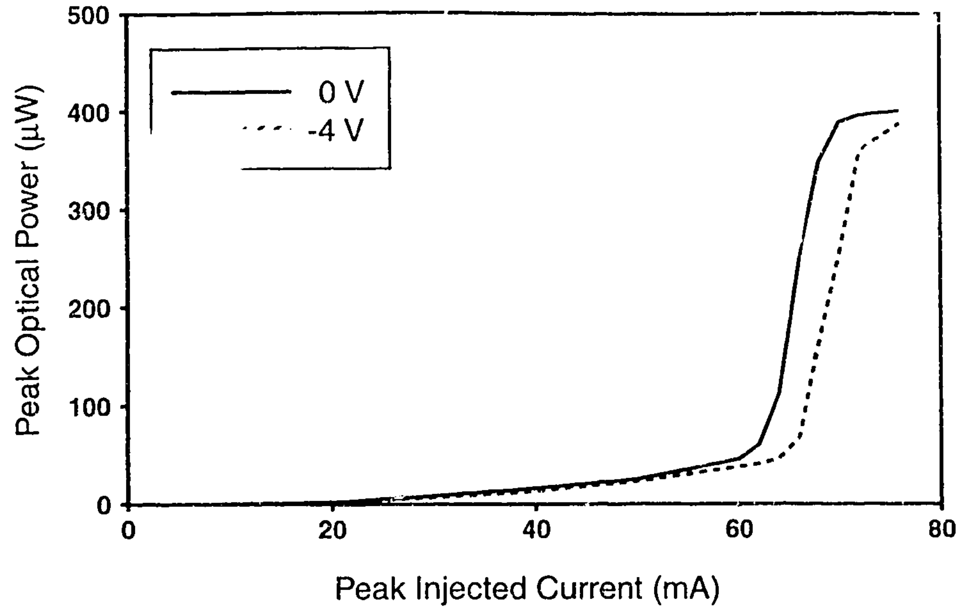


Figure 5.12 Experimental power versus current curves for 300 μm cavity LAMDA ridge laser with 0 V (solid line) and -4 V (dashed line) applied to the modulator. The threshold current is increased by about 5 mA with the application of the bias, and about 6 dB of optical power modulation is achieved.

the resulting photocurrent was collected between the middle and bottom contact terminals. No bias was applied to the bottom quantum well for these measurements. The incident beam was chopped and the photocurrent measurements were made using a lock-in amplifier. The polarization of the light could be controlled by rotating the half-wave plate. Typical results of the measurements are shown in Figure 5.13; the measured photocurrent for each curve is normalized to unit peak magnitude. The solid curve shows the measured absorption characteristic for the top quantum well stack. This curve is consistent with theory, which predicts, for 10 nm wells, exciton absorption peaks at 851.2 nm (N=1 electron-heavy hole) and 846.3 nm (N=1 electron-light hole). The measured characteristics for the bottom, 8 nm quantum well are poor and show high absorption at long wavelengths (>860 nm). The reason for this is unclear. The TE-polarized scan shows evidence of an absorption peak at 843 nm, in agreement the predicted position of the N=1 electron-heavy hole exciton transition. This feature disappears when the incident light is TM-polarized, also consistent with theory. However, the general mismatch of the form of the observed absorption spectra for the 8 nm quantum well with the predicted spectra (see Figure 5.7) indicates that this quantum well may be of poor quality.

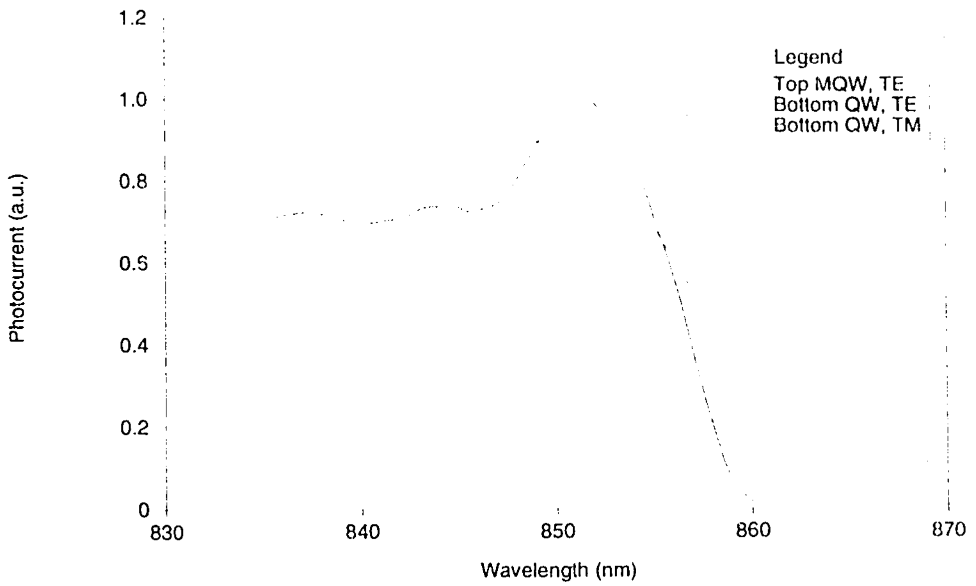


Figure 5.13 Measured absorption (photocurrent) spectra of the top (solid line) quantum wells and bottom quantum well (dashed line) and TE-polarized light, and the bottom quantum well with TM-polarized light (dotted line).

5.9 Modulation Bandwidth of Single-Cavity LAMDA Laser/Modulator

Although CW operation of the fabricated LAMDA lasers could not be achieved, and hence a bandwidth evaluation of the modulated lasers could not be carried out, we can nevertheless examine the theoretical bandwidth limit of the single-cavity LAMDA laser/modulator device. The frequency limitations of the single-cavity configuration are not immediate clear, since, while the electroabsorption mechanism itself is RC-limited, this mechanism interacts with a large population of resonating photons of finite mean lifetime. To gain some insight into the frequency response of the device, we therefore conduct a small signal analysis using the appropriate rate equations for the carrier and photon populations [8]:

$$\frac{dn}{dt} = \frac{i}{eV_{QW}} - \frac{n}{\tau_{sp}} - \Gamma_g \frac{c_o}{n_{cav}} g \frac{P}{V_{QW}} \qquad \textbf{(5.44.a)}$$

$$\frac{dP}{dt} = \left\{ \Gamma_g g - \Gamma_a \alpha_a - \left(1 - \Gamma_g - \Gamma_a\right)\alpha_c - \frac{1}{l}\ln\left(\frac{1}{R}\right) \right\} \frac{c_o}{n_{cav}} P \qquad \textbf{(5.44.b)}$$

where P is the number of photons in the lasing mode, i is the injected current assuming overflow mechanisms are negligible, V_{QW} is the volume of the gain quantum well(s), and τ_{sp} is an effective spontaneous emission-limited carrier lifetime. Other parameters are as previously defined. Note that no spontaneous emission rate term is included in (5.44.b), since we assume that the spontaneous emission coupling to the lasing mode is negligibly small. Let us now assume that all [illegible] quantities vary only slightly around equilibrium (static) values. If we denote these static values by the subscript o then we can write:

$$P = P_o + \Delta P \qquad \textbf{(5.45.a)}$$

$$n = n_o + \Delta n \qquad \textbf{(5.45.b)}$$

$$\alpha_a = \alpha_{a,o} + \Delta\alpha_a \qquad \textbf{(5.45.c)}$$

$$g = g_o + g'\Delta n \qquad \textbf{(5.45.d)}$$

where:

$$g' \equiv \frac{\partial g}{\partial n} \qquad \textbf{(5.46)}$$

The parameter g' is referred to as the differential gain. Substituting (5.45) into (5.44), we obtain:

$$\frac{d(\Delta n)}{dt} = \frac{i}{eV_{QW}} - \frac{n_o}{\tau_{sp}} - \frac{\Delta n}{\tau_{sp}} - \Gamma_g \frac{c_o}{n_{cav}} g_o \frac{P_o}{V_{QW}} - \Gamma_g \frac{c_o}{n_{cav}} g_o \frac{\Delta P}{V_{QW}} - \Gamma_g \frac{c_o}{n_{cav}} g'\Delta n \frac{P_o}{V_{QW}} \textbf{(5.47.a)}$$

$$\frac{d(\Delta P)}{dt}=\left\{\Gamma_g g_o+\Gamma_g g'\Delta n-\Gamma_a\alpha_{a,o}-\Gamma_a\Delta\alpha_a-\left(1-\Gamma_g-\Gamma_a\right)\alpha_r-\frac{1}{l}\ln\left(\frac{1}{R}\right)\right\}\frac{c_o}{n_{cav}}P_o$$

$$+\left\{\Gamma_g g_o-\Gamma_a\alpha_{a,o}-\left(1-\Gamma_g-\Gamma_a\right)\alpha_r-\frac{1}{l}\ln\left(\frac{1}{R}\right)\right\}\frac{c_o}{n_{cav}}\Delta P \qquad \textbf{(5.47.b)}$$

where all products of differentials (Δ terms) have been neglected as a first-order approximation. Equations (5.47.a&b) can be simplified by first writing their steady-state solutions, i.e. their solutions when the time derivatives and the Δ terms are equal to zero. In this way we obtain the following:

$$n_o=\tau_{sp}\left\{\frac{i}{eV_{QW}}-\Gamma_g\frac{c_o}{n_{cav}}g_o\frac{P_o}{V_{QW}}\right\} \qquad \textbf{(5.48.a)}$$

$$\Gamma_g g_o-\Gamma_a\alpha_{a,o}-\left(1-\Gamma_g-\Gamma_a\right)\alpha_r-\frac{1}{l}\ln\left(\frac{1}{R}\right)=0 \qquad \textbf{(5.48.b)}$$

Substitution of (5.48) into (5.47) yields:

$$\frac{d(\Delta n)}{dt}=-\frac{\Delta n}{\tau_{sp}}-\Gamma_g\frac{c_o}{n_{cav}}g_o\frac{\Delta P}{V_{QW}}-\Gamma_g\frac{c_o}{n_{cav}}g'\Delta n\frac{P_o}{V_{QW}} \qquad \textbf{(5.49.a)}$$

$$\frac{d(\Delta P)}{dt}=\left\{\Gamma_g g'\Delta n-\Gamma_a\Delta\alpha_a\right\}\frac{c_o}{n_{cav}}P_o \qquad \textbf{(5.49.b)}$$

Let us now assume harmonic solutions of equations (5.49), i.e. :

$$\Delta\alpha_a=\alpha_1 e^{j\omega t} \qquad \textbf{(5.50.a)}$$

$$\Delta P=-P_1 e^{j\omega t} \qquad \textbf{(5.50.b)}$$

$$\Delta n = n_1 e^{j\omega t} \tag{5.50.c}$$

where we have assumed that the modulation of P is out of phase with the modulation of α_a. The variables with subscripts 1 denote the magnitudes of the harmonic functions. Using (5.50) we can rewrite (5.49) as follows:

$$j\omega n_1 = -\frac{n_1}{\tau_{sp}} + \Gamma_g \frac{c_o}{n_{cav}} g_o \frac{P_1}{V_{QW}} - \Gamma_g \frac{c_o}{n_{cav}} g' n_1 \frac{P_o}{V_{QW}} \tag{5.51.a}$$

$$-j\omega P_1 = \left\{ \Gamma_g g' n_1 - \Gamma_a \alpha_1 \right\} \frac{c_o}{n_{cav}} P_o \tag{5.51.b}$$

We want to examine the frequency dependence of the photon population modulation, i.e.:

$$\frac{\Delta P}{\Delta \alpha_a} = \frac{P_1}{\alpha_1} \tag{5.52}$$

We do this by solving (5.51.a) for n_1 and substituting into (5.51.b). The result is:

$$\frac{P_1}{\alpha_1} = \frac{\Gamma_a \frac{c_o}{n_{cav}} P_o \left\{ j\omega + \frac{1}{\tau_{sp}} + \Gamma_g \frac{c_o}{n_{cav}} g' \frac{P_o}{V_{QW}} \right\}}{\left\{ -\omega^2 + \left(\Gamma_g \frac{c_o}{n_{cav}} \right)^2 g_o g' \frac{P_o}{V_{QW}} \right\} + j\omega \left(\frac{1}{\tau_{sp}} + \Gamma_g \frac{c_o}{n_{cav}} g' \frac{P_o}{V_{QW}} \right)} \tag{5.53}$$

This expression is quite similar to that obtained for a directly current-modulated quantum well laser with a fixed intra-cavity absorption, except that a term to the first order in ω exists in the numerator. We can define a resonance or relaxation oscillation frequency:

$$\omega_r \cong \Gamma_g \frac{c_o}{n_{cav}} \sqrt{g_o g' \frac{P_o}{V_{QW}}} \qquad (5.54)$$

which defines the approximate bandwidth of the device, since the frequency response rolls off at a rate of about -40 dB/decade above this frequency. Note that a large bandwidth requires a large intra-cavity absorption in the static state (i.e. reverse bias on the modulator), since this leads to large values of g_o and P_o. Furthermore, bandwidth is enhanced by a large value of Γ_g and g', both of which are promoted by using a large number of gain quantum wells.

5.10 Conclusions We have conceived, fabricated and tested a monolithic semiconductor structure which can achieve laser and electroabsorption modulation functions simultaneously at the same wavelength. This device was fabricated using a single, uniform epitaxial growth and relatively simple post-growth processing. The device is unique in that three contact terminals are provided, allowing for independent biasing of the gain and modulation active regions of the p-i-n-i-p-doped device. Computer models were developed and used to design the AlGaAs structure; we have presented modeled results which are in good agreement with the experimental observations of the laser threshold currents in cleaved devices of varying lengths. Threshold modulation by varying the reverse bias applied to the bottom quantum well stack was demonstrated; about 6 dB of intensity modulation, at -4 V bias, was achieved. A discussion of the frequency characteristics of the single-cavity laser/modulator device was also presented. The work of this chapter demonstrates that multi-functional monolithic integration of optoelectronic circuits can be achieved without resorting to extraordinary fabrication methods, such as patterned epitaxy and regrowth.

A number of improvements to the design of the structure are possible. The laser threshold currents could be reduced by redesigning the heterostructure profile to achieve a higher mode confinement in the 10 nm quantum wells, and by making the lower quantum well narrower (e.g. 7 nm) to reduce residual intra-cavity losses at zero bias. The latter measure might permit the incorporation of more than one quantum well in the modulator, which should improve the quality of the wells due to the smoothing effect of growing thin heterostructure layers. We might also improve the quality of the middle contact electrode by ramping the doping to a high value just before the AlAs etch stop layer, and by using the lowest possible Al content in the waveguide layer. These measures might lead to lower resistive heating at the contact and make CW laser operation possible. If such operation could be achieved, high-frequency modulation of single-cavity devices, and segmented multi-functional integrated photonic circuits, might be demonstrated.

Chapter 6

Spontaneous Emission Coupling to Radiation and Guided Modes of Planar Waveguide Structures

The calculation of spontaneous emission rates is important for many purposes. For example, the principle of waveguide capture of spontaneous emission is important for certain sensor applications [96]-[101]. Noise in erbium-doped fiber amplifiers is generated by captured spontaneous emission which is subsequently amplified [102]. Further, the dynamic behaviour and the spectral characteristics of semiconductor lasers depend on the total rate of spontaneous emission, and more particularly on the fraction of spontaneously emitted photons which are coupled into the lasing mode (the so-called spontaneous emission factor) [103],[104].

The rate of spontaneous emission from an atomic dipole in a homogeneous dielectric medium is well known [8]. Since the rate is proportional to $n^3/\varepsilon = n/\varepsilon_o$, inhomogeneities in refractive index, n, affect the spontaneous emission rate locally. In particular, the existence of guided modes complicates the calculation of spontaneous emission rates, since they must be treated in a mathematically different fashion than radiation modes, which extend throughout space. This issue was encountered in Chapter 5 when we derived expressions for optical gain and the spontaneous emission rate in quantum well lasers. In that case, we simplified the theoretical treatment by calculating these parameters in a uniform dielectric cavity. In practice, the rates of absorption, and spontaneous and stimulated emission, are affected by the presence of a waveguide, particularly with respect to guided modes which are profoundly different in form than radiation modes, as discussed in Chapter 2.

There have been theoretical treatments of such problems using approximate quantum mechanical [105],[106] or semi-classical approaches [107],[108] which may, in fact, be adequate for many applications. However, many of these models seem pedagogically unsatisfying since they do not clearly differentiate the contributions of the bound and radiating mode fields to spontaneous emission. Other studies tend to focus on Fabry-Perot structures with high reflectivity mirrors, where the primary focus is on the pattern of radiated power, while the consideration of guided modes, if they exist, is secondary [109]-[112]. A recent paper [113] demonstrated that the spontaneous emission characteristics of compound planar structures with guided modes may be treated by adding a high index substrate to the structure, converting bound modes into leaky modes which may then be handled uniformly with the true radiation modes of the structure In [113], the radiated power was again the primary concern, but the effects of the guided mode fields were conveniently and accurately included. However, applications wherein the guided modes themselves are of primary concern may be better served by a different approach.

In this chapter, we present a theoretical consideration of spontaneous emission in the vicinity of a symmetric slab waveguide. Quantum mechanical expressions for the spontaneous emission rate are derived and solved by direct substitution of explicit functions for the complete set of bound and radiating modes. In this way, the contributions of the guided and radiating field components to spontaneous emission are clearly differentiated and separately accessible, in a more direct way, albeit for a simpler case, than the method presented in [113]. Extension of the treatment to more complicated structures is discussed in the chapter conclusions. We apply the present results to an analysis of spontaneous emission coupling to the guided modes of a planar erbium-doped amplifier, a subject of considerable current interest [114]. Spontaneous emission factors (the fraction of the total emission coupled to specific guided modes) are calculated for the cases of erbium-doping of the guide and of the cladding. It is shown that when guided mode optical pumping of the erbium is used, high spontaneous emission coupling to the guided signal modes is obtained when the guide layer is doped, and that the coupling to TM modes is greater than that to TE modes when the atomic dipoles are assumed randomly oriented. Negligible spontaneous emission factors are obtained when the cladding layers are erbium-doped.

The work of this chapter is presented in [115].

6.1 Theoretical Formulation We begin by finding the correct expression for the quantized electric field in a cubic dielectric box of side L and index n_2 (Figure 6.1) which has a slab of thickness a and refractive index n_1 placed in the center perpendicular to $\hat{x}$. We shall eventually take the limit $L \rightarrow \infty$. Following Yariv [8], the orthogonal function expansion of the field has the form:

$$\bar{E}=\sum_{knm}\bar{E}_{knm}=\sum_{knm}-jN_k\left\{a^{+}_{knm}\frac{e^{-j(\beta^y_n y+\beta^z_m z)}}{L}\bar{F}_k(x)-a_{knm}\frac{e^{j(\beta^y_n y+\beta^z_m z)}}{L}\bar{F}^{*}_k(x)\right\} \qquad (6.1)$$

where $\bar{F}_k(x)\exp[\mp j(\beta^y_m y+\beta^z_n z)]$ is one of a complete set of modes of the slab waveguide structure found by solving Maxwell's equations with the appropriate boundary conditions. Evanescent modes are excluded by the arguments of section 2.8. The exponential functions reflect the fact that mode fields can propagate in arbitrary directions in the y-z plane, and the $1/L$ factors normalize these functions. Periodic boundary conditions are assumed to apply:

$$\beta^y_n=\frac{2\pi n}{L} \quad ; \quad \beta^z_m=\frac{2\pi m}{L} \qquad (6.2)$$

where n and m are integers. In the limit of large L, β^y_n and β^z_m become continuous. The factors a^{+}_{knm} and a_{knm} are interpreted as creation and annihilation operators, the phase factor $-j$ follows from a phaseless mode vector potential, and the coefficients N_k are obtained by correct normalization of the mode energy as follows.

The energy of mode nmk is given classically by:

$$\frac{1}{2}\iiint_V\left(\varepsilon\left|\bar{E}_{nmk}\right|^2+\mu\left|\bar{H}_{nmk}\right|^2\right)dV=\iiint_V\varepsilon\left|\bar{E}_{nmk}\right|^2 dV \qquad (6.3)$$

where we have used:

$$\iiint_V\varepsilon\left|\bar{E}_{nmk}\right|^2 dV=\iiint_V\mu\left|\bar{H}_{nmk}\right|^2 dV \qquad (6.4)$$

which is generally true for harmonic solutions of Maxwell's equations. In our case, it can be shown that due to the periodicity conditions (6.2) and the commutation relation:

$$\left[a_k,a_k^{+}\right]=a_k a_k^{+}-a_k^{+}a_k=1, \qquad (6.5)$$

$$\iiint_V \varepsilon\left|\vec{E}_{nmk}\right|^2 dV = \left\{a_{knm}a^{\dagger}_{knm} + a^{\dagger}_{knm}a_{knm}\right\}N_k^2 \int_{L/2}^{L/2} \varepsilon(x)\left|\vec{F}_k(x)\right|^2 dx$$

$$= \left(1 + 2p_{knm}\right)N_k^2 \int_{L/2}^{L/2} \varepsilon(x)\left|\vec{F}_k(x)\right|^2 dx \qquad \textbf{(6.6)}$$

where p_{knm} is the photon number operator. Now, (6.6) must equal the quantized energy of the mode, $\hbar\omega_{knm}(1/2 + p_{knm})$, where $\hbar\omega_{knm}/2$ is the zero-point energy and ω_{knm} is the radial frequency of the mode.

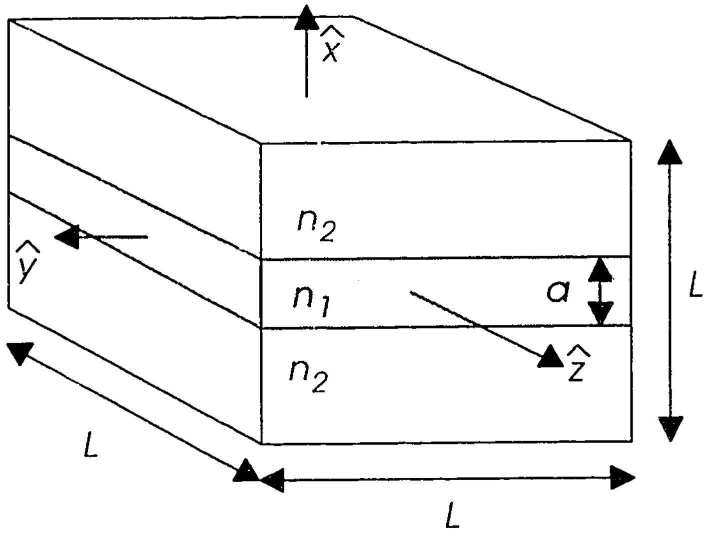


Figure 6.1 The symmetric slab dielectric waveguide. A guide layer of refractive index n_1 and thickness a is clad by material of index n_2; $x = 0$ at the centre of the guide layer. Spontaneous emission rates are examined in the limit as $L \to \infty$.

Therefore, the normalization constants are given by:

$$N_k^2 = \frac{\hbar\omega_{knm}}{2\int_{-L/2}^{L/2} \varepsilon(x)\left|\vec{F}_k(x)\right|^2 dx} = \frac{\hbar\omega_{knm}}{2\varepsilon_o\int_{-L/2}^{L/2} n^2(x)\left|\vec{F}_k(x)\right|^2 dx} \qquad \textbf{(6.7)}$$

In the Appendix we have listed the appropriate functions $\bar{F}$, derived for guided and radiating TE and TM modes, together with their appropriate normalization expressions obtained by evaluating (6.7) in the limit as $L \to \infty$.

The rate of transition of a system from state $|2, p_l\rangle$ to state $|1, p_l + 1\rangle$, where 2,1 denote the initial and final states of an emitting atom, and p_l denotes the number of photons, p, in the lth optical mode, is given by the Golden Rule:

$$W_l = \frac{2\pi}{\hbar} \left| \langle 1, p_l + 1 | H_l' | 2, p_l \rangle \right|^2 \delta(E_2 - E_1 - \hbar\omega_l) \tag{6.8}$$

taken in the limit of purely harmonic emission (δ is the Dirac delta), an acceptable approximation for narrow-linewidth atomic emitters. E_l and E_2 are the final and initial atomic eigenenergies, and H_l' is the perturbation Hamiltonian causing the transition. In the dipole approximation:

$$H_l' = -e\bar{E}_l \bullet \bar{r} \tag{6.9}$$

where $\bar{E}_l$ is the time-independent electric field of the lth optical mode, and $\bar{r}$ is the position operator. Using (6.1), the transition rate (6.8) becomes in our case:

$$W_l \to W_{knm} = \frac{2\pi e^2 N_k^2}{\hbar L^2} \left| \langle 1, p_{knm} + 1 | a_{knm}^{+} \bar{F}_k(x) \bullet \bar{r} | 2, p_{knm} \rangle \right|^2 \delta(E_2 - E_1 - \hbar\omega_{knm})$$

$$= \frac{2\pi e^2 N_k^2}{\hbar L^2} \left| \langle 1 | \bar{F}_k\,(x) \bullet \bar{r} | 2 \rangle \right|^2 (p_{knm} + 1)\delta(E_2 - E_1 - \hbar\omega_{knm}) \tag{6.10}$$

where we have used:

$$\langle p_{knm} + 1 | a_{knm}^{+} | p_{knm} \rangle = \sqrt{p_{knm} + 1} \tag{6.11}$$

From (6.10) we obtain the spontaneous emission rate:

$$W_{knm} = \frac{2\pi e^2 N_k^2}{\hbar L^2}\left|\left\langle 1\right|\vec{F}_k\,(x)\bullet\vec{r}\left|2\right\rangle\right|^2 \delta(E_2 - E_1 - \hbar\omega_{knm}) \qquad \textbf{(6.12)}$$

by dropping the term proportional to p_{knm} which describes stimulated emission. From now on, we use W_{knm} equal to (6.12) and understand that it refers to spontaneous emission only.

Equations (6.7) and (8) define the rate of spontaneous emission into the *knm*th optical mode; the total rate is obtained by summing over all possible modes. The total rate of spontaneous emission into radiation modes is given by:

$$W_{rad} = \sum_{knm_{rad}} W_{knm}\,; \qquad \textbf{(6.13)}$$

and the total emission rate into bound modes is:

$$W_{bound} = \sum_{knm_{bound}} W_{knm} \qquad \textbf{(6.14)}$$

The total rate of spontaneous emission is then:

$$W = W_{rad} + W_{bound} \qquad \textbf{(6.15)}$$

We first derive W_{rad}. Using (6.12), (A.3) and:

$$\delta(E_2 - E_1 - \hbar\omega_{knm}) = \frac{1}{\hbar}\delta(\omega_o - \omega_{knm}) \qquad \textbf{(6.16)}$$

where:

$$\omega_o \equiv \frac{E_2 - E_1}{\hbar} \qquad \textbf{(6.17)}$$

(6.13) becomes:

$$W_{rad} = \frac{\pi e^2}{\hbar n_2^2 \varepsilon_o L^3} \sum_{knm_{rad}} \left| \left\langle 1 \right| \vec{F}_k(x) \bullet \vec{r} \left| 2 \right\rangle \right|^2 \omega_{knm} \delta(\omega_o - \omega_{knm})$$

$$= \frac{\pi e^2}{\hbar n_2^2 \varepsilon_o L^3} \sum_{\bar{P}} \sum_{Q_{rad}} \sum_{\varphi} \sum_{\beta} \left| \left\langle 1 \right| \vec{F}_{\bar{P}}(\omega, \beta, Q, x) \bullet \vec{r} \left| 2 \right\rangle \right|^2 \omega \delta(\omega_o - \omega) \qquad \textbf{(6.18)}$$

where $\bar{P}$=TE (even), TE (odd), TM (even) and TM (odd), and we have simply rewritten $\omega_{knm} \rightarrow \omega$. The n and m components of the sum are replaced by β and φ (Figure 2.a), while the k component is replaced by Q, defined as:

$$Q^2 = \frac{n_2^2 \omega^2}{c_o^2} - \beta^2 \qquad \textbf{(6.19)}$$

In accordance with our new sum parameters we have rewritten $\vec{F}_k(x) \rightarrow \vec{F}_{\bar{P}}(\omega, Q, \beta, x)$. Now, considering $0 \le \varphi \le 2\pi$ and $\beta \ge 0$, we obtain by the standard limiting procedure for large L:

$$\frac{1}{L^3} \sum_{Q_{rad}} \sum_{\varphi} \sum_{\beta} \rightarrow \frac{1}{4\pi^3} \int_{Q=0}^{Q_{max}} \int_{\varphi=0}^{2\pi} \int_{\beta=0}^{\infty} \beta d\beta d\varphi dQ \qquad \textbf{(6.20)}$$

where:

$$Q_{max} = n_2 \omega / c_o \qquad \textbf{(6.21)}$$

is the limiting value of Q for radiating modes. Performing a coordinate transformation using (6.19), and neglecting material dispersion, we can rewrite (6.20) as:

$$\frac{1}{L^3} \sum_{Q_{rad}} \sum_{\varphi} \sum_{\beta} \rightarrow \frac{n_2^2}{4\pi^3 c_o^2} \int_{Q=0}^{Q_{max}} \int_{\varphi=0}^{2\pi} \int_{\omega=Qc_o/n_2}^{\infty} \omega d\omega d\varphi dQ \qquad \textbf{(6.22)}$$

Therefore, (6.18) becomes:

$$W_{rad} = \frac{e^2\omega_o^2}{4\pi^2\varepsilon_o\hbar c_o^2}\sum_p \int_{Q=0}^{Q_{max}}\int_{\varphi=0}^{2\pi}\left|\langle 1|\vec{F}_p(\omega_o,Q,x)\bullet\vec{r}|2\rangle\right|^2 d\varphi dQ \tag{6.23}$$

where now $Q_{max} = n_2\omega_o / c_o$, and we have taken $\vec{F}_p(\omega,Q,\beta,x) \rightarrow \vec{F}_p(\omega_o,Q,x)$ since fixing $\omega = \omega_o$ makes β a function of Q only.

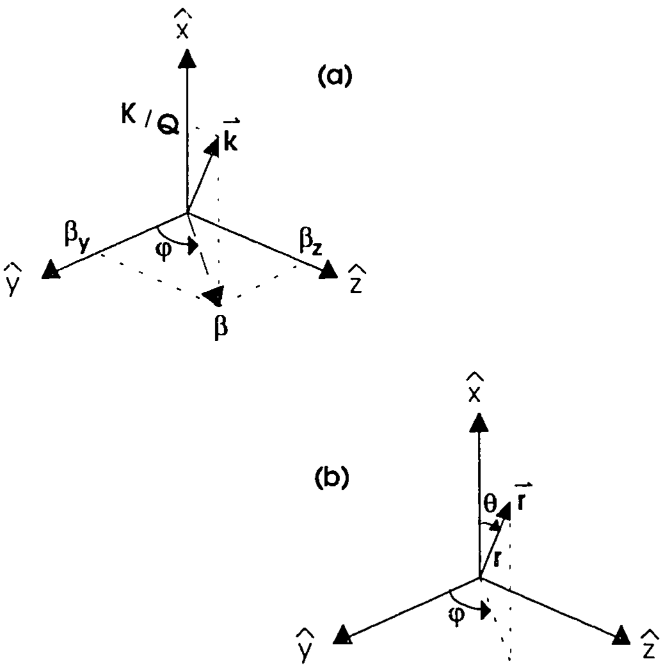


Figure 6.2.a Coordinate representation of general mode wave vector $\vec{k}$. **.b** Coordinate representation of atomic dipole moment $\vec{r}$.

The field-dipole matrix element:

$$\left|\langle 1|\vec{F}_{\vec{p}}(\omega_o,Q,x)\bullet\vec{r}|2\rangle\right|^2 \tag{6.24}$$

depends on the mode polarization (hence the subscript $\bar{P}$) and on the atomic dipole magnitude and orientation. The atomic dipole may have any orientation; see Figure 6.2.b. We write:

$$\left|\left\langle 1\left|\bar{F}_{\bar{P}}(\omega_o, Q, x)\bullet\bar{r}\right|2\right\rangle\right|^2 = r_{12}^2\left|\bar{F}_{\bar{P}}(\omega_o, Q.x_a)\bullet\hat{r}\right|^2 \qquad \textbf{(6.25)}$$

where:

$$r_{12}^2 \equiv \left|\left\langle 1|\bar{r}|2\right\rangle\right|^2 \qquad \textbf{(6.26)}$$

and x_a is the atomic position: the field magnitude is approximately constant over the dimensions of an atom. We may then rewrite (6.23) as:

$$W_{rad} = \frac{e^2\omega_o^2 r_{12}^2}{4\pi^2\varepsilon_o\hbar c_o^2}\sum_{\bar{P}}\int_{Q=0}^{Q_{max}}\int_{\varphi=0}^{2\pi}\left|\bar{F}_{\bar{P}}(\omega_o, Q, x_a)\bullet\hat{r}\right|^2 d\varphi dQ \qquad \textbf{(6.27)}$$

The integral over Q must be computed numerically, whereas that over φ can be analytically evaluated for a given $\bar{P}$ and $\hat{r}$. For example, for TE modes where $\hat{r}$ is arbitrarily oriented (say $\hat{r} = \hat{y}$) in the $\hat{y} - \hat{z}$ plane:

$$\int_{\varphi=0}^{2\pi}\left|\bar{F}_{\bar{P}=TE_{e,o}}(\omega_o, Q, x_a)\bullet\hat{r}\right|^2 d\varphi = \left|\bar{F}_{\bar{P}=TE_{e,o}}(\omega_o, Q, x_a)\right|^2\int_{\varphi=0}^{2\pi}\cos^2\varphi d\varphi = F_{\bar{P}=TE_{e,o}}^2(\omega_o, Q, x_a)\pi$$

$$\textbf{(6.28)}$$

Horizontal ($\hat{y} - \hat{z}$ plane) dipoles can couple to both TE and TM modes, as both have electric field components in the plane, whereas vertical dipoles (parallel to $\hat{x}$) can couple only to TM modes. Both components of $\bar{F}$ must be considered in evaluating (6.27) for the cases $\bar{P}$=TM (even) and TM (odd). For example, for TM modes and vertical dipoles, (6.27) becomes:

$$W_{rad} = \frac{e^2\omega_o^2 r_{12}^2}{4\pi^2\varepsilon_o \hbar c_o^2} \sum_{P=TM_{e,o}} \int_{Q=0}^{Q_{max}} \int_{\varphi=0}^{2\pi} \left| \left(\bar{F}_P(\omega_o, Q, x_a) \bullet \hat{x} \right) \right|^2 d\varphi dQ$$

$$= \frac{e^2\omega_o^2 r_{12}^2}{4\pi^2\varepsilon_o \hbar c_o^2} \sum_{P=TM_{e,o}} \int_{Q=0}^{Q_{max}} \int_{\varphi=0}^{2\pi} F_x^2(\omega_o, Q, x_a) d\varphi dQ$$

$$= \frac{e^2\omega_o^2 r_{12}^2}{2\pi\varepsilon_o \hbar c_o^2} \sum_{\bar{P}=TM_{e,o}} \int_{Q=0}^{Q_{max}} F_x^2(\omega_o, Q, x_a) dQ \qquad \textbf{(6.29)}$$

where F_x is the $\hat{x}$-component of $\bar{F}_{\bar{P}=TM_{e,o}}$ in (A.1.c,d).

We evaluate the emission rate for bound modes using equations (6.12) and (6.14):

$$W_{bound} = \frac{2\pi e^2}{\hbar L^2} \sum_{knm_{bound}} N_k^2 \left| \left\langle 1 \right| \bar{F}_k(x) \bullet \bar{r} \left| 2 \right\rangle \right|^2 \delta(E_2 - E_1 - \hbar\omega_{knm}) \qquad \textbf{(6.30)}$$

Using (6.16) and (6.17), we can rewrite (6.30) as:

$$W_{bound} = \frac{2\pi e^2}{\hbar L^2} \sum_{\bar{P}} \sum_{K} \sum_{\varphi} \sum_{\beta} N_{\bar{P}}^2(\omega, K, \beta) \left| \left\langle 1 \right| \bar{F}_{\bar{P}}(\omega, K, \beta, x) \bullet \bar{r} \left| 2 \right\rangle \right|^2 \delta(\omega_o - \omega) \textbf{(6.31)}$$

where again we have rewritten $\omega_{knm} \rightarrow \omega$. The new lateral propagation constant parameter, K, is given by:

$$K^2 = \frac{n_1^2\omega^2}{c_o^2} - \beta^2 \qquad \textbf{(6.32)}$$

Bound modes exist for certain discrete values of K in the range $0 \leq K \leq q$, where:

$$q^2 \equiv K^2 - Q^2 = \left(n_1^2 - n_2^2\right)\omega^2/c_o^2 \qquad \textbf{(6.33)}$$

The allowed values of K are the solutions of equations (A.5), and the normalization constants of the bound modes, $N_{\bar{P}}^2(\omega, K, \beta)$, are given by equations (A.6). Taking the limit of large L, we can write:

$$\frac{1}{L^2}\sum_{\varphi}\sum_{\beta} \rightarrow \frac{1}{4\pi^2}\int_{\varphi=0}^{2\pi}\int_{\beta=0}^{\infty}\beta d\beta d\varphi = \frac{1}{4\pi^2}\int_{\varphi=0}^{2\pi}\int_{\omega=0}^{\infty}\left(\frac{\beta}{v_g^K}\right)d\omega d\varphi \qquad \textbf{(6.34)}$$

where v_g^K is the group velocity of mode K. Therefore, from (6.31) and using our previous arguments regarding the field-dipole matrix element, we finally obtain:

$$W_{bound} = \frac{e^2 r_{12}^2}{2\pi\hbar^2}\sum_{\bar{P}}\sum_{K} N_{\bar{P}}^2(\omega_o, K)\left(\frac{\beta}{v_g^K}\right)\int_{\varphi=0}^{2\pi}\left|\bar{F}_{\bar{P}}(\omega_o, K, x_a)\bullet\hat{r}\right|^2 d\varphi \qquad \textbf{(6.35)}$$

where we have taken $N_{\bar{P}}^2(\omega, K, \beta) \rightarrow N_{\bar{P}}^2(\omega_o, K)$ and $\bar{F}_{\bar{P}}(\omega, K, \beta, x) \rightarrow \bar{F}_{P}(\omega_o, K, x)$. It can be shown that, in the absence of material dispersion, the factor β/v_g^K is given by:

$$\frac{\beta}{v_g^K} = \frac{\omega_o}{c_o^2}\left\{\Gamma_K n_1^2 + (1-\Gamma_K)n_2^2\right\} \qquad \textbf{(6.36)}$$

where Γ_K is the mode power confinement in the core layer, given by (A.7) in the Appendix. Note that:

$$\beta/v_g^K \approx n_1^2\omega_o/c_o^2 \qquad \textbf{(6.37)}$$

is a good approximation except for modes close to cut-off. The integration over φ may again be performed analytically for each $\bar{P}$ and $\hat{r}$.

Equations (6.15), (6.27) and (6.35) together yield the total rate of spontaneous emission from a dipole emitter located at an arbitrary position relative to the waveguide. Furthermore, the emission rates into radiating and guided modes are given separately by (6.27) and (6.35). The rate of emission into any *particular* guided mode is directly calculable using (6.35) for a single choice of $\overline{P}$ and K.

6.2 Example Calculations In Figures 6.3-6.8 we show calculated spontaneous emission rates for various slab waveguides as a function of the $\hat{x}$-position of the emitting atom. In all cases, the guide refractive index n_1=1.50, and the transition wavelength $\lambda_o=2\pi c_o/\omega_o$=1.55 μm. The spontaneous emission rates are normalized with respect to the total rate expected in a homogeneous medium of refractive index n_2, given by equation (8.3-9) in [8].

In Figure 6.3, we show the total spontaneous emission rate, obtained by numerical solution of equations (6.15), (6.27), (6.35) and (6.36), using (A.1)-(A.7), for n_2=1.47 and $a=2$ μm. This waveguide supports one bound TE mode and one bound TM mode. The x-axis in the plot is the position of the radiating atom in multiples of a, so that the guide/cladding boundary is located at 0.5. The solid line shows the emission rate for horizontal atomic dipoles, the dashed line shows that for vertical dipoles, while the dotted line depicts the rate for randomly-oriented dipoles. For atoms located far into the cladding, the normalized emission rates approach one, as expected since these atoms are essentially unaffected by the distant index variation. The rates in the guide are higher than those in the cladding by roughly a factor of n_1/n_2; this is also expected since the emission rate in a homogeneous medium is linear with respect to index of refraction. Note, however, the marked difference in the emission rate characteristics for horizontal and vertical dipoles in the region of the guide/cladding interface. For horizontal dipoles, the emission rate is a continuous function of position, while that for vertical dipoles is discontinuous at the index interface. This, of course, is due to the fact that the horizontal dipoles couple to mode electric fields which lie in the waveguide plane and which are therefore continuous with respect to x, while the vertical dipoles couple to the $\hat{x}$-components of the TM mode electric fields, which are discontinuous at the index step. The discontinuity in the emission rate of vertical dipoles dominates the characteristic for randomly-oriented dipoles, leading to the result that the rate of spontaneous emission is highest for atoms located immediately outside the waveguide core, and lowest for atoms immediately inside the guide/cladding boundary. In fact, the rate of emission there is actually lower than that deep in the cladding.

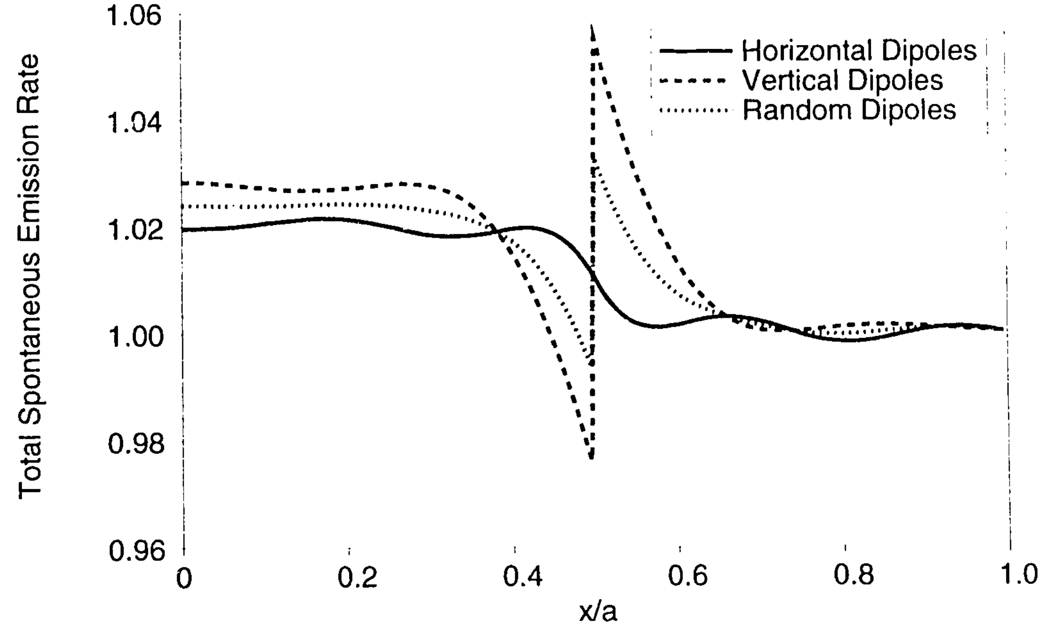


Figure 6.3 Total spontaneous emission rates as a function of $\hat{x}$-position of emitting atom (/*a*) for a symmetric slab waveguide with n_1=1.50, n_2=1.47, a=2 μm, normalized with respect to the rate of emission in a homogeneous medium of index n_2. The waveguide centre is at x=0; therefore **the guide/cladding interface is located at 0.5**. The emission wavelength is 1.55 μm, and the structure is monomode. Solid line: emission rate from horizontal dipoles; dashed line: emission rate from vertical dipoles; dotted line: emission rate from randomly-oriented dipoles.

This behaviour persists as we increase the width of the guide layer; see Figure 6.4. Here, we show the total spontaneous emission rate from randomly-oriented dipoles for guide widths of $a = 2$ μm (one TE, TM mode), 6 μm (three TE, TM modes) and 24 μm (ten TE, TM modes); again, n_2=1.47. As the waveguide becomes wider, the total emission rate at the center of the guide more closely approaches that expected for a homogeneous region of index n_1, and the range of emission rate oscillations decreases.

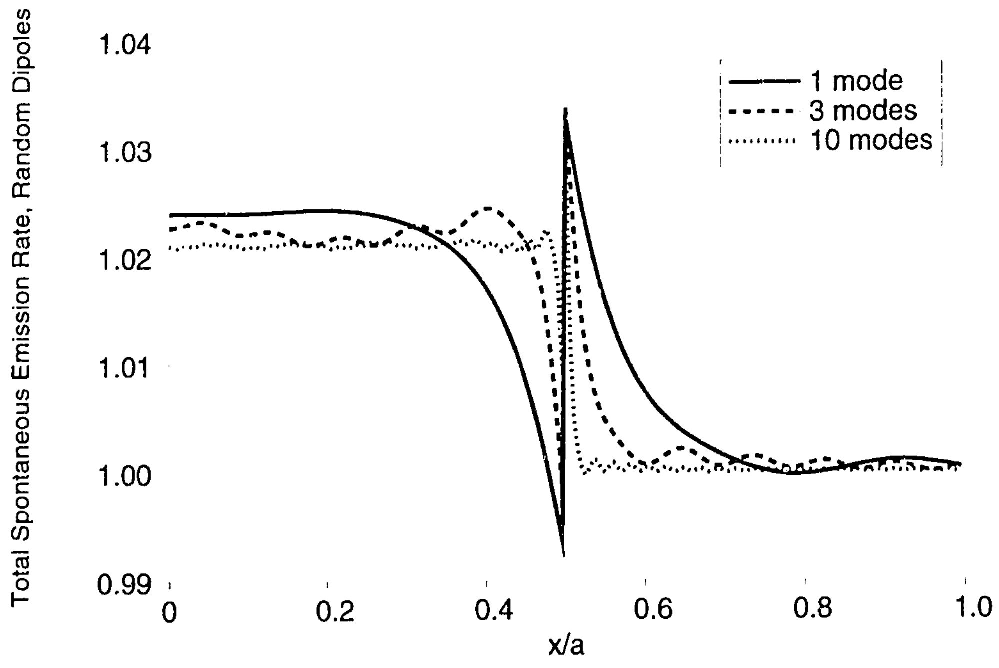


Figure 6.4 Total emission rates from randomly-oriented dipoles for n_1=1.50 and n_2=1.47; a=2 μm (solid line, one TE and TM mode); a=6 μm (dashed line, three modes); a=24 μm (dotted line, ten modes).

In Figures 6.5 and 6.6, we show the rates of emission into the bound and radiating modes, respectively, of a 2 μm-wide guide where n_2=1.47. The discontinuity in the E_x field of the TM mode is apparent in the rate curve for vertical dipoles. Note that the rate of emission into bound modes for vertical dipoles is higher than that for horizontal dipoles, as expected from classical theory.

In Figure 6.7 we show the total spontaneous emission rate from randomly-oriented dipoles for a guide width of 2 μm, where the cladding index is set to n_2=1.47 (solid line, one TE and TE mode), n_2=1.40 (dashed line, two TE and TM modes) and n_2=1.30 (dotted line, two TE and TM modes). As the cladding index is reduced, the normalized emission rate in the guide increases (since the rate in the

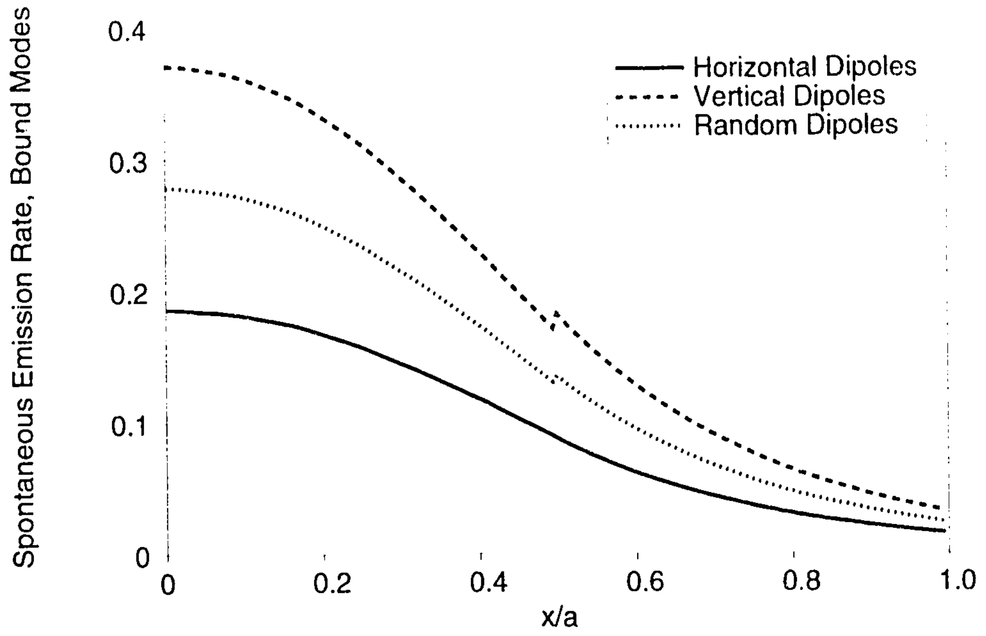


Figure 6.5 Rates of spontaneous emission into bound modes for $n_1=1.50$ and $n_2=1.47$; $a=2$ μm; horizontal dipoles (solid line); vertical dipoles (dashed line); randomly-oriented dipoles (dotted line).

cladding is reduced), and the magnitude of the swing in emission rate about the index interface increases. The rate of emission into bound modes from randomly-oriented dipoles for the same cases is shown in Figure 6.8. As the cladding index is reduced, and the numerical aperture of the waveguide is increased, a greater proportion of the emitted photons are coupled into the bound modes, as we would expect. Note, however, that as the guide/cladding index difference rises, the probability of bound mode emission capture from atoms immediately outside the waveguide increases with respect to that from atoms just inside the interface.

6.3 Calculation of Spontaneous Emission Capture by Guided Modes in Planar Rare Earth-Doped Waveguides Integrated waveguides doped with rare earths, especially erbium, are of great current interest for application as optical amplifiers and lasers. These devices use the guided modes at a pump wavelength to excite imbedded

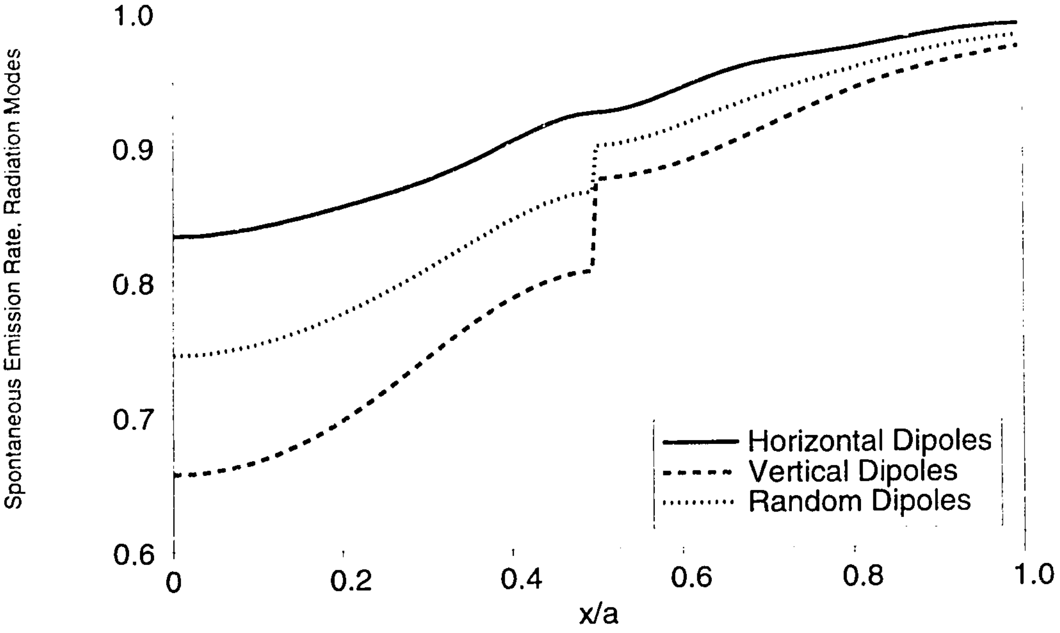


Figure 6.6 Rates of spontaneous emission into radiating modes for n_1=1.50 and n_2=1.47; a=2 μm; horizontal dipoles (solid line); vertical dipoles (dashed line); randomly-oriented dipoles (dotted line).

rare earth ions, which then contribute, via stimulated emission, to enhancement of the guided mode at a longer, signal wavelength. The spontaneous emission factor - defined as the spontaneously-emitted power coupled into guided modes, or a particular guided mode, as a fraction of the total spontaneous emission power - is an important performance parameter of such devices, since amplifier linear gain (i.e. in the low input signal power regime) is reduced with increasing spontaneous emission factor, due to the effect of amplified spontaneous emission (ASE) [116]. The spontaneous emission factor, which we label α, is given by:

$$\alpha = \frac{\int_0^\infty W_{MOl}(x)\zeta(x)dx}{\int_0^\infty \{W_{rad}(x) + W_{bound}(x)\}\zeta(x)dx} \qquad \textbf{(6.38)}$$

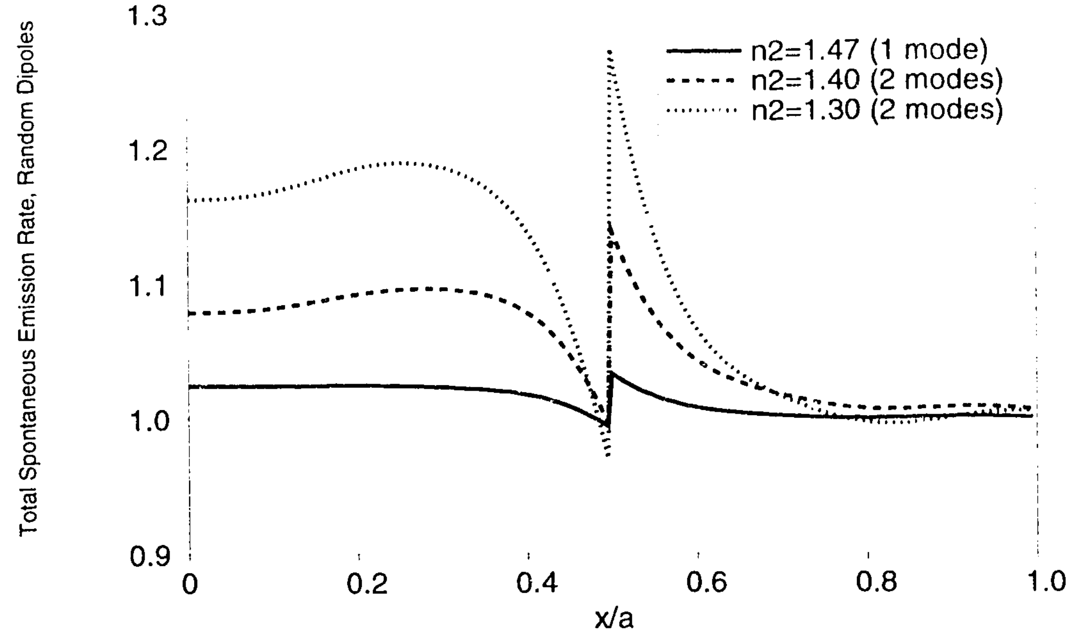


Figure 6.7 Total rates of spontaneous emission for n_1=1.50 and a=2 μm; n_2=1.47 (solid line, one TE and TM mode); n_2=1.40 (dashed line, two modes); n_2=1.30 (dotted line, two modes).

where W_{rad} and W_{bound} are given by equations (6.27) and (6.35), respectively, averaged over the orthogonal dipole orientations. The subscript MOI in the numerator stands for mode(s) of interest, and W_{MOI} is given by evaluation of (6.35) for a specific bound mode or modes, again averaged with respect to dipole orientation. The function $\zeta(x)$ is a distribution function for dipole emitters. In semiconductor lasers and amplifiers, a function equivalent to $\zeta(x)$ is derived from the distribution of injected carriers; here, the rare-earth doping profile and the guided modes at the pump wavelength together determine $\zeta(x)$. In what follows, we consider the common case of Er^{3+}, pumped at 980 nm wavelength, for emission at (approximately) 1550 nm. Two different doping situations are examined: uniform doping of the guide (core), and uniform doping of the cladding. The distribution function $\zeta(x)$ is taken as:

$$\zeta(x) = \sum_{v_{even}} \frac{1}{N_{even}} \chi(x) \left|\psi_v(x)\right|^2 \qquad (6.39)$$

where the summation is over the *even* modes (TE and TM) of the pump light: N_{even} is the number of such modes, $\psi_\nu(x)$ are the intensity-normalized pump mode profiles, and $\chi(x)$ is a function representing the erbium doping profile, e.g.:

$$\chi(x) = 1 \quad |x| \le a/2, \quad 0 \quad |x| > a/2 \tag{6.40}$$

for the case of doping in the core.

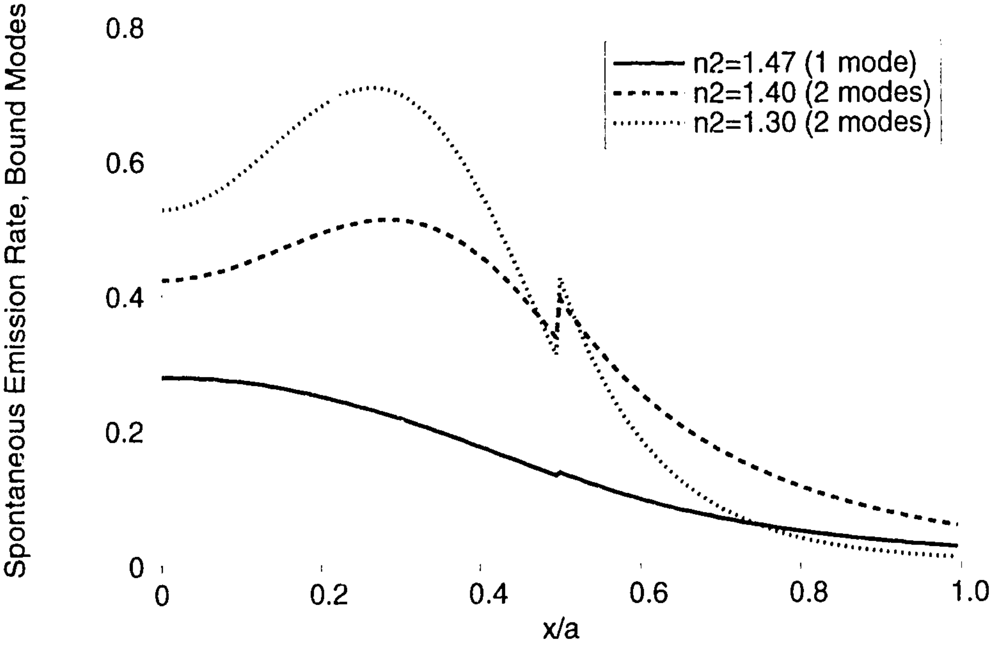


Figure 6.8 Rates of spontaneous emission into bound modes for n_1=1.50 and a=2 μm; n_2=1.47 (solid line, one TE and TM mode); n_2=1.40 (dashed line, two modes); n_2=1.30 (dotted line, two modes).

Uniform weighting of the even pump modes approximates the situation of a narrow pump source centered on the guide. Note that if *both* the guide and the cladding were doped, i.e.:

$$\chi(x) = 1, \quad 0 \le |x| < \infty \tag{6.41}$$

we would then have:

$$\int_{-\infty}^{\infty} \zeta(x)dx = 1 \tag{6.42}$$

In Figures 6.9-6.12, we consider a case where the cladding index n_2=1.40 and the guide width $a = 3$ µm; the guide index n_l is varied from 1.42 to 1.52. To reduce the number of features in the plots and make the effects of the appearance of new modes, at both 0.98 µm and 1.55 µm, with increasing n_l more clear, we average the cut-off point of TE and TM modes so that they appear simultaneously. We first consider the conventional case where the core layer is erbium-doped. In Figure 6.9, we calculate the spontaneous emission factor where the modes of interest for the two cases shown are, respectively, the TE and TM fundamental modes at 1.55 µm wavelength. The spontaneous emission factor is higher for the TM mode than for the TE mode, by a factor of almost exactly two in all cases. This is an artifact of choosing randomly-oriented atomic dipoles, since vertical dipoles, which couple only to TM fields, emit radiation which is more efficiently captured by the bound modes. A similar effect on stimulated emission and gain is indicated according to equation (6.10). In planar amplifier geometries, such polarization-dependent gain has been seen experimentally [117]; the magnitude of the observed gain dependence (up to 0.8 dB) is more than can be readily explained by the difference in confinement factors between TE and TM modes. This polarization effect is not observed in fibers as both quasi-polarizations have two electric field components. We note that α is roughly constant with respect to guide index; more precisely, α for the fundamental mode increases with guide index, but decreases suddenly at certain points as the number of bound modes increases. The large drop at n_l=1.438 corresponds to the appearance of a second even mode at 980 nm. Since two pump modes are now available, the erbium in the guide is more uniformly excited, and since two 1.55 µm modes are already sustained at this point, spontaneous emission coupling to the first order mode increases at the expense of coupling to the fundamental mode. This is confirmed in Figure 6.10, where now the "modes of interest" defining the spontaneous emission factor include *all* of the guided modes. A much smaller drop is observed here,

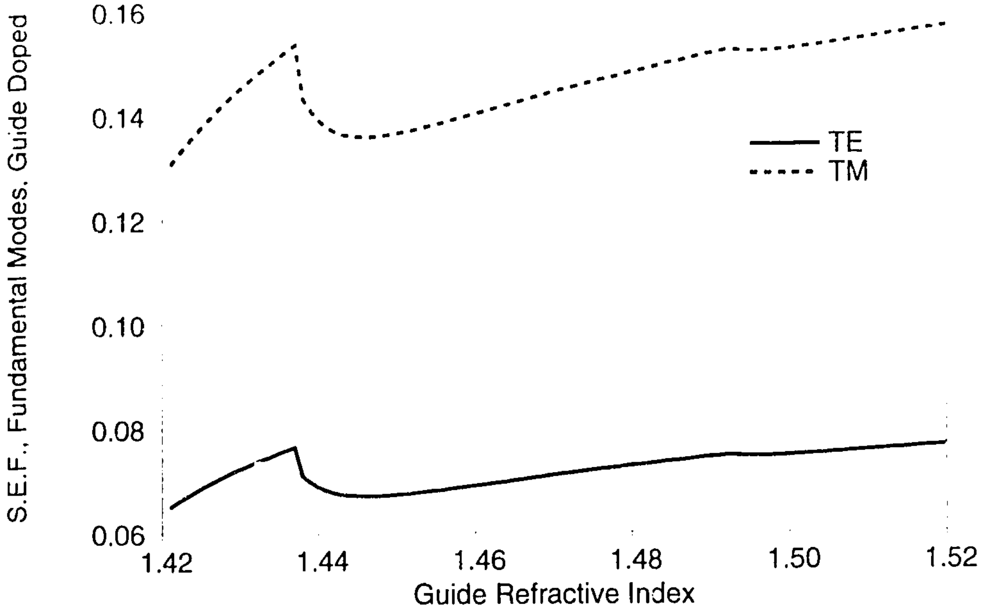


Figure 6.9 Spontaneous emission factors for TE and TM fundamental 1.55 μm modes of an erbium-doped planar amplifier pumped with 0.98 μm guided light, as a function of guide index. The guide is 3 μm wide and the cladding index n_2=1.40. The core layer is erbium-doped.

indicating that coupling to the first order mode at 1.55 μm has, in fact, increased. Going back to Figure 6.9, the smaller drop at n_l=1.493 corresponds to the onset of a second order mode at 1.55 μm. This is seen in Figure 6.10 as a large increase in α when all the guided modes are considered. In general, total captured spontaneous emission increases whenever new guided modes appear at 1.55 μm; the slight increases in the characteristics in Figure 6.10 at n_l=1.424 correspond to the appearance of the first order mode at 1.55 μm.

When the erbium is in the cladding rather than the core, the situation changes dramatically. In Figures 6.11 and 6.12 we show the TE and TM spontaneous emission factors, defined for the fundamental modes in Figure 6.11, and for all of the guided modes in Figure 6.12. The values of α are now greatly reduced; this is attributable to the reduced magnitude of the guided modes in the cladding, and the increased coupling to radiation modes. In Figure 6.11 we see that α for the

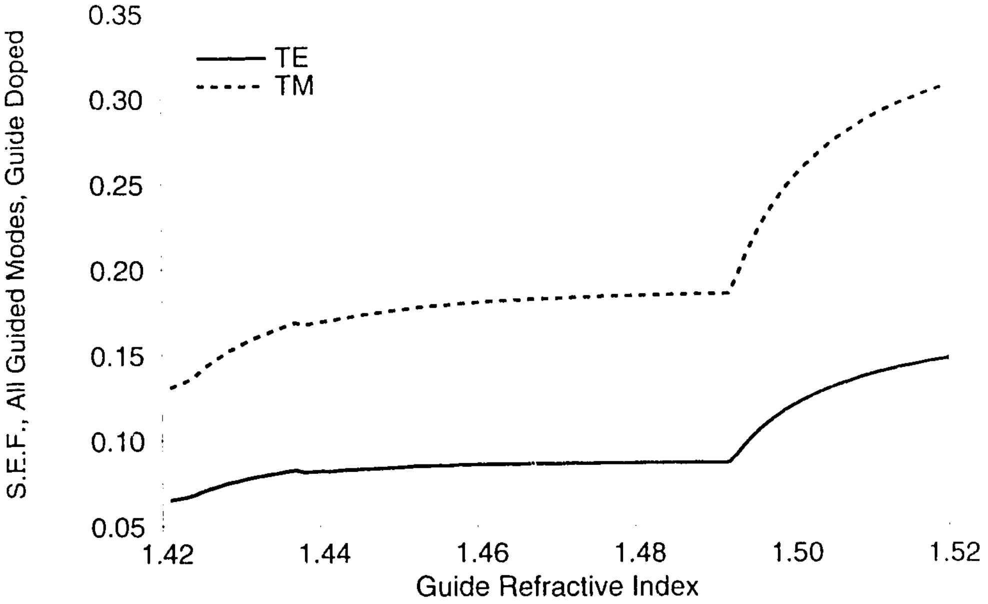


Figure 6.10 Spontaneous emission factors versus guide index for all guided TE and TM modes; the guide is 3 μm wide and the cladding index n_2=1.40. The core is erbium-doped.

fundamental modes are monotonically decreasing with increasing guide index, and are less than 0.1% in all cases. In Figure 6.12, we note that a similar characteristic is obtained with the appearance of every new mode at 1.55 μm, and that the resulting peak in α increases in magnitude with every new mode. This is attributable to the greater penetration into the cladding of the higher order modes. Thus, when the cladding is doped, significant spontaneous emission coupling to the guided modes occurs only near the cut-off of a high-order mode, while no significant coupling to the fundamental modes occurs at all. It is interesting that the ratio of α for TM modes versus that for TE is now greater than two, due to the discontinuous increase in the TM fields at the guide-cladding interface. It is also interesting that no effects due to the onset of new 980 nm modes are apparent in Figures 6.11 and 6.12; this implies that when these modes appear they contribute equal relative increases in spontaneous emission into the bound modes and into the radiation modes, since:

$$\int_0^\infty W_{rad}(x)\zeta(x)dx >> \int_0^\infty W_{bound}(x)\zeta(x)dx \tag{6.43}$$

when the cladding is doped. Note that while the spontaneous emission coupling to bound modes is greatly reduced when the cladding is doped, rather than the core, the gain in the fundamental modes is concomitantly degraded due to the reduced overlap of the signal and pump modes with the doped region. Therefore, the reduced gain versus pump power efficiency effectively prohibits this configuration for effective amplifier operation.

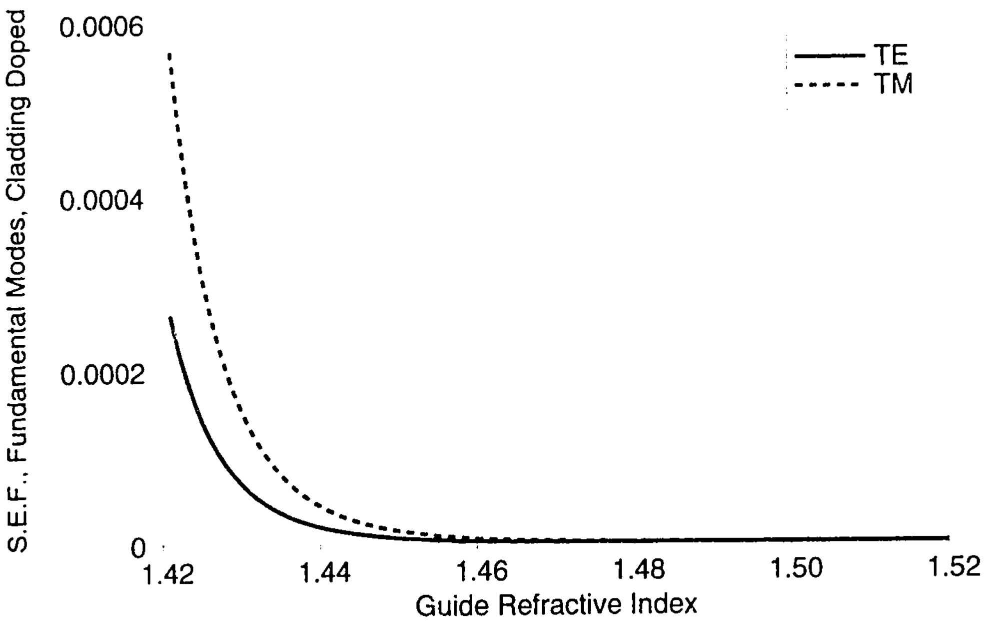


Figure 6.11 Spontaneous emission factors of fundamental modes versus guide index when the cladding is erbium-doped; the guide is 3 μm wide and the cladding index n_2=1.40.

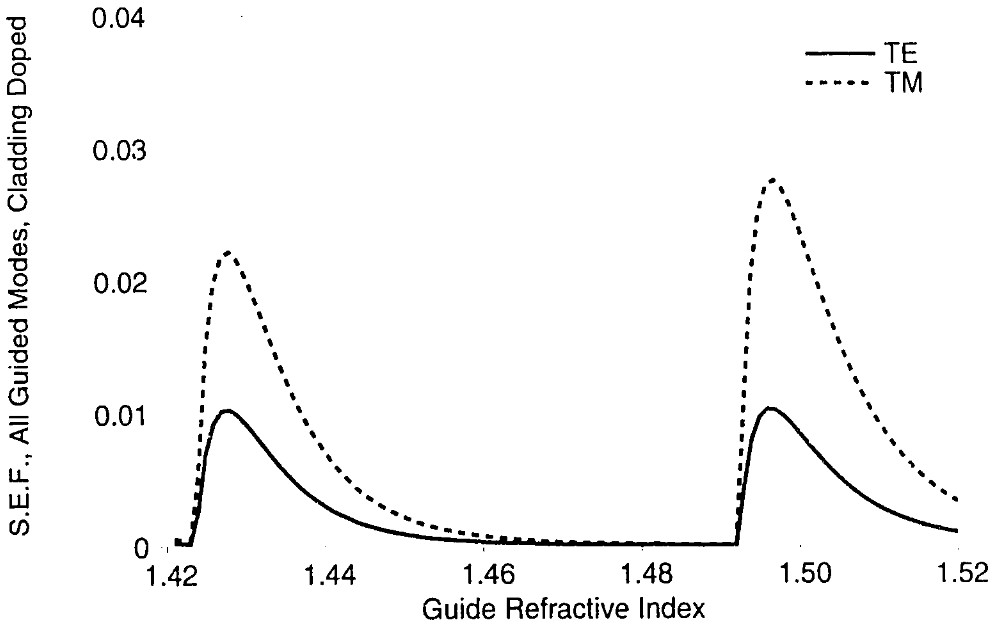


Figure 6.12 Spontaneous emission factors of all guided modes versus guide index when the cladding is erbium-doped; the guide is 3 μm wide and the cladding index $n_2=1.40$.

6.4 Conclusions In this chapter, we have presented an analysis of spontaneous emission in the vicinity of a symmetric slab waveguide. The analysis uses explicit functions for the complete set of bound and radiating TE and TM modes, and hence the rate of emission into specific modes from an arbitrarily-located emitter is easily found. We have used the model to calculate spontaneous emission factors in an optical-guided-wave-pumped planar erbium-doped waveguide; the planar geometry is sufficient to approximate the behaviour of at least weakly-guiding three-dimensional structures. We found that when the guide layer is uniformly doped, significant coupling to the bound modes exists, and the spontaneous emission factor for the fundamental modes is relatively constant as a function of n_1/n_2, although the spontaneous emission factor is higher for TM modes than for TE when randomly-oriented atomic dipoles are assumed. When the cladding layers are doped, significant coupling to bound modes exists only for values of n_1/n_2 near the cut-off condition of high-order modes, and no significant coupling to the fundamental modes can be obtained.

Our model may be extended to more complicated, multi-layer planar devices by using a numerical method to find the guided and radiating modes; a transfer matrix approach is suitable [118]. Equation (6.27) for the radiation modes remains valid since these modes are always identically normalized, though n_2 in (6.19) and (A.3) must be replaced with the index of the semi-infinite cap and substrate layers (if these differ, then (6.27) must of course be decomposed to deal separately with the substrate modes and the true radiation modes.) Likewise, equation (6.35) for the bound modes remains valid if (6.36) is properly modified. The normalization of the bound modes would have to be done numerically, but this would not present a problem since these modes are of finite spatial extent. Such a model would be similar in many ways to that presented in [113], except that the rate of emission into guided modes would be more readily calculable. This could be of some use in examining novel devices such as reported in [119].

Chapter 7

Conclusions

The work of this thesis has been concerned with two aspects of photonic integrated circuits: compactness, and multi-functionality. Compactness will be an important concern in the future as increasingly complex circuit implementation is attempted and higher device densities are required. Deeply-etched rib waveguides will probably be required to realize such complex circuits, since they can be sharply bent without suffering radiation losses and hence are uniquely suited to space-efficient optical distribution. We have considered the implementation of multi-mode couplers in such waveguides. Multi-mode couplers are an important component of many integrated guided-wave devices, such as the Mach-Zehnder switch designed and fabricated in this work. We showed that extremely compact (<100 μm) high-performance couplers can be realized by a proper optimization of the access waveguide geometry and the coupler width. A Mach-Zehnder switch in InP which used such small multi-mode couplers demonstrated high contrast and low on-chip insertion loss, and was of a record small size. We also showed that bending a multi-mode waveguide can influence the guided mode propagation characteristics in a useful way, and that the radius of a circular bend can therefore be used as a new design parameter for some structures. We used this concept to design and fabricate a multi-mode coupler for (de)multiplexing 850 nm and 1550 nm light. Impressive performance was achieved using a very small device (<150 μm). This is the first device yet reported which uses a circularly bent multi-mode waveguide to useful advantage.

Multi-functionality is another important aspect of future monolithically integrated photonic circuits. We have designed the first semiconductor structure which is capable of laser/modulator/detector/amplifier functions, such that the local functionality is flexible and determined only by electrical biasing, and which does not require semiconductor regrowth or patterned epitaxy. Though other authors have reported integrated devices with superior performance, we have nevertheless demonstrated that only a minimum of extraordinary fabrication measures need be employed to realize monolithic optoelectronic multi-functionality.

Finally, we have performed a theoretical evaluation of spontaneous emission processes in the vicinity of a symmetric slab waveguide. This work demonstrates the inherent polarization-sensitivity of emission processes in non-uniform dielectric spaces lacking radial symmetry, and provides a starting point for rigorous investigations of gain and spontaneous emission in integrated, semi-planar rare-earth doped waveguide geometries.

Optoelectronic integrated circuits are only now reaching the stage of viable commercial implementation, and so far have been largely restricted to optical transmitter and receiver modules. As telecommunications networks migrate toward the SONET standard, and as optical fiber links continue to proliferate down to lower levels of the networks, new integrated optoelectronic circuit components will begin to be deployed. The most important such class of circuits in the forseeable future will be optical time division and wavelength division multiplexing components, which will require the hybrid and/or monolithic integration of a wide variety of active and passive devices. Semiconductor lasers, of which quantum well lasers are now the most important class, have reached a relatively mature state, as have modulators and detectors, although methods of integrating these two classes of devices on one semiconductor substrate are being actively investigated. One particularly important research effort in the field of active devices is the search for polarization-insensitive operation. Strained quantum wells have emerged as the leading candidate for high-performance polarization insensitive modulators and detectors [120], although other interesting means of fabricating polarization-insensitive quantum wells have been suggested [121].

Ironically, passive optical components, the backbone of all integrated photonic systems, have recently been the center of renewed focus. Multi-mode waveguide couplers have been truly understood only in the last few years, and this understanding has lead not only to record small size and high-performance components, but also to new device concepts, such as single-stage *NXN* Mach-Zehnder switches. Integrated components for both coarse and fine wavelength division multiplexing components are of increasing interest, and research to this end has examined a wide variety of device types, including multi-mode and diffractive elements [122]. Material systems for useful passive component fabrication, and cost-effective means of waveguide circuit fabrication, are also the subjects of considerable research. Organic and polymer materials for direct waveguide writing are being explored in many laboratories [123]. Some such materials also hold the promise of thermal or electro-optic modulation capabilities. These efforts are sure to receive increased attention as the demand for integrated photonic components in telecommunications networks rises.

References

[1] L. Mollenauer, E. Lichtman, M. Neubelt & G. Harvey, *"Demonstration, using sliding-frequency guiding filters, of error-free soliton transmission over more than 20 Mm at 10 Gbit/s, single channel, and over more than 13 Mm at 20 Gbit/s in a two-channel WDM,"* Electron. Lett., vol. 29, 1993, pp. 910 - 911

[2] T. Widdowson & A. Ellis, *"20 Gbit/s soliton transmission over 125 Mm,"* Electron. Lett., vol. 30, 1994, pp. 1866 - 1868

[3] J. Zucker, K. Jones, B. Miller, M. Young, U. Koren, B. Tell & K. Brown-Goebeler, *"Interferometric quantum well modulators with gain,"* IEEE J. Lightwave Technol., vol. 10, 1992, pp. 924 - 932

[4] J. Verdiell, T. Koch, B. Miller, M. Young, U. Koren, F. Storz & K. Brown-Goebeler, *"A WDM receiver photonic integrated circuit with net on-chip gain,"* IEEE Photon. Technol. Lett., vol. 6, 1994, pp. 960 - 962

[5] U. Hilbk, T. Hermes, P. Meissner, F. Westphal, G. Jacumeit, R. Stenzel & G. Unterborsch, *"First system experiments with a monolithically integrated tunable polarization diversity heterodyne receiver OEIC on InP,"* IEEE Photon. Technol. Lett., vol. 7, 1995, pp. 129 - 131

[6] L. Spiekman, Y. Oei, E. Metaal, F. Groen, I. Moerman & M. Smit, *"Extremely small multimode interference couplers and ultrashort bends on InP by deep etching,"* IEEE Photon. Technol. Lett., vol. 6, 1994, pp. 1008 - 1010

[7] D. Marcuse, *Theory of dielectric optical waveguides*, Academic Press, 1974

[8] A. Yariv, *Quantum electronics*, 3rd Edition, J. Wiley & Sons, 1988, pp. 600 - 650

[9] K. Ebeling, *Integrated opto-electronics*, Springer-Verlag, 1993, pp. 88 - 104

[10] B. Saleh & M. Teich, *Photonics*, J. Wiley & Sons, 1991, pp. 238 - 271

[11] M. Born & E. Wolf, Principles of optics, 6th Edition, Pergamon Press, 1980

[12] A. Hardy & M. Ben-Artzi, *"Expansion of an arbitrary field in terms of waveguide modes,"* IEE Proc.-Optoelectron., vol. 141, 1994, pp. 16 - 20

[13] P. Lorrain & D. Corson, Electromagnetic fields and waves, 2nd Edition, W. H. Freeman & Co., 1970, pp. 139 - 141 & 400 - 402

[14] R. Burden, J. Faires & A. Reynolds, Numerical analysis, 2nd Edition, PWS Publishers, 1981, pp. 476 - 489

[15] M. Jones & M. Patrick, *LANZ*: Software for solving the large sparse symmetric generalized eigenproblem (FORTRAN code), 1992, public domain

[16] M. Jones & M. Patrick, *"The use of Lanczos' method to solve the large generalized symmetric definite eigenvalue problem,"* Technical Report 89-67, Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, Virginia, U.S.A., 1989

[17] D. Marcuse, Light transmission optics, 2nd Edition, Van Nostrand Press, 1982, pp. 10

[18] M. Jones, *"Variational methods for dielectric waveguides,"* in P. Robson & P. Kendall (Eds.), *Rib waveguide theory by the spectral index method*, Research Studies Press and Wiley, 1990, pp. 35 - 49

[19] D. Marcuse, *"Solution of the vector wave equation for general dielectric waveguides by the Galerkin method,"* IEEE J. Quantum Electron., vol. 28, 1992, pp. 459 - 465

[20] B. Azizur Rahman & J. Davies, *"Finite-element analysis of optical and microwave waveguide problems,"* IEEE Trans. Microwave Theory & Techniques, vol. MIT-32, 1984, pp. 20 - 28

[21] P. Silvester & R. Ferrari, *Finite elements for electrical engineers*, 2nd Edition, Cambridge University Press, 1990, pp. 220 - 221

[22] C. Vassallo, *"Improvement of finite difference methods for step-index optical waveguides,"* IEE Proc.-Optoelectron., vol. 139, 1992, pp. 137 - 142

[23] P.-L. Liu & B.-J. Li, *"Semi-vectorial beam-propagation method for analyzing polarized modes of rib waveguides,"* IEEE J. Quantum Electron., vol. 28, 1992, pp. 778 - 782

[24] T. Benson, P. Kendall, M. Matin & M. Stern, *"Polarisation correction applied to scalar analysis of semiconductor rib waveguides,"* IEE Proc.-Optoelectron., vol. 139, 1992, pp. 39 - 41

[25] K. Chiang, *"Performance of the effective-index method for the analysis of dielectric waveguides,"* Opt. Lett., vol. 16, 1991, pp. 714 - 716

[26] P. Kendall, M. Adams, S. Ritchie & M. Robertson, *"Theory for calculating approximate values for the propagation constants of an optical rib waveguide by weighting the refractive indices,"* IEE Proc.-Optoelectron., vol. 134, 1987, pp. 699 - 702

[27] S. Seki, T. Yamanaka & K. Yokoyama, *"Two-dimensional analysis of optical waveguides with a nonuniform finite difference method,"* IEE Proc.-Optoelectron., vol. 138, 1991, pp. 123 - 127

[28] R. Fenner, *Finite element methods for engineers*, Unwin Bros. Ltd., 1975, pp. 37 - 70

[29] M. Koshiba, H. Saitoh, M. Eguchi & H. Hirayama, *"Simple scalar finite element approach to optical rib waveguides,"* IEE Proc.-Optoelectron., vol. 139, 1992, pp. 166 - 171

[30] W. Allegretto, Department of Mathematics, University of Alberta, Edmonton, Canada

[31] M. Feit & J. Fleck, *"Propagating beam theory of optical fiber cross coupling,"* J. Opt. Soc. Am., vol. 71, 1981, pp. 1361 - 1372

[32] D. Yevick & B. Hermansson, *"Split-step finite difference analysis of rib waveguides,"* Electron. Lett., vol. 25, 1989, pp. 461 - 462

[33] R. Accornero, M. Artiglia, G. Coppa, P. Di Vita, G. Lapenta, M. Potenza & P. Ravetto, *"Finite difference methods for the analysis of integrated optical waveguides,"* Electron. Lett., vol. 26, 1990, pp. 1959 - 1960

[34] G. Hadley, *"Wide-angle beam propagation using Pade approximant operators,"* Opt. Lett., vol. 17, 1992, pp. 1426 - 1428

[35] J. McMullin, *"Explicit split-step propagating beam method,"* IEEE Photon. Technol. Lett., vol. 6, 1994, pp. 534 - 536

[36] M. Munowitz & D. Vezzetti, *"Numerical procedures for constructing equivalent slab waveguides. An alternative approach to effective-index theory,"* IEEE J. Lightwave Technol., vol. 9, 1991, pp. 1068 - 1073

[37] C. Rolland, G. Mak, K. Fox, D. Adams, A. SpringThorpe & D. Yevick, *"Analysis of strongly guiding rib waveguide s-bends: theory and experiment,"* Electron. Lett., vol. 25, 1989, pp. 1256 - 1257

[38] O. Bryngdahl, *"Image formation using self-imaging techniques,"* J. Opt. Soc. Am., vol. 63, 1973, pp. 416 - 418

[39] R. Ulrich & G. Ankele, *"Self-imaging in homogeneous planar optical waveguides,"* Appl. Phys. Lett., vol. 27, 1975, pp. 337 - 339

[40] M. Bachmann, P. Besse & H. Melchior, *"General self-imaging properties in NXN multimode interference couplers including phase relations,"* Appl. Opt., vol. 33, 1994, pp. 3905 - 3911

[41] L. Soldano, F. Veerman, M. Smit, B. Verbeek, A. Dubost & E. Pennings, *"Planar monomode optical couplers based on multimode interference effects,"* IEEE J. Lightwave Technol., vol. 10, 1992, pp. 1843 - 1850

[42] P. Besse, M. Bachmann & H. Melchior, *"Phase relations in multi-mode interference couplers and their application to generalized integrated Mach-Zehnder optical switches,"* 6th Eur. Conf. Int. Opt., Neuchatel, Switzerland, 1993

[43] A. Ferreras, F. Rodriguez, E. Gomez-Salas, J. de Miguel & F. Hernandez-Gil, *"Useful formulas for multimode interference power splitter/combiner design,"* IEEE Photon. Technol. Lett., vol. 5, 1993, pp. 1224 - 1227

[44] P. Besse, M. Bachmann, H. Melchior, L. Soldano & M. Smit, *"Optical bandwidth and fabrication tolerances of multimode interference couplers,"* IEEE J. Lightwave Technol., vol. 12, 1994, pp. 1004 - 1009

[45] M. Papuchon & A. Roy, *"Electrically active optical bifurcation: BOA,"* Appl. Phys. Lett., vol. 31, 1977, pp. 266 - 267

[46] C. Caldera, C. De Bernardi, G. Destefanis, M. Meliga, S. Morasca, C. Rigo & A. Stano, *"3-dB couplers integrated in InGaAlAs/InP for coherent applications,"* Proc. Eur. Conf. Opt. Comm., 1990, pp. 357 - 360

[47] E. Pennings, R. Deri, A. Scherer, R. Bhat, T. Hayes, N. Andreadakis, M. Smit, L. Soldano & R. Hawkins, *"Ultracompact, low-loss directional couplers on InP based on self-imaging by multimode interference,"* Appl. Phys. Lett., vol. 59, 1991, pp. 1926 - 1928

[48] O. Prat, Y. Bourbin, A. Enard & M. Papuchon, *"Diagrammes de commutation du coupleur a jonction Y,"* Revue Technique Thomson-CSF, vol. 19, 1987, pp. 551 - 573

[49] S. Korotky, *"Three-space representation of phase-mismatch switching in coupled two-state optical systems,"* IEEE J. Quantum Electron., vol. 22, 1986, pp. 952 - 958

[50] C. Rolland, R. Moore, F. Shepherd & G. Hillier, *"10 Gbit/s 1.56 μm multiquantum well InP/InGaAsP Mach-Zehnder optical modulator,"* Electron. Lett., vol. 29, 1993, pp. 471 - 472

[51] C. Rolland, G. Mak, W. Bardyszewski & D. Yevick, *"Improved extinction ratio of waveguide electroabsorption modulators induced by an InGaAs absorbing layer,"* IEEE J. Lightwave Technol., vol. 10, 1992, pp. 1907 - 1911

[52] C. Rolland, D. Adams, D. Yevick & B. Hermansson, *"Optimization of strongly guiding semiconductor rib waveguide y-junctions,"* IEEE Photon. Technol. Lett., vol. 2, 1990, pp. 404 - 406

[53] S. Adachi, *"Refractive indices of III-V compounds: key properties of InGaAsP relevant to device design,"* J. Appl. Phys., vol. 53, 1982, pp. 5863 - 5869

[54] E. Perea, E. Mendez & G. Fonstad, *"Electroreflectance of indium gallium arsenide phosphide lattice matched to indium phosphide,"* Appl. Phys. Lett., vol. 36, 1980, pp. 978 - 980

[55] R. Nahory, M. Pollack, W. Johnston, Jr. & R. Barns, *"Band gap vs. composition and demonstration of Vegard's law for* $In_{1-x}Ga_xAs_yP_{1-y}$ *lattice matched to InP,"* Appl. Phys. Lett., vol. 33, 1978, pp. 650 - 657

[56] G. Alman, L. Molter, H. Shen & M. Dutta, *"Refractive index approximations from linear perturbation theory for planar MQW waveguides,"* IEEE J. Quantum Electron., vol. 28, 1992, pp. 650 - 657

[57] J. Zucker, K. Jones, T. Chin, B. Tell & K. Brown-Goebeler, *"Strained quantum wells for polarization-independent electrooptic waveguide switches,"* IEEE J. Lightwave Technol., vol. 10, 1992, pp. 1926 - 1930

[58] N. Agrawal, C. Bornholdt, H.-J. Ehrke, D. Franke, D. Hoffman, F. Kappe, R. Langenhorst, G. Mekonnen, F. Reier & C. Weinert, *"Semiconductor based fast optical switches: principles, structures, performance,"* 5th Optoelectron. Conf., Makuhari, Japan, 1994

[59] C. Janz, B. Keyworth, W. Allegretto, R. MacDonald, M. Fallahi, G. Hillier & C. Rolland, *"Mach-Zehnder switch employing an ultra-compact directional coupler,"* Eur. Conf. Lasers & Electro-Opt., Amsterdam, the Netherlands, 1994

[60] C. Janz, B. Keyworth, W. Allegretto, R. MacDonald, M. Fallahi, G. Hillier & C. Rolland, *"Mach-Zehnder switch using an ultra-compact directional coupler in a strongly-confining rib structure,"* IEEE Photon. Technol. Lett., vol. 6, 1994, pp. 981 - 983

[61] T. Kitagawa, K. Hattori, Y. Hibino, Y. Ohmori & M. Horiguchi, *"Laser oscillation in Er-doped silica-based planar ring resonator,"* Proc. Eur. Conf. Opt. Com. 1992, pp. 907 - 910

[62] W. Wang, S. Honkanen, S. Najafi & A. Tervonen, *"New integrated optical ring resonator in glass,"* Electron. Lett., vol. 28, 1992, pp. 1967 - 1968

[63] T. Negami, H. Haga & S. Yamamoto, *"Guided-wave optical wavelength demultiplexer using an asymmetric y-junction,"* Appl. Phys. Lett., vol. 54, 1989, pp. 1080 - 1082

[64] T. Hoekstra, H. van Weerden, P. Lambeck & Th. Popma, *"Bent asymmetric y-branch integrated optical broadband wavelength multi/demultiplexer,"* IEEE Photon. Technol. Lett., vol. 6, 1994, pp. 83 - 85

[65] M. Heiblum & J. Harris, *"Analysis of curved optical waveguides by conformal transformation,"* IEEE J. Quantum Electron., vol. 11, 1975, pp. 75 - 83

[66] I. Goyal, R. Gallawa & K. Ghatak, *"Bent planar waveguides and whispering gallery modes: a new method of analysis,"* IEEE J. Lightwave Technc , vol. 8, 1990, pp. 768 - 774

[67] A. Delage, K. McGreer & E. Rainville, *"Modelling of circular wave guides,"* Can. J. Phys., vol. 70, 1992, pp. 1092 - 1098

[68] M. Smit, E. Pennings & H. Blok, *"A normalized approach to the design of low-loss optical waveguide bends,"* IEEE J. Lightwave Technol., vol. 11, 1993, pp. 1737 - 1742

[69] Lord Rayleigh, *"The problem of the whispering gallery,"* The London, Edinburgh and Dublin Philos. Mag. J. Sci., series 6, vol. 20, 1910, pp. 1001 - 1004

[70] A. Kumar, R. Gallawa & I. Goyal, *"Modal characteristics of bent dual mode planar optical waveguides,"* IEEE J. Lightwave Technol., vol. 12, 1994, pp. 621 - 624

[71] M. Aoki, H. Sano, M. Suzuki, M. Takahashi, K. Uomi & A. Takai, *"Novel structure MQW electroabsorption modulator/DFB-laser integrated device fabricated by selective area MOCVD growth,"* Electron. Lett., vol. 27, 1991, pp. 2138 - 2140

[72] T. Kato, T. Sasaki, K. Komatsu & I. Mito, *"DFB-LD/modulator integrated light source by bandgap energy controlled selective MOVPE,"* Electron. Lett., vol. 28, 1992, pp. 153 - 154

[73] K. Sato, I. Kotaka, K. Wakita, Y. Kondo & M. Yamamoto, *"Strained-InGaAsP MQW electroabsorption modulator integrated DFB laser,"* Electron. Lett., vol. 29, 1993, pp. 1087 - 1089

[74] A. Ramdane, A. Ougazzaden, F. Devaux, F. Delorme, M. Schneider & J. Landreau, *"Very simple approach for high-performance DFB laser-electroabsorption modulator monolithic integration,"* Electron. Lett., vol. 30, 1994, pp. 1980 - 1981

[75] S. O'Brien, J. Shealy & G. Wicks, *"Monolithic integration of an (Al) GaAs laser and an intracavity electroabsorption modulator using selective partial interdiffusion,"* Appl. Phys. Lett., vol. 58, 1991, pp. 1363 - 1365

[76] Y. Arakawa & A. Yariv, *"Theory of gain, modulation response and spectral linewidth in AlGaAs quantum well lasers,"* IEEE J. Quantum Electron., vol. 21, 1985, pp. 1666 - 1674

[77] R. Yan, S. Corzine, L. Coldren & I. Suemine, *"Corrections to the expression for gain in GaAs,"* IEEE J. Quantum Electron., vol. 26, 1990, pp. 213 - 216

[78] M. Asada, A. Kameyama & Y. Suematsu, *"Gain and intervalence band absorption in quantum-well lasers,"* IEEE J. Quantum Electron., vol. 20, 1984, pp. 745 - 753

[79] Z.-M. Li, K. Dzurko, A. Delage & S. McAlister, *"A self-consistent two-dimensional model of quantum-well semiconductor lasers: optimization of a GRIN-SCH SQW laser structure,"* IEEE J. Quantum Electron., vol. 28, 1992, pp. 792 - 803

[80] R. Nagarajan, T. Kamiya & A. Kurobe, *"Band filling in GaAs/AlGaAs multiquantum well lasers and its effect on the threshold current,"* IEEE J. Quantum Electron., vol. 25, 1989, pp. 1161 - 1170

[81] R. Olshansky, C. Su, J. Manning & W. Powazinik, *"Measurement of radiative and non-radiative recombination rates in InGaAsP and AlGaAs light sources,"* IEEE J. Quantum Electron., vol. 20, 1984, pp. 838 - 854

[82] E. Zielinski, H. Schweizer, S. Hausser, R. Stuber, M. Pilkuhn & G. Weimann, *"Systematics of laser operation in GaAs/AlGaAs multiquantum well structures,"* IEEE J. Quantum Electron., vol. 23, 1987, pp. 969 - 975

[83] J. Fouquet & R. Burnham, *"Recombination dynamics in GaAs/$Al_xGa_{1-x}As$ quantum well structures,"* IEEE J. Quantum Electron., vol. 22, 1986, pp. 1799 - 1810

[84] J. Fouquet & A. Siegman, *"Room-temperature photoluminescence times in GaAs/$Al_xGa_{1-x}As$ molecular beam epitaxy multiple quantum well structure,"* Appl. Phys. Lett., vol. 46, 1985 pp. 280 - 282

[85] T. Wood, *"Multiple quantum well (MQW) waveguide modulators,"* IEEE J. Lightwave Technol., vol. 6, 1988, pp. 743 - 757

[86] B. Bransden & C. Joachain, *Introduction to quantum mechanics,* Longman Scientific & Technical, Essex, UK, 1989, pp. 422 - 424

[87] D. Miller, J. Weiner & D. Chemla, *"Electric-field dependence of linear optical properties in quantum well structures: waveguide electroabsorption and sum rules,"* IEEE J. Quantum Electron., vol.22, 1986, pp. 1816 - 1829

[88] P. Stevens, M. Whitehead, G. Parry & K. Woodbridge, *"Computer modeling of the electric field dependent absorption spectrum of multiple quantum well material,"* IEEE J. Quantum Electron., vol. 24, 1988, pp. 2007 - 2016

[89] G. Lengyel, K. Jelley & R. Engelmann, *"A semi-empirical model for electroabsorption in GaAs/AlGaAs multiple quantum well modulator structures,"* IEEE J. Quantum Electron., vol. 26, 1990, pp. 296 - 304

[90] A. Ghatak, R. Gallawa & I. Goyal, *"Accurate solutions to Schrodinger's equation using modified Airy functions,"* IEEE J. Quantum Electron., vol. 28, 1992, pp. 400 - 403

[91] G. Bastard, E. Mendez, L. Chang & L. Esaki, *"Variational calculations on a quantum well in an electric field,"* Phys. Rev. B, vol. 28, 1983, pp. 3241 - 3245

[92] K. Nakamura, A. Shimizu, K. Fujii, M. Koshiba & K. Hayata, *"Numerical anaylsis of the absorption and refractive index change in arbitrary semiconductor quantum-well structures,"* IEEE J. Quantum Electron., vol. 28, 1992, pp. 1670 - 1677

[93] A. Harwitt, J. Harris, Jr. & A. Kapitulnik, *"Calculated quasi-eigenstates and quasi-eigenenergies of quantum well superlattices in an applied electric field,"* J. Appl. Phys., vol. 60, 1986, pp. 3211 - 3213

[94] S. Adachi, *"GaAs, AlAs and $Al_xGa_{1-x}As$: material parameters for use in research and device applications,"* J. Appl. Phys., vol. 58, 1985, pp. R1

[95] P. McIlroy, A. Kurobe & Y. Uematsu, *"Analysis and application of theoretical gain curves to the design of multi-quantum-well lasers,"* IEEE J. Quantum Electron., vol. 21, 1985, pp. 1958 - 1963

[96] B. MacCraith, V. Ruddy, C. Potter, B. O'Kelly, & J. McGilp, *"Optical waveguide sensor using evanescent wave excitation of fluorescent dye in sol-gel glass,"* Electron. Lett., vol. 27, 1991, pp. 1247 - 1248

[97] Y. Zhou, P. Laybourn, J. MaGill & R. De La Rue, *"Evanescent fluorescence biosensor using ion-exchanged buried waveguides and the enhancement of peak fluorescence,"* Biosensors & Bioelectron., Vol. 6, 1991, pp. 595 - 607

[98] J. Kimmel & D. Christensen, *"FDTD modeling in the design of optical chemical sensor structures,"* Int. Conf. Solid-State Sensors and Actuators, San Francisco, U.S.A., 1991

[99] D. Christensen, S. Dyer, J. Kimmel & J. Herron, *"Evanescent coupling in a waveguide fluoroimmunosensor,"* Conf. Fiber Optic Medical and Fluorescent Sensors and Applications, Los Angeles, U.S.A., 1992

[100] J. Kimmel, & D. Christensen, *"Finite-difference time domain modeling and experimental characterization of planar waveguide fluorescence sensors,"* 3rd Conf. Chemical, Biochemical and Environmental Fiber Sensors, Boston, U.S.A., 1992

[101] S. Lal & M. Yappert, *"Development, characterization, and application of a double-waveguide evanescent sensor,"* Applied Spectroscopy, vol. 45, pp. 1607 - 1612

[102] E. Desurvire & J. Simpson, *"Amplification of spontaneous emission in Erbium-doped single-mode fibers,"* IEEE J. Lightwave Technol., vol. 7, 1989, pp. 835 - 845

[103] Y. Suematsu & K. Furuya, *"Theoretical spontaneous emission factor of injection lasers,"* Trans. IECE Japan, vol. E60, 1977, pp. 467 - 472

[104] K. Petermann, *"Calculated spontaneous emission factor for double heterostructure injection lasers with gain-induced waveguiding,"* IEEE J. Quantum Electron., vol. 19, 1979, pp. 566 - 570

[105] M. Philpott, *"Fluorescence from molecules between mirrors,"* Chem. Phys. Lett., vol. 19, 1973, pp. 435 - 439

[106] S. Ching, H. Lai & K. Young, *"Dielectric microspheres as optical cavities: thermal spectrum and density of states,"* J. Opt. Soc. Am. B, vol. 4, 1987, pp. 2004 - 2009

[107] H. Kreuwal, Planar Waveguide Sensors for the Chemical Domain, Doctoral Dissertation, University of Trent, Netherlands, 1988, pp. 107 - 121

[108] D. Marcuse, *"Launching light into fiber cores from sources located in the cladding,"* IEEE J. Lightwave Technol., vol. 6, 1988, pp. 1273 - 1279

[109] S. Brorson, H. Yokoyama & E. Ippen, *"Spontaneous emission rate alteration in optical waveguide structures,"* IEEE J. Quantum Electron., vol. 26, 1990, pp. 1492 - 1499

[110] K. Ujihara, A. Nakamura, O. Manba & X.-P. Feng, *"Spontaneous emission in a very short optical cavity with plane-parallel dielectric mirrors,"* Jap. J. Appl. Phys., vol. 30, 1991, pp. 3388 - 3398

[111] F. De Martini, M. Marrocco & P. Mataloni, *"Spontaneous emission in the optical microscopic cavity,"* Phys. Rev. A, vol. 43, 1991, pp. 2480 - 2497

[112] G. Bjork, *"On the spontaneous lifetime change in an ideal planar microcavity - transition from a mode continuum to quantized modes,"* IEEE J. Quantum Electron., vol. 30, 1994, pp. 2314 - 2318

[113] C. Lin, D. Deppe & C. Lei, *"Role of waveguide light emission in planar microcavities,"* IEEE J. Quantum Electron., vol. 30, 1994, pp. 2304 - 2313

[114] E. Snoeks, A. Lagendijk, M. Brongersma & A. Polman, *"Spontaneous emission rate of erbium implanted in silica near a dielectric interface,"* Paper QWD35, Fifth European Quantum Electronics Conference, Amsterdam, The Netherlands, 1994

[115] C. Janz & J. McMullin, *"Spontaneous emission coupling to radiation and guided modes of planar waveguide structures,"* IEEE J. Quantum Electron., to be published, July, 1995

[116] P. Brosson, *"Analytical model of a semiconductor optical amplifier,"* IEEE J. Lightwave Technol., vol. 12, 1994, pp. 49 - 54

[117] G. Nykolak, M. Haner, P. Becker, J. Shmulovich & Y. Wong, *"Systems evaluation of an Er^{3+}-doped planar waveguide amplifier,"* IEEE Photon. Technol. Lett., vol. 5, 1993, pp. 1185 - 1187

[118] H. A. Macleod, Thin Film Optical Filters, 2nd. Ed. New York: McGraw-Hill, 1986

[119] B. Pezeshki, F. Tong, J. Kash & D. Kisker, *"Vertical cavity devices as wavelength selective waveguides,"* IEEE J. Lightwave Technol., vol. 12, 1994, pp. 1791 - 1801

[120] Y. Chen, J. Zucker, N. Sauer & T. Chang, *"Polarization-independent strained InGaAs/InGaAlAs quantum-well phase modulators,"* IEEE Photon. Technol. Lett., vol. 4, 1992, pp. 1120 - 1123

[121] T. Yamaguchi, T. Morimoto, K. Akeura, K. Tada & Y. Nakano, *"Polarization-independent waveguide modulator using a novel quantum well with mass-dependent width,"* IEEE Photon. Technol. Lett., vol. 6, 1994, pp. 1442 - 1444

[122] J. Aarnio, P. Heimala & I. Stuns, *"Wide passband wavelength multi/demultiplexer at 1.3/1.55 μm based on etched Fresnel mirror,"* IEE Proc.-Optoelectron., vol. 139, 1992, pp. 228 - 236

[123] B. Keyworth, D. Corazza, W. Fuchs & J. McMullin, *"Low-cost direct-write multimode polymer waveguides,"* LEOS Meeting Int. Optoelectron., Lake Tahoe, Nevada, 1994

Appendix

Here we state the expressions for the mode fields of the symmetric slab waveguide, and the corresponding normalization constants. We must consider TE and TM bound and radiating modes. The parameter q is given by (6.33).

The radiation modes are described by:

$$\vec{F}_k(x) \rightarrow \vec{F}(\omega,Q,x) = \left\{\begin{array}{l} A\cos\left[\sqrt{q^2+Q^2}\,x\right] \\ \cos[Q(x-a/2)+\delta] \end{array}\right. \hat{y} \quad \begin{array}{c} 0 \le x < a/2 \\ x \ge a/2 \end{array} \tag{A.1.a}$$

for even TE modes;

$$\vec{F}_k(x) \rightarrow \vec{F}(\omega,Q,x) = \left\{\begin{array}{l} A\sin\left[\sqrt{q^2+Q^2}\,x\right] \\ \sin[Q(x-a/2)+\delta] \end{array}\right. \hat{y} \quad \begin{array}{c} 0 \le x < a/2 \\ x \ge a/2 \end{array} \tag{A.1.b}$$

for odd TE modes;

$$\vec{F}_k(x) \rightarrow \vec{F}(\omega,Q,x)$$

$$= \left\{\begin{array}{ll} \dfrac{An_2c_o}{n_1^2\omega}\left[\left(\sqrt{\dfrac{n_2^2\omega^2}{c_o^2}-Q^2}\cos\left[\sqrt{q^2+Q^2}\,x\right]\right)\hat{x}+j\left(Q\sin\left[\sqrt{q^2+Q^2}\,x\right]\right)\hat{z}\right] & 0 \le x < a/2 \\ \dfrac{c_o}{n_2\omega}\left[\left(\sqrt{\dfrac{n_2^2\omega^2}{c_o^2}-Q^2}\cos[Q(x-a/2)+\delta]\right)\hat{x}+j\left(Q\sin[Q(x-a/2)+\delta]\right)\hat{z}\right] & x \ge a/2 \end{array}\right. \tag{A.1.c}$$

for even TM modes, and:

$$
= \begin{cases} \dfrac{An_2c_o}{n_1^2\omega}\left[\left(\sqrt{\dfrac{n_2^2\omega^2}{c_o^2}-Q^2}\,\sin\left[\sqrt{q^2+Q^2}\,x\right]\right)\hat{x}-j\left(Q\cos\left[\sqrt{q^2+Q^2}\,x\right]\right)\hat{z}\right] & 0 \le x < a/2 \\ \dfrac{c_o}{n_2\omega}\left[\left(\sqrt{\dfrac{n_2^2\omega^2}{c_o^2}-Q^2}\,\sin\left[Q(x-a/2)+\delta\right]\right)\hat{x}-j\left(Q\cos\left[Q(x-a/2)+\delta\right]\right)\hat{z}\right] & x \ge a/2 \end{cases}
$$

(A.1.d)

for odd TM modes. In (A.1), a is the full width of the guide (n_1) layer; definitions are given only for $x \ge 0$ since the slab waveguide is symmetric. The parameters A and δ are obtained by satisfying the continuity conditions at the index interfaces:

$$
A = \begin{cases} Q/\sqrt{Q^2+q^2\sin^2(a\sqrt{q^2+Q^2}/2)} & \textit{Even} \\ & ; \\ Q/\sqrt{Q^2+q^2\cos^2(a\sqrt{q^2+Q^2}/2)} & \textit{Odd} \end{cases} \qquad \textbf{(A.2.a)}
$$

for TE modes;

$$
A = \begin{cases} \dfrac{n_1^2Q}{\sqrt{Q^2\left\{n_1^4\cos^2(a\sqrt{q^2+Q^2}/2)+n_2^4\sin^2(a\sqrt{q^2+Q^2}/2)\right\}+n_2^4q^2\sin^2(a\sqrt{q^2+Q^2}/2)}} & \textit{Even} \\ & ; \\ \dfrac{n_1^2Q}{\sqrt{Q^2\left\{n_1^4\sin^2(a\sqrt{q^2+Q^2}/2)+n_2^4\cos^2(a\sqrt{q^2+Q^2}/2)\right\}+n_2^4q^2\cos^2(a\sqrt{q^2+Q^2}/2)}} & \textit{Odd} \end{cases} \qquad \textbf{(A.2.b)}
$$

for TM modes;

$$\delta = \begin{cases} \pm\cos^{-1}\left[A\cos(a\sqrt{q^2+Q^2}/2)\right] & Even \\ \pm\cos^{-1}\left[(A\sqrt{q^2+Q^2}/Q)\cos(a\sqrt{q^2+Q^2}/2)\right] & Odd \end{cases}; \qquad \textbf{(A.2.c)}$$

for TE modes, and

$$\delta = \begin{cases} \pm\cos^{-1}\left[(n_2^2/n_1^2)A\cos(a\sqrt{q^2+Q^2}/2)\right] & Even \\ \pm\cos^{-1}\left[(n_2^2A\sqrt{q^2+Q^2}/n_1^2Q)\cos(a\sqrt{q^2+Q^2}/2)\right] & Odd \end{cases}; \qquad \textbf{(A.2.d)}$$

for TM modes, where:

$$\pm \rightarrow \left\langle \begin{array}{l} \sin(a\sqrt{q^2+Q^2}/2) \geq 0 \rightarrow + \\ \sin(a\sqrt{q^2+Q^2}/2) < 0 \rightarrow - \end{array} \right. \qquad \textbf{(A.2.e)}$$

in (A.2.c) and (A.2.d). The normalization constant for radiating modes is obtained as:

$$N_k^2 \rightarrow N^2(\omega) = \frac{\hbar\omega}{2n_2^2\varepsilon_o L} \qquad \textbf{(A.3)}$$

The bound mode functions are given by:

$$\vec{F}_k(x) \rightarrow \vec{F}(\omega,K,x) = \begin{cases} \cos Kx \\ \cos(Ka/2)\exp[-\sqrt{q^2-K^2}(x-a/2)] \end{cases} \hat{y} \qquad \begin{array}{c} 0 \leq x < a/2 \\ x \geq a/2 \end{array}$$

(A.4.a)

for even TE modes;

$$\vec{F}_k(x) \rightarrow \vec{F}(\omega,K,x) = \begin{cases} \sin Kx \\ \sin(Ka/2)\exp[-\sqrt{q^2-K^2}(x-a/2)] \end{cases} \hat{y} \qquad \begin{array}{c} 0 \leq x < a/2 \\ x \geq a/2 \end{array}$$

(A.4.b)

for odd TE modes;

$$\bar{F}_k(x) \to \bar{F}(\omega, K, x)$$

$$= \begin{cases} \dfrac{\sqrt{\dfrac{n_1^2\omega^2}{c_o^2} - K^2}}{n_1^2}\cos(Kx)\hat{x} + j\dfrac{K}{n_1^2}\sin(Kx)\hat{z} & 0 \le x < a/2 \\ \dfrac{\cos(Ka/2)}{n_2^2}\left\{\sqrt{\dfrac{n_1^2\omega^2}{c_o^2} - K^2}\,\hat{x} + j\sqrt{q^2 - K^2}\,\hat{z}\right\}e^{-\sqrt{q^2-K^2}(x-a/2)} & x \ge a/2 \end{cases} \tag{A.4.c}$$

for even TM modes, and:

$$\bar{F}_k(x) \to \bar{F}(\omega, K, x)$$

$$= \begin{cases} \dfrac{\sqrt{\dfrac{n_1^2\omega^2}{c_o^2} - K^2}}{n_1^2}\sin(Kx)\hat{x} - j\dfrac{K}{n_1^2}\cos(Kx)\hat{z} & 0 \le x < a/2 \\ \dfrac{\sin(Ka/2)}{n_2^2}\left\{\sqrt{\dfrac{n_1^2\omega^2}{c_o^2} - K^2}\,\hat{x} + j\sqrt{q^2 - K^2}\,\hat{z}\right\}e^{-\sqrt{q^2-K^2}(x-a/2)} & x \ge a/2 \end{cases} \tag{A.4.d}$$

for odd TM modes. In (A.4); K must satisfy the equations:

$$\begin{matrix}+\\-\end{matrix} K \begin{matrix}\tan\\\cot\end{matrix}(Ka/2) = \sqrt{q^2 - K^2} \tag{A.5.a}$$

for TE modes, and:

$$\begin{matrix}+\\-\end{matrix} K \begin{matrix}\tan\\\cot\end{matrix}(Ka/2) = (n_1/n_2)^2\sqrt{q^2 - K^2} \tag{A.5.b}$$

for TM modes, where $+\tan$ is used for even modes, and $-\cot$ is used for odd modes.

For the bound modes, normalization yields:

$$N_k^2 \rightarrow N^2(\omega,K) = \begin{cases} \dfrac{\hbar\omega/\varepsilon_o}{n_1^2\left\{a+\dfrac{\sin Ka}{K}\right\}+\dfrac{2n_2^2\cos^2(Ka/2)}{\sqrt{q^2-K^2}}} & Even \\[3ex] \dfrac{\hbar\omega/\varepsilon_o}{n_1^2\left\{a-\dfrac{\sin Ka}{K}\right\}+\dfrac{2n_2^2\sin^2(Ka/2)}{\sqrt{q^2-K^2}}} & Odd \end{cases} \qquad \textbf{(A.6.a)}$$

for TE modes, and:

$$N_k^2 \rightarrow N^2(\omega,K)$$

$$= \begin{cases} \dfrac{\hbar\omega/\varepsilon_o}{\dfrac{\omega^2}{c_o^2}\left\{a+\dfrac{\sin Ka}{K}\right\}+\dfrac{K^2}{n_1^2}\left\{a-\dfrac{\sin Ka}{K}\right\}+\dfrac{2\left(\dfrac{n_2^2\omega^2}{c_o^2}-2K^2\right)\cos^2(Ka/2)}{n_2^2\sqrt{q^2-K^2}}} & Even \\[3ex] \dfrac{\hbar\omega/\varepsilon_o}{\dfrac{\omega^2}{c_o^2}\left\{a-\dfrac{\sin Ka}{K}\right\}+\dfrac{K^2}{n_1^2}\left\{a+\dfrac{\sin Ka}{K}\right\}+\dfrac{2\left(\dfrac{n_2^2\omega^2}{c_o^2}-2K^2\right)\sin^2(Ka/2)}{n_2^2\sqrt{q^2-K^2}}} & Odd \end{cases} \qquad \textbf{(A.6.b)}$$

for TM modes.

The mode power confinement factors used in (6.37) are given by:

$$\Gamma_K = \frac{\dfrac{a\sqrt{q^2 - K^2}}{2} + \sin^2\left(\dfrac{Ka}{2}\right)}{\dfrac{a\sqrt{q^2 - K^2}}{2} + 1} \tag{A.7.a}$$

for even TE modes;

$$\Gamma_K = \frac{\dfrac{a\sqrt{q^2 - K^2}}{2} + \cos^2\left(\dfrac{Ka}{2}\right)}{\dfrac{a\sqrt{q^2 - K^2}}{2} + 1} \tag{A.7.b}$$

for odd TE modes;

$$\Gamma_K = \frac{\dfrac{n_2^2}{n_1^2}\left\{\dfrac{a\sqrt{q^2 - K^2}}{2} + \dfrac{n_2^2}{n_1^2}\sin^2\left(\dfrac{Ka}{2}\right)\right\}}{\dfrac{n_2^2}{n_1^2}\left\{\dfrac{a\sqrt{q^2 - K^2}}{2} + \dfrac{n_2^2}{n_1^2}\sin^2\left(\dfrac{Ka}{2}\right)\right\} + \cos^2\left(\dfrac{Ka}{2}\right)} \tag{A.7.c}$$

for even TM modes, and:

$$\Gamma_K = \frac{\dfrac{n_2^2}{n_1^2}\left\{\dfrac{a\sqrt{q^2 - K^2}}{2} + \dfrac{n_2^2}{n_1^2}\cos^2\left(\dfrac{Ka}{2}\right)\right\}}{\dfrac{n_2^2}{n_1^2}\left\{\dfrac{a\sqrt{q^2 - K^2}}{2} + \dfrac{n_2^2}{n_1^2}\cos^2\left(\dfrac{Ka}{2}\right)\right\} + \sin^2\left(\dfrac{Ka}{2}\right)} \tag{A.7.d}$$

for odd TM modes.