Integrated Optical Grating-Based Matched Filters for Fiber-Optic Communications

by

Thomas E. Murphy

B.S. Electrical Engineering, Rice University (1994)
B.A. Physics, Rice University (1994)

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

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February, 1996

Signature of Author . . . .
Department of Electrical Engineering and Computer Science
October, 1996

Certified by . . . .
Henry I. Smith
Thesis Supervisor

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

Integrated Optical Grating-Based Matched Filters for Fiber-Optic Communications

by

Thomas E. Murphy

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

Master of Science

Abstract

The unavoidable presence of noise in optical communications systems makes it necessary to use filters in practical optical receivers. Noise, especially amplified spontaneous emission noise from optical amplifiers, is what ultimately limits the sensitivity of an optical communications system. This thesis describes an optical matched filter which operates on the principle of Bragg reflection from an integrated grating structure. A matched filter is predicted to yield the highest possible signal to noise ratio, which would allow receivers to more closely approach the theoretical limit in receiver sensitivity. We analyze the predicted behavior of the matched filter device, and present a complete description of the device design, including coupled mode theory, insertion loss minimization, and calculation of fabrication tolerances. We describe a fabrication process which addresses some of the unique nanolithography challenges associated with the device, and present some fabricated structures which demonstrate the feasibility of this process.

Thesis Supervisor: Henry I. Smith
Title: Keithly Professor of Electrical Engineering, MIT

Acknowledgments

It has been my pleasure to work under the expert mentorship of Professor Hank Smith in the NanoStructures Lab. I feel that one of the most important and sometimes difficult steps in learning how to conduct research is to decide how to formulate problems, work independently, and to set goals for oneself. More than any other professor that I know, Hank has fostered my independence while at the same time remaining approachable and encouraging. I would like to thank Hank for giving me the freedom to work on a project of my choosing that I have found truly rewarding and always interesting.

The work described in this thesis would not have been possible without the help of my coworkers and comrades in the optics sub-group. Brent Little provided many valuable discussions of waveguide coupling and optical physics and was always willing to take the time to explain things to me. Juan Ferrera has been a consistent source of good advice, on topics ranging from laboratory techniques to system administration. Mike Lim has been very generous with his time in helping me with the fabrication tasks involved in building these devices, and he has also been a great source of suggestions and speculations. If I tried to enumerate the ways in which Jay Damask has helped me with this work, this acknowledgments section would be as long as one of the chapters. From the inception of this project, Jay has been a source of encouragement and enlightenment. I must also thank him for his careful reading of this thesis. I have no doubt that much of the success of this work can be attributed to these very talented and dedicated individuals that I have had the opportunity to work with over the past two years.

Beyond my colleagues in the optics sub-group, there are several other members of the NSL who helped this project considerably. Most notably, Jim Carter and Mark Mondol have been very valuable sources of technical assistance.

I would like to thank Jeff Livas and Steve Chinn for originally motivating this project, and supporting my work in this direction and Hermann Haus for his invaluable physical insight and his willingness to discuss just about any aspect of the project.

I feel it necessary to mention also the friends who have been with me in the past few years. Although they have perhaps not contributed directly to this work, their support and encouragement outside of my research is part of what keeps me going those times when research isn't going as well as I would like. In particular, I would like to recognize Kent Pryor, my erstwhile roommate and chemistry consultant, Erik Thoen, my present apartment-mate, cooking companion, and chauffeur, and Richard Barron. Additionally I would like to thank other members of the NSL who have been both companions and coworkers, including Tim and Sofia Savas, Maya Farhoud, and David Carter. You have all helped me out, often without even knowing it, and I hope I can reciprocate.

Finally, I would like to thank my parents for their constant support over the years. The encouragement and love that they have selflessly and tirelessly invested in me is undoubtedly the greatest source of my ambition, inspiration, dedication and motivation. They have taught me far more than words can express.

TABLE OF CONTENTS

Chapter 1: Introduction

The subject of this thesis is the design and analysis of an optical matched filter for fiber optic communications. The principle goal of a matched filter is to improve the sensitivity of an optical receiver by optimally filtering a communications signal. A matched filter could potentially allow optical receivers to approach the theoretical limit of sensitivity, which is governed by the noise generated when the signal is optically amplified. The integrated Bragg grating filter could provide improved sensitivity over currently used filters, in a package that can be integrated with other optical and electronic components of the receiver system. We have attempted in this thesis to describe all aspects of the design of such a filter, ranging from a theoretical analysis of the interaction of light with a Bragg grating to the more practical aspects of fabrication and device construction.

Chapter 2 introduces the problem of noise in modern optical communications systems, and describes the basic relationships between the noise and bit error rates. Without presenting a rigorous analysis of filtering and detection, we describe the need for optical filters and discuss the projected benefits of a matched optical filter.

Chapter 3 is devoted entirely to analyzing the reflection from an integrated Bragg grating, starting from the coupled mode equations, which are later derived in Chapter 4. We begin with an analysis of the spectral response of a Bragg grating, demonstrating how a weak grating filter has a spectrum that is matched to the spectrum of an optical communications signal. Next, we describe the temporal response of the Bragg grating by analyzing precisely how the shape of a single bit is changed when it is reflected from a Bragg grating.

Chapter 4 describes those topics of guided wave optics that are necessary for an understanding of this work, including an overview of dielectric waveguides, a summary of commonly used integrated waveguide structures, and a derivation of the coupled mode equations describing a Bragg grating.

Chapter 5 describes the physics of codirectional waveguide coupling, and analyzes a device design that utilizes a codirectional coupler to spatially separate the reflected light from a Bragg grating from the input signal.

Chapter 6 addresses the practical issues of device optimization, with the goal of achieving a design yielding a low total insertion loss. This chapter analyzes several sources of loss, and provides practical design criteria for minimizing loss where possible.

Chapter 7 describes some of the tolerances of the matched filter device. We begin by examining how the device performance is affected by deviations from the nominal waveguide design parameters. The remainder of Chapter 7 describes a fabrication process that addresses these tolerance issues.

CHAPTER 2: THE BENEFIT OF MATCHED FILTERS IN OPTICAL COMMUNICATIONS

2.1 The Need for Filters in Communications Systems

Filters are needed in optical communications systems for a variety of reasons. For example, in wavelength division multiplexing (WDM), an optical fiber can be used to carry several different communications signals, each centered at a unique carrier frequency. WDM systems allow several signals to be closely packed in frequency, spanning the available bandwidth of the system. In order for such a system to be practical, a filter is needed at the receiving end to separate one signal of interest from the other signals which are centered at different frequencies. For wavelength demultiplexing, an appropriate filter would have a bandpass characteristic, with a flat spectral response over some narrow band and very low response outside of that band. Such a filter would allow undistorted transmission of a single optical signal, blocking all other signals.

More generally, filters are needed in any system where noise has been introduced in the signal to be detected, regardless of whether other carrier frequencies are present. For the purpose of filtering a noisy signal, a flat-response bandpass filter is not the most suitable choice because such a filter would transmit all of the noise within the filter bandwidth. Because the noise source typically overlaps with the signal to be detected, the optimal filter would instead have a spectral response which is similar to that of the signal to be detected. This thesis seeks to describe the design, analysis, and construction of such a filter.

This chapter discusses some of the sources of noise in optical communications systems, and the components which give rise to that noise. The effect of noise on the performance and sensitivity of optical receivers is discussed, and the potential benefits of a matched optical filter are described.

2.2 Noise in Communications Systems

Some degree of noise is inevitably present in any communications system because of the quantized nature of the photons or electrons which carry the signal. This noise, called shot noise, results from statistical fluctuations in the number of signal carriers. Naturally, shot noise becomes more of a limiting factor when the number of signal carriers is small. Additionally some amount of thermal noise is present when a signal is detected electronically. One of the more important sources of noise in modern optical communications systems is the noise associated with amplified spontaneous emission, which arises in all optical amplifiers.

The effect of noise in a communications system is usually quantified in terms of a bit-error-rate (BER), which is the frequency at which errors occur in transmission. In modern communications systems, a BER of $<10^{-9}$ is deemed acceptable. This corresponds to an average of less than one erroneous bit per 125 megabytes of transmitted data. External error correction techniques can be employed to detect and correct these infrequent errors allowing for essentially error free communications. The bit-error-rate is directly related to the signal-to-noise ratio of the detected signal. Accordingly, a lower BER can be achieved by increasing the input signal power. The sensitivity of an optical receiver system is described by the amount of average signal power required to achieve a sufficiently low BER. Figure 2.1 illustrates the relationship between bit-error-rate and signal power for a typical optical communications system [1]. In this example, in order to achieve a sufficiently low BER of 10^{-9}, the signal which arrives at the receiver must have an average power of at least 100 photons per bit. A *more* sensitive receiver would exhibit a similar exponential relationship between power and BER, but the corresponding curve would be shifted to the left indicating that less power is required to achieve the same bit-error-rate.

The thick dashed line in figure 2.1 represents the theoretical performance of an "ideal" preamplified receiver system, as described in [2, 3]. The sensitivity limit of an optical receiver depends upon how the optical signal is encoded, how it is detected, and which sources of noise are dominant in the receiver [4]. Fundamentally, the sensitivity of an optical receiver is limited by the statistics of photon counting. That is, the number of photons detected in a single bit interval follows a Poisson distribution [5]. However, this shot-noise limit cannot be easily approached in a practical receiver system because of the inevitable presence of thermal noise and other receiver noise when detecting the weak optical signal over a short time interval. For this reason, preamplified receiver systems are the most likely candidates for high bit-rate communications in the future. The limit plotted in 2.1 assumes that the optical signal is preamplified by an ideal optical amplifier with high enough gain that thermal noise, receiver noise, and shot noise are negligible. In this case, the sensitivity of the receiver is limited only by the amplified spontaneous emission noise generated in the optical amplifier.

Noise is usually only a problem in communications systems when the signal has been significantly attenuated. The problem of attenuation in optical communications was conventionally addressed by a process of signal regeneration. In signal regeneration, the signal is detected electronically before it becomes too weak to recognize, and the detected digital signal is then retransmitted at a higher power. With the advent of the erbium doped fiber amplifier, the cumbersome and costly signal regeneration process was replaced by an optical amplification stage which boosts the signal power without detecting, correcting, and retransmitting the signal.

Optical amplifiers are also used as preamplifiers in receiver systems. A simplified block diagram of an optically preamplified system is provided in Fig. 2.3. In this configuration, the optical amplifier amplifies the weak optical signal before detection. Because the signal power which reaches the detector is relatively large, the effect of shot noise and thermal noise in the electronic detector is usually negligible. However, the process of amplification introduces an additional source of noise, amplified spontaneous emission (ASE). As the name implies, ASE is associated with the random emission of photons in the active medium and subsequent amplification of these photons. ASE can be viewed as a broad spectrum background noise source that would be present at the output of the amplifier even in the absence of an incident signal [6]. The bandwidth of this ASE noise is governed only by the gain bandwidth of

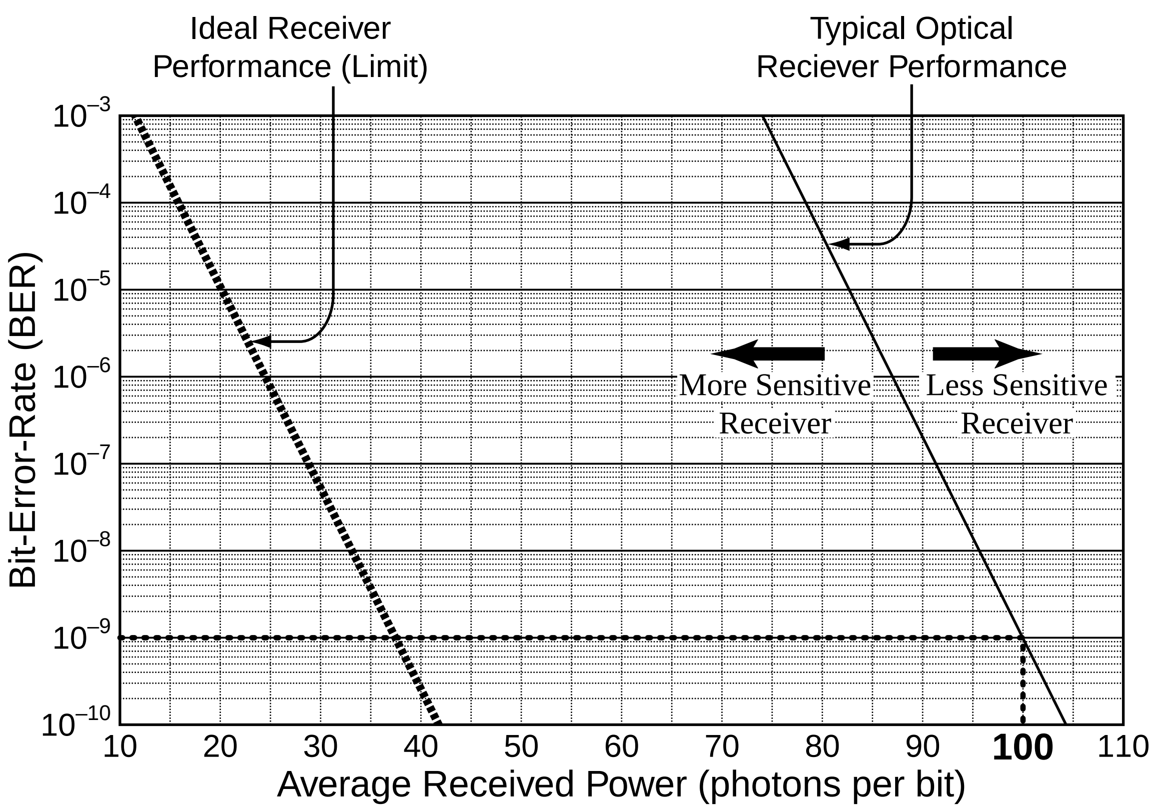


Figure 2.1 The relationship between signal power and bit-error-rate (BER) for a typical optical communications system. A more sensitive system would require less signal power to achieve the same error performance. In this figure, an acceptable BER of 10^{-9} can be achieved with an optical signal power of 100 photons per bit. The thick dashed line indicates the predicted performance of an ideal preamplified receiver. For the ideal case, the BER is limited by the amplified spontaneous emission noise introduced in the optical amplifier, as described in [2].

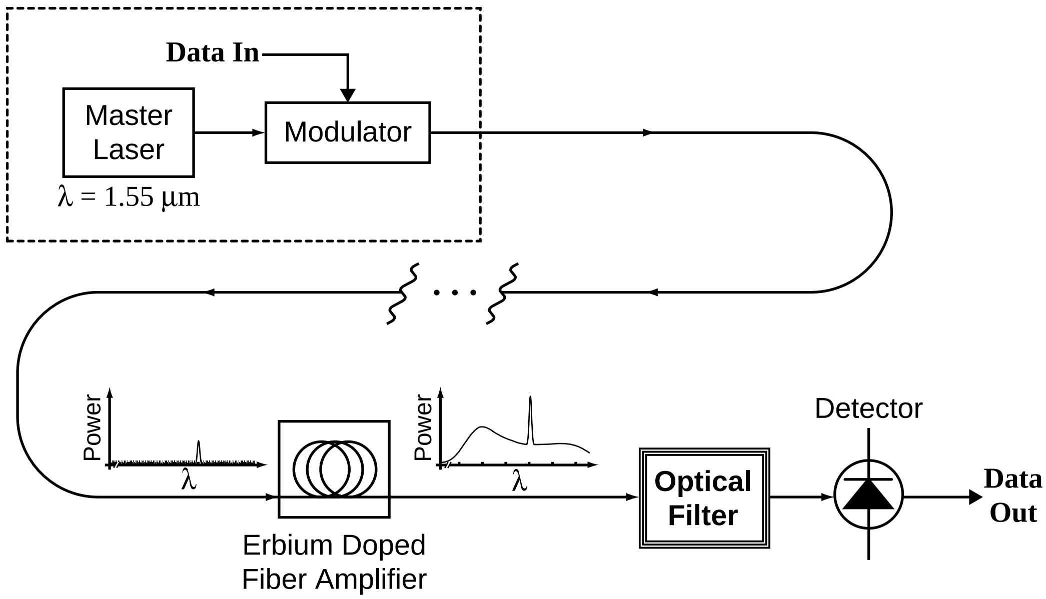


Figure 2.2 Block diagram of a typical preamplified optical receiver system. The weak signal is preamplified with an erbium doped fiber amplifier, which boosts the signal power but introduces broad-band ASE noise which must be filtered out before the signal is detected.

the optical amplifier, which is typically much wider than the bandwidth of the communications signal, as depicted in Fig. 2.3. The presence of ASE in optically amplified systems and receivers necessitates the use of a narrow-band filter to remove as much of the ASE as possible before the signal is detected.

2.3 A Matched Filter for Optical Communications

Norbert Wiener studied the problem of filtering of noisy signals and determined that for filtering out broad band noise, the optimal filter is one that has a spectral response that exactly matches the spectral response of the original unadulterated data signal [7]. Although it is impossible to completely remove the noise from the signal, such a matched filter can be shown to yield the highest possible signal-to-noise ratio (SNR).

2.3.1 Encoding Binary Information in Optical Signals

To understand the spectrum of an optical communications signal, it is necessary to know how a sequence of binary data is encoded into an optical signal. There are several methods of encoding information on an optical signal: one can modulate the amplitude, the phase, the frequency, or the polarization of an optical carrier signal.

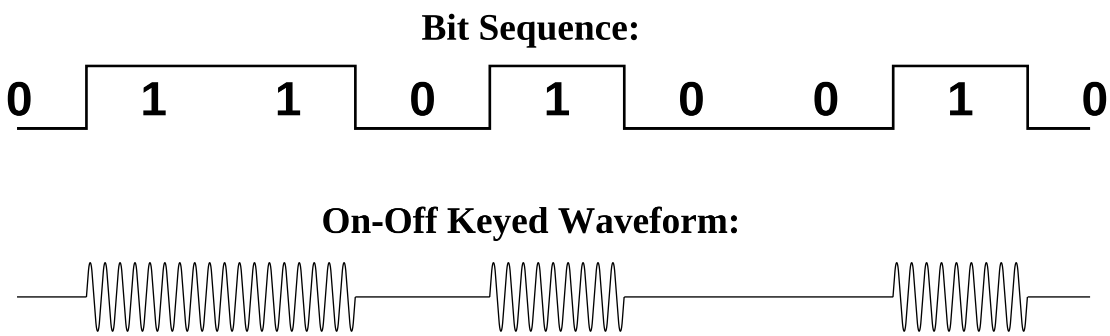


Figure 2.3 Encoding binary data using on-off keying. In common optical communications systems, the carrier frequency is several orders of magnitude higher than the communications bit rate. The frequency of the optical carrier frequency has been greatly exaggerated in this figure for illustrative purposes.

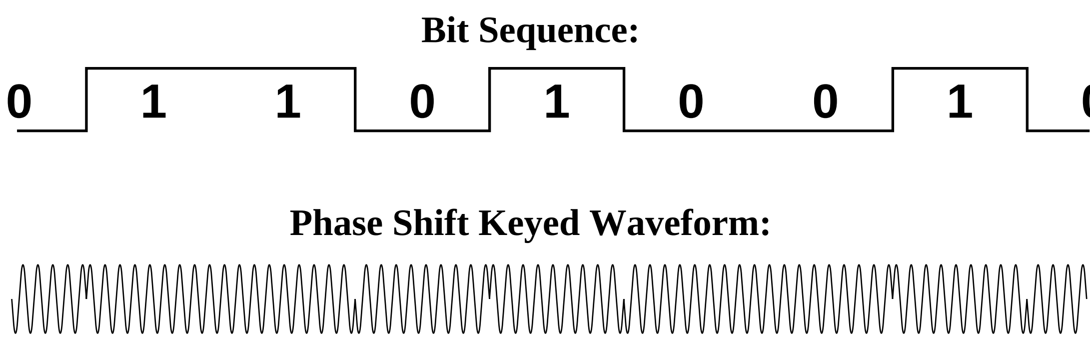


Figure 2.4 Encoding binary data using phase-shift keying. The information is stored in the phase of the optical carrier signal. In this case, the difference between a 0 and 1 is indicated by a 180 degree phase shift. As in Fig. 2.3, the period of the optical carrier has been greatly exaggerated in this figure for illustrative purposes.

The simplest method of sending binary information on an optical signal is amplitude-shift-keyed (ASK) modulation. ASK is analogous to amplitude modulation in radio transmission. To represent a sequence of binary data, the amplitude of an optical carrier is switched between two or more discrete amplitude levels. A special case of amplitude shift keying is on-off keying (OOK), in which there are only two discrete levels between which the carrier is switched, the lower level being 0. In OOK, a binary sequence of 0's and 1's can be represented by switching an optical carrier signal on and off at the data transmission rate. A binary 1 would be represented by a non-zero amplitude state, existing for the duration of one bit timeslot. Such a scheme is also commonly referred to as non-return-to-zero (NRZ) encoding, because between two consecutive binary 1's the amplitude does not return to zero. Figure 2.3 illustrates how a binary sequence of ones and zeros would be represented in an OOK optical signal.

Another way to encode binary information is to modulate the phase of the carrier, keeping the amplitude constant. Such a scheme is known as phase-shift-keying (PSK). Figure 2.4 illustrates how a binary signal would be represented in a PSK scheme, with a phase shift of 180° between 1's and 0's. In order to distinguish the 0's from 1's, it is necessary for the receiver to have a stable local oscillator which is phase locked to the transmitter. The value of each bit is then determined by comparing the phase with that of the reference signal. Because of the difficulty of generating a stable local oscillator, a variation of PSK called differential-phase-shift-keying (DPSK), is sometimes used. In this modified scheme, ones and zeros are represented by the difference in phase between adjacent bits. A reference oscillator is no longer necessary because the receiver does not need to determine the absolute phase, only the relative difference between adjacent bits. Figure 2.5 depicts a sequence of data encoded with DPSK, where a phase change of 180° between adjacent bits indicates a 1 and a phase change of 0° indicates a binary 0. Note that in Figs. 2.3, 2.4, and 2.5 the period of the optical oscillations has been greatly exaggerated for the purpose of illustration. In a modern optical communications system there would be of order 10^4 optical cycles per bit of information.

Yet another way to represent binary data is to encode information in the frequency of the carrier, which is called frequency-shift-keying (FSK). The spectrum of a FSK waveform is more complicated than that of a PSK or ASK signal, and therefore the construction of a matched filter for FSK signals is not straightforward. For this reason, generation of a matched filter for FSK signals is not considered in this work.

2.3.2 Spectrum of Binary Optical Signals

The spectrum of a sequence of discrete pulses comprising an optical signal can be derived from the spectrum of an isolated pulse by a process of superposition. For the case of an OOK signal, the spectrum of an isolated pulse can be found by computing the Fourier spectrum. We will consider one pulse of duration T and amplitude A_0, centered at time t=0:

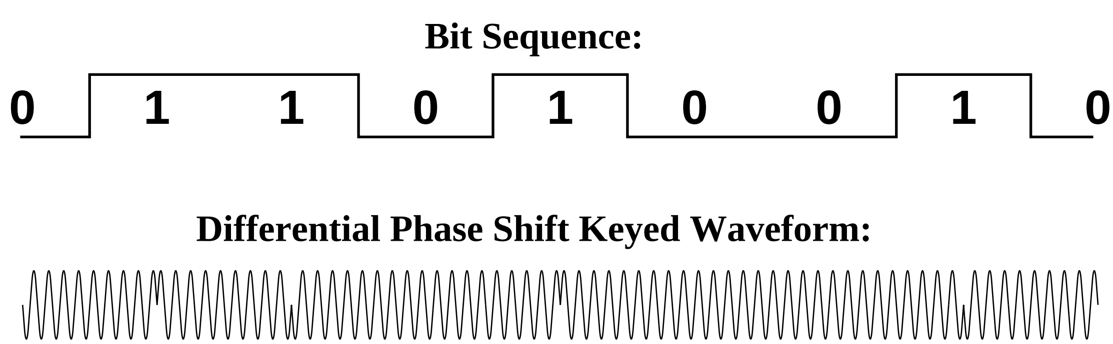


Figure 2.5 Encoding binary data using differential-phase-shift keying. A binary 1 is encoded as a 180 degree phase shift, and a binary zero is indicated by no phase shift. As in Figs. 2.3 and 2.4, the period of the optical carrier has been greatly exaggerated in this figure for illustrative purposes.

$$f(t) = \begin{cases} A_0 e^{-i\omega_0 t} & |t| \le \dfrac{T}{2} \\ 0 & |t| > \dfrac{T}{2} \end{cases} \tag{2.1}$$

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt = A_0 T \frac{\sin\left(\frac{1}{2}(\omega - \omega_0)T\right)}{\frac{1}{2}(\omega - \omega_0)T} \tag{2.2}$$

The spectrum given in equation 2.2 is plotted in Fig. 2.6. As expected, the spectral width of the signal is inversely proportional to T, which is the bit duration. The signal spectrum has a characteristic sin(x)/x (also known as sinc(x)) shape*, and is centered about the optical carrier frequency of ν_0. The carrier wavelength is typically around 1.55 μm (194 THz in frequency), which is the wavelength at which optical fibers have the lowest propagation loss. A typical data stream might contain data modulated at 10 Gb/s (Gigabits per second), corresponding to a bit duration of 100 ps and a signal bandwidth of 10 GHz.

Of course, a real communications signal containing information is not an isolated pulse, but rather contains a random sequence of such pulses. It can be shown that for a random sequence of

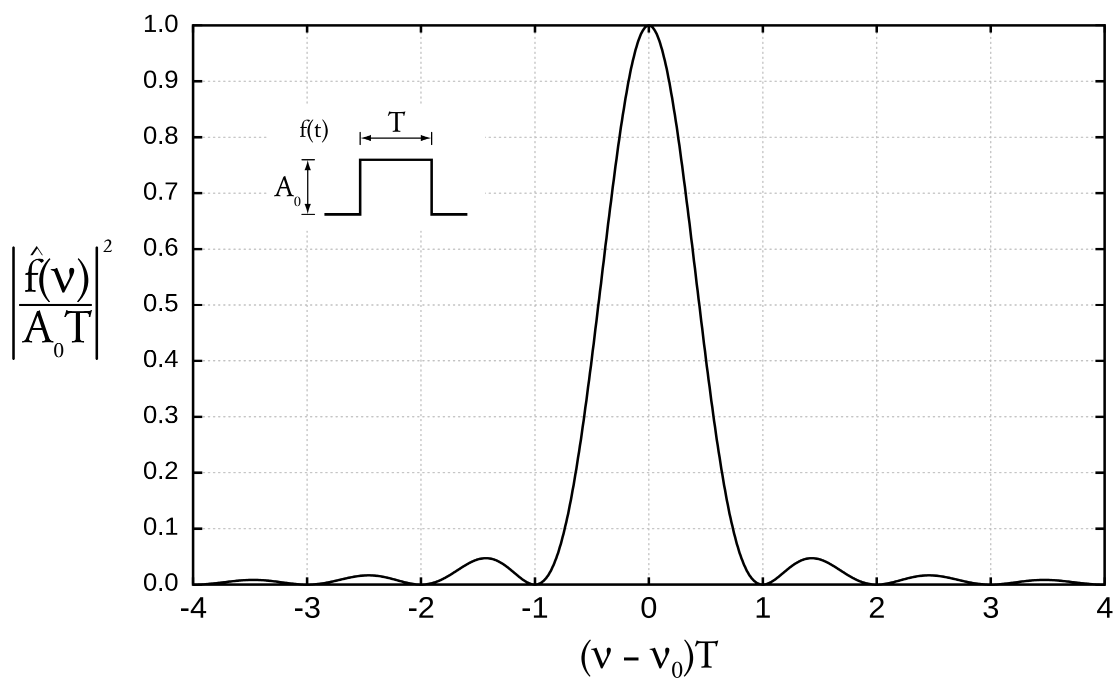


Figure 2.6 The spectrum of an OOK or PSK encoded optical signal has a characteristic sin(x)/x shape, centered about the carrier frequency. This figure plots the spectrum of an isolated positive bit of duration T and amplitude A_0.

* We have adopted the convention that sinc(x) denotes sin(x)/x. It should be noted, however, that some authors choose to define the sinc(x) function as sin(πx)/πx.

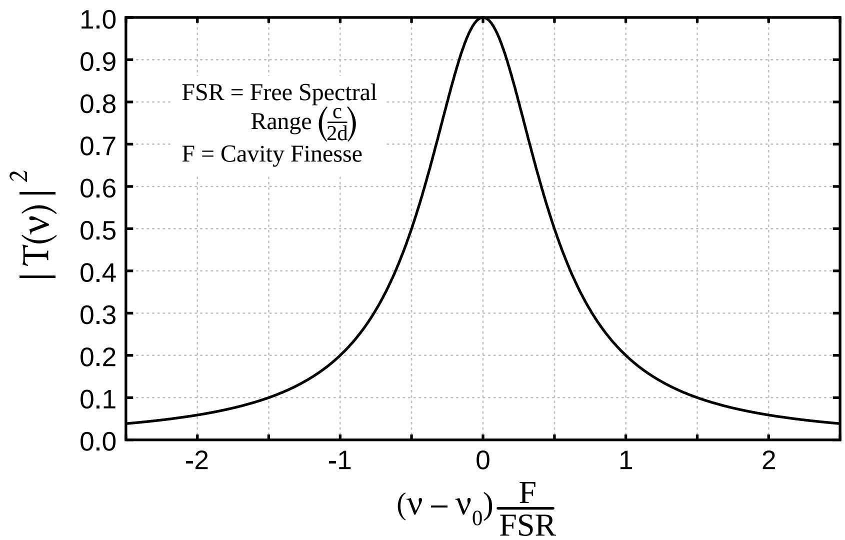


Figure 2.7 The power spectrum of a Fabry-Perot filter is described by a Lorentzian, as depicted in this figure. FSR represents the free-spectral range of the Fabry-Perot cavity, which is the frequency separation between subsequent transmission peaks. F is the cavity finesse, which is related to the mirror reflectivities. The Lorentzian FWHM depends on these two quantities, as shown.

pulses, the statistically averaged power spectrum of the random signal is identical to that of a single isolated pulse [7]. A similar analysis of PSK or DPSK signals reveals that they too have a characteristic sinc(x) spectral shape whose width is proportional to the communications bit rate, as with OOK signals.

An ideal filter for noise filtering in optical communications would have a spectral response identical to the response plotted in Fig. 2.6. For OOK or PSK systems, such a filter would have a characteristic sinc(x) shaped response, centered about the optical carrier frequency

2.3.3 The Alternative to Matched Optical Filters

In practice, it can be difficult to achieve a filter response that is precisely matched to the communications signal, and communications engineers must settle for one that resembles, but does not match, the signal spectrum. One common solution is to use a fiber Fabry-Perot filter, which is constructed from two polished and coated fiber facets separated by a gap, forming a resonator structure. Such a filter has a characteristic Lorentzian response, as shown in Fig. 2.7. The Lorentzian bandwidth is determined by the finesse and the free spectral range of the resonant cavity. The fiber Fabry-Perot filter has the advantages that it is easily integrated with other fiber-optic components, and it can be tuned by changing the size of the cavity. However, if the free spectral range is too small, a secondary broadband filter must be used to remove auxiliary resonance peaks in the transmission spectrum. A further restriction of Fabry-Perot filters is that the finesse is currently limited to about 1000 [8]. In addition to optical filtering, an electronic filter can also be used after the signals have been detected.

2.3.4 Benefits of an Integrated Optical Matched Filter

As described in Section 2.3, the benefit of a matched optical filter is that it provides the optimal spectral response for filtering a specified signal. Figure 2.8 presents some receiver sensitivity data reported by J. Livas at Lincoln Laboratories. [9, 10] The data points in Fig. 2.8 represent measured bit-error-rates for an optically preamplified receiver system operating at 10 Gb/s using OOK encoding. The results of Fig. 2.8 represent some of the most sensitive optical receiver results reported in the literature to date: in order to achieve a BER of 10^{-9}, an average signal level as low as 80 photons per bit is permissible. The solid line in Fig. 2.8 represents the theoretical limit on receiver sensitivity, determined solely by the amount of amplified spontaneous emission generated by the optical preamplifier, assuming a matched filter is used before detection [3]. The difference between the solid line and the reported data points in Fig. 2.8 can be attributed to non-optimal filtering of the optical signal before detection. As indicated, using a matched optical filter could potentially allow for a BER of 10^{-9} with only 38 photons per bit of average signal power. This represents an increase in receiver sensitivity by about a factor of 2. Similar improvements can be expected for PSK/DPSK systems. This improvement would mean that only half as much power is required to achieve the same receiver performance.

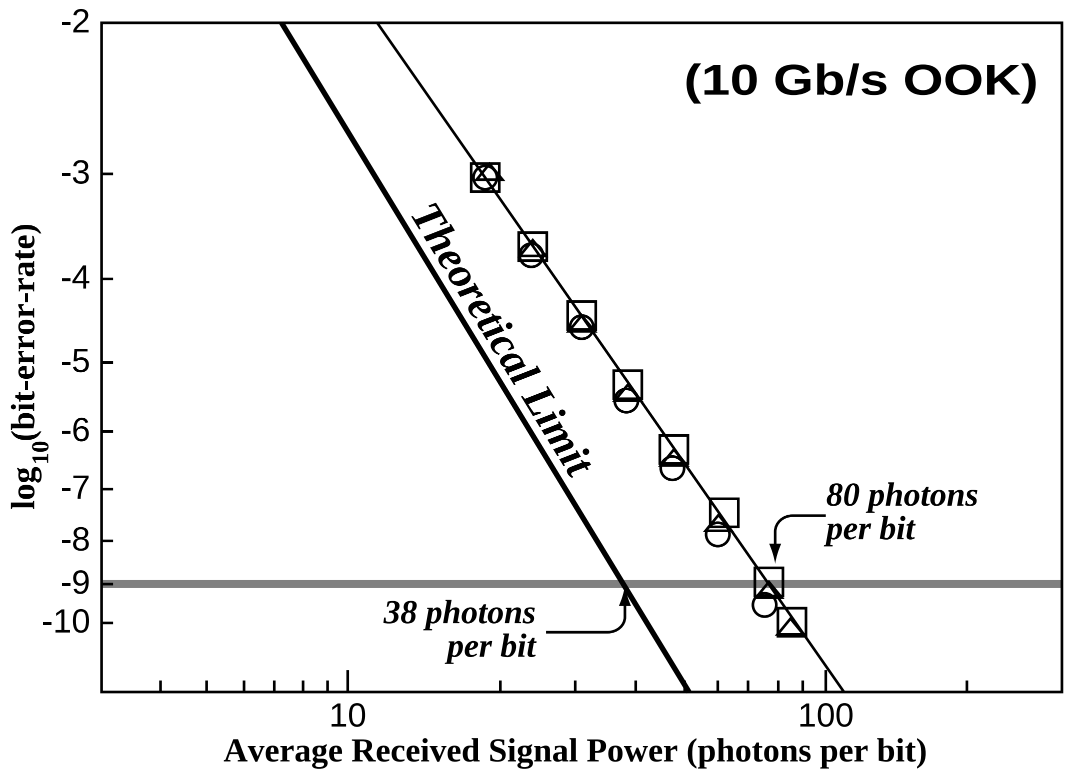


Figure 2.8 Bit-error-rate performance of an on-off keyed optical communications system operating at 10 Gb/s with an optically preamplified receiver [9]. The solid line indicates the theoretical upper limit to the receiver sensitivity for this type of system. The difference between these curves can be attributed to non-ideal filtering of the optical signal.

An integrated optical matched filter could have several advantages over conventionally used fiber Fabry-Perot filters. If the filter is fabricated on a chip, it could, in principle, be integrated with other optical or electronic components of the detector. The matched filters presented in this thesis could provide an integrated alternative to fiber Fabry-Perot filters and electronic post-detection filters.

2.4 Summary

Optical amplifiers have revolutionized optical communications and will likely replace more costly repeaters in future optical communications systems. However, the unavoidable presence of amplified spontaneous emission in optical amplifiers necessitates the use of narrow-band optical filters. A matched optical filter could play an important role in all-optical communications systems for filtering noise from an optical signal. The integrated optical matched filters described in the remainder of this thesis could allow for optical receivers with improved sensitivity, meaning less optical power would be required to attain the desired level of performance. Such an increase in receiver sensitivity could enable information providers to reach more customers or use longer fiber-optic cables without amplifiers or repeaters. Integrated optical matched filters could provide an advantage in performance over currently used fiber components, in a package that could be conveniently combined with other optical or electronic components of optical receivers.

CHAPTER 3: THE BRAGG GRATING FILTER

A *Bragg grating* can be loosely defined as a periodic structure that scatters radiation in a way that generates constructive interference. One of the simplest and earliest examples of such reflection is the diffraction of x-rays from the atomic lattice of a solid crystal. Because the atoms are assembled in a regular periodic array, the scattered x-rays from the lattice can constructively interfere at certain angles. The condition for constructive interference is that the path length difference between scattered light from two adjacent atoms in the lattice be an integral number of wavelengths [11]. The concept of constructive interference and wave propagation in periodic structures arises in a variety of physical situations including periodic antenna arrays, crystal diffraction, and even the quantum mechanical interaction of electrons with a semiconductor crystal. This work considers a very simple Bragg grating structure in which light is confined to propagate in only one dimension. The one dimensional Bragg grating is a periodic structure which coherently reflects light travelling in one direction into the opposite direction.

This chapter describes how a coherent grating patterned in a waveguide can act as a reflective filter. Coupled mode theory is used to analyze the effect of a weak grating patterned in a guiding optical structure. Using coupled mode theory, the reflection spectrum of a grating is derived, and it is shown that a properly designed grating can have a spectral response that matches that of a communications signal. Finally, we analyze the temporal response of the grating to an incident square pulse representing one bit of information to determine how a grating filter distorts the bits of an optical signal.

3.1 Contradirectional Coupling with a Grating

Optical communications systems typically employ guided wave devices where light is confined and directed in a dielectric structure rather than propagating freely in space. The topic of guided wave optics will be reviewed in Chapter 4; a complete understanding of guided wave optics is not necessary for this chapter. The principle of Bragg reflection can also apply to guided wave devices. If a periodic perturbation is introduced in the guiding dielectric medium, a portion of the incident light can be reflected into the counterpropagating mode. As with other forms of Bragg reflection, in order for the reflected light to constructively interfere, the path difference between reflected light from subsequent grating periods must be an integral number of wavelengths. Strongest reflection occurs when the path difference is precisely one wavelength, meaning the grating period must be one half of the wavelength:

$$\Lambda = \frac{\lambda_0}{2n_{eff}} \tag{3.1}$$

Here Λ is the grating period, λ_0 is the free-space wavelength, and λ_0/n_{eff} is the wavelength of light inside of the dielectric material. (n_{eff} is the effective index of refraction of the guiding dielectric structure.) The principle of reflection from a Bragg grating in a waveguide structure is illustrated schematically in Fig. 3.1. Reflections from subsequent teeth combine in phase to generate constructive interference in reflection.

Another way to analyze the constructive interference from a grating is to consider the problem in k-space, as illustrated in Fig. 3.2. The waveguide supports a forward and a backward guided mode, which have propagation constants $+\beta_0$ and $-\beta_0$ respectively. If a grating with wavevector $2\beta_0$ is added to the waveguide, it can couple the forward and backward travelling waves because the difference in propagation constant between the forward and backward travelling modes is precisely $2\beta_0$.

One way to construct a grating is to induce a periodic change in the index of refraction of the guiding dielectric structure. This technique is often used to construct gratings in optical fibers [12]. If the core of the fiber is composed of a photorefractive glass, then illuminating the fiber with a standing wave pattern generates a periodic modulation of the index of refraction of the fiber core. Another way to build a grating is to physically construct a geometric corrugation of the waveguide boundary, as shown in Fig. 3.1 [13].

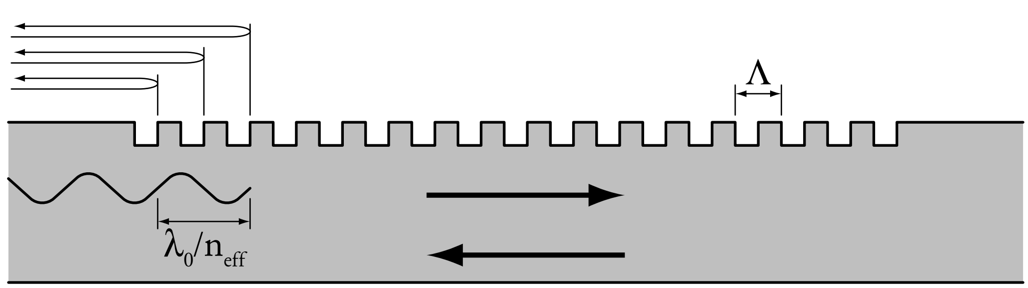


Figure 3.1 A Bragg grating constructed on the surface of an optical waveguide. Because the period of the grating is precisely half of the wavelength of light in the waveguide, the reflections from subsequent teeth in the grating generate constructive interference.

3.1.1 Coupled Mode Equations

The response of a grating can be calculated by treating the grating as a small perturbation of an otherwise normal waveguide. Rather than analyzing the reflection from each tooth of the grating separately, we instead treat the grating as a distributed reflector, parameterized by a grating strength, κ, which describes the rate at which energy is transferred between the otherwise decoupled forward and backward modes. In the absence of a grating, the propagation of light in a waveguide can be described in terms of the following equations:

$$\begin{aligned} \frac{d}{dz} a_+(z) &= +i\beta a_+(z) \\ \frac{d}{dz} a_-(z) &= -i\beta a_-(z) \end{aligned} \tag{3.2}$$

where $a_+(z)$ and $a_-(z)$ represent the mode amplitudes of the forward and backward travelling waves, respectively, and β is the propagation constant of the waveguide at a certain wavelength. The Cartesian coordinate axes have been oriented such that the waveguide points in the z direction, and an implicit time dependence of $e^{-i\omega t}$ has been factored out of all quantities. Notice that in the absence of a grating structure, the equations of motion for forward and backward travelling waves are completely decoupled. The solution to equations 3.2 is a pair of independent travelling waves in the +z and –z directions.

When a grating is added to the guiding structure the equations of motion for $a_+(z)$ and $a_-(z)$ are no longer decoupled. The modified coupled mode equations are:

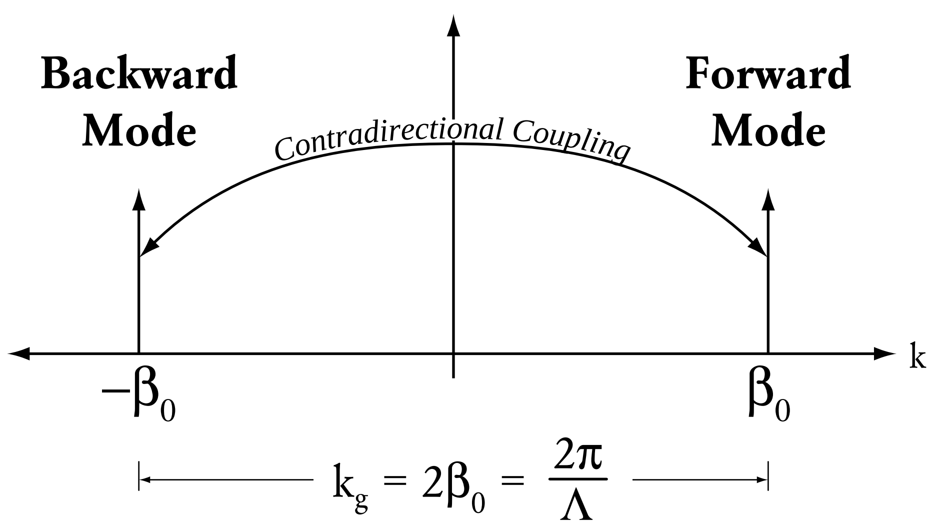


Figure 3.2 The presence of a grating in a waveguide structure can couple energy between the forward and backward travelling modes, provided the k-vector of the grating is exactly twice as large as the propagation constant of the light in the waveguide, as shown above.

$$\frac{d}{dz}a_+(z) = +i\beta a_+(z) + \kappa e^{+ik_g z}a_-(z)$$
$$\frac{d}{dz}a_-(z) = -i\beta a_-(z) + \kappa^* e^{-ik_g z}a_+(z) \tag{3.3}$$

The first terms in these equations are identical to those of equations 3.2, but a second term is present to represent the effect of the grating. In equation 3.3, κ is a quantity known as the grating strength, which is a measure of how much reflection is generated per unit length along the grating and k_g is the k-vector of the grating, described by:

$$k_g = \frac{2\pi}{\Lambda} \tag{3.4}$$

A mathematical derivation of the coupled mode equations 3.3 will be presented in Chapter 4. This chapter seeks only to describe the physical nature of these equations and to examine their solutions.

3.1.2 Solution of Coupled Mode Equations

In solving the coupled mode equations, it is useful to factor out the rapidly oscillating components from the mode amplitudes $a_+(z)$ and $a_-(z)$. The rapid oscillations are removed by making the following change of variables:

$$A_+(z) = a_+(z)e^{-\frac{i}{2}k_g z}$$
$$A_-(z) = a_-(z)e^{+\frac{i}{2}k_g z} \tag{3.5}$$

$A_+(z)$ and $A_-(z)$ represent slowly varying mode envelope functions, after the rapid optical oscillations have been factored out. After making this substitution, the coupled mode equations 3.3 simplify to:

$$\frac{d}{dz}A_+(z) = +i\delta A_+(z) + \kappa A_-(z)$$
$$\frac{d}{dz}A_-(z) = -i\delta A_-(z) + \kappa^* A_+(z) \tag{3.6}$$

where δ is a measure of the deviation from the Bragg condition, given by

$$\delta \equiv \beta - \frac{1}{2}k_g = \beta - \frac{\pi}{\Lambda} \tag{3.7}$$

The coupled mode equations for $A_+(z)$ and $A_-(z)$ can be written in the convenient matrix form:

$$\frac{d}{dz}\begin{bmatrix} A_+(z) \\ A_-(z) \end{bmatrix} = \begin{bmatrix} +i\delta & \kappa \\ \kappa^* & -i\delta \end{bmatrix}\begin{bmatrix} A_+(z) \\ A_-(z) \end{bmatrix} \tag{3.8}$$

Notice that after changing variables to $A_+(z)$ and $A_-(z)$, the resulting equations of motion comprise a system of coupled linear differential equations. Equation 3.8 is a linear vector differential equation which can be solved in the conventional way by computing the eigenvectors and eigenvalues of the system of equations. Using this method of eigenvector decomposition, the solution is:

$$\begin{bmatrix} A_+(z) \\ A_-(z) \end{bmatrix} = \begin{bmatrix} \mathbf{v_1} & \mathbf{v_2} \end{bmatrix} \begin{bmatrix} e^{\lambda_1 z} & 0 \\ 0 & e^{\lambda 2z} \end{bmatrix} \begin{bmatrix} \mathbf{v_1} & \mathbf{v_2} \end{bmatrix}^{-1} \begin{bmatrix} A_+(0) \\ A_-(0) \end{bmatrix} \tag{3.9}$$

where $\mathbf{v_1}$ and $\mathbf{v_2}$ are the eigenvectors the of the system and λ_1 and λ_2 are the corresponding eigenvalues. The eigenvalues and associated eigenvectors are given by:

$$\begin{aligned} \lambda_1 &= +\gamma \\ \lambda_1 &= -\gamma \end{aligned} \tag{3.10}$$

$$\begin{aligned} \mathbf{v_1} &= \begin{bmatrix} i\delta + \gamma \\ \kappa^* \end{bmatrix} \\ \mathbf{v_2} &= \begin{bmatrix} -\kappa \\ i\delta + \gamma \end{bmatrix} \end{aligned} \tag{3.11}$$

where we have defined the quantity γ as:

$$\gamma \equiv \sqrt{|\kappa|^2 - \delta^2} \tag{3.12}$$

Substituting the eigenvalues and eigenvectors from equations 3.10 and 3.11 into the vector solution in equation 3.9, the following simplified solution for the mode amplitudes is obtained:

$$\begin{bmatrix} A_+(z) \\ A_-(z) \end{bmatrix} = \begin{bmatrix} \cosh(\gamma z) + i\dfrac{\delta}{\gamma}\sinh(\gamma z) & \dfrac{\kappa}{\gamma}\sinh(\gamma z) \\ \dfrac{\kappa^*}{\gamma}\sinh(\gamma z) & \cosh(\gamma z) - i\dfrac{\delta}{\gamma}\sinh(\gamma z) \end{bmatrix} \begin{bmatrix} A_+(0) \\ A_-(0) \end{bmatrix} \tag{3.13}$$

This solution can then be written in terms of the original mode amplitude quantities, $a_+(z)$ and $a_-(z)$, by using the relationships given in equations 3.5:

$$\begin{bmatrix} a_+(z) \\ a_-(z) \end{bmatrix} = \begin{bmatrix} e^{+\frac{i}{2}k_g z} & 0 \\ 0 & e^{-\frac{i}{2}k_g z} \end{bmatrix} \begin{bmatrix} \cosh(\gamma z) + i\dfrac{\delta}{\gamma}\sinh(\gamma z) & \dfrac{\kappa}{\gamma}\sinh(\gamma z) \\ \dfrac{\kappa^*}{\gamma}\sinh(\gamma z) & \cosh(\gamma z) - i\dfrac{\delta}{\gamma}\sinh(\gamma z) \end{bmatrix} \begin{bmatrix} a_+(0) \\ a_-(0) \end{bmatrix} \tag{3.14}$$

3.2 Reflection Spectral Response of Bragg Grating

Equation 3.14 is a closed-form solution to the coupled mode equations, which gives the mode amplitudes at any point z, given the initial forward and backward mode amplitudes at z = 0. Often, the boundary conditions on a_+ and a_- at z = 0 are not simultaneously known. Consider a grating extending from z = 0 to L, with some signal incident from the left. The two boundary conditions for this problem are that no signal is incident from the right hand side of the grating ($a_-(L) = 0$), and that some known signal is incident from the left ($a_+(0) = 1$.) From these two boundary conditions we wish to derive the transmission and reflection coefficients for the grating segment. Substituting these boundary conditions into the matrix equation 3.14 gives:

$$\begin{bmatrix} t \\ 0 \end{bmatrix} = \begin{bmatrix} e^{+\frac{i}{2}k_g L} & 0 \\ 0 & e^{-\frac{i}{2}k_g L} \end{bmatrix} \begin{bmatrix} \cosh(\gamma L) + i\frac{\delta}{\gamma}\sinh(\gamma L) & \frac{\kappa}{\gamma}\sinh(\gamma L) \\ \frac{\kappa^*}{\gamma}\sinh(\gamma L) & \cosh(\gamma L) - i\frac{\delta}{\gamma}\sinh(\gamma L) \end{bmatrix} \begin{bmatrix} 1 \\ r \end{bmatrix} \tag{3.15}$$

where r and t represent the amplitude reflection and transmission coefficients of the grating, respectively. Equation 3.15 represents two linear equations with two unknown quantities, r and t. It is straightforward to solve for $r(\delta)$ from the second equation:

$$r(\delta) = \frac{\frac{\kappa^*}{\gamma}\tanh(\gamma L)}{1 - i\frac{\delta}{\gamma}\tanh(\gamma L)} \tag{3.16}$$

The quantities γ and δ are defined in equations 3.7 and 3.12. The case of $\delta = 0$ corresponds to an incident signal whose wavelength exactly meets the Bragg condition. When $\delta = 0$, the magnitude of the reflection coefficient simplifies to:

$$|r(\delta = 0)|^2 = \tanh^2(|\kappa|L) \tag{3.17}$$

Figure 3.3 plots the reflection spectrum $|r(\delta)|^2$ of a Bragg grating for several different κL values.

3.2.1 Strong Grating

When the κL product is substantially larger than 1, the reflection spectrum is characterized by a plateau shaped response, centered about the Bragg frequency ($\delta = 0$). Within a band of frequencies centered around the Bragg frequency, essentially all of the incident signal is reflected by the grating. This band is usually referred to as the stop-band of the grating. The spectral width of the stop-band is proportional to the grating strength κ. The edges of the stopband are given by $\delta = \pm\kappa$. Outside the stopband, the spectral response exhibits decaying oscillations away from the Bragg frequency.

3.2.2 Weak Grating

When the κL product is smaller than 1, the reflection spectrum is characterized by a nearly sinc shaped response, centered at the Bragg frequency. This result can be derived by expanding equation 3.16 to lowest order in the quantity κ/δ. A complete Taylor expansion of the reflection spectrum in this limit is somewhat involved, so only the final result is presented here:

$$r(\delta) \approx -\frac{\kappa^*}{\delta} e^{i\delta L} \sin(\delta L) \tag{3.18}$$

$$\frac{|\kappa|}{\delta} << 1 \tag{3.19}$$

Thus, under certain conditions, the reflection spectrum of a Bragg grating has a characteristic $\sin(\delta L)/\delta$ shape, with a linear phase term. The expression given in equation 3.18 is valid even for strong Bragg gratings ($\kappa L > 1$), at points far outside of the stopband where equation 3.19 is satisfied. Although it is not obvious at first glance, for weak Bragg gratings (with $\kappa L < 1$), equation 3.18 is a good approximation

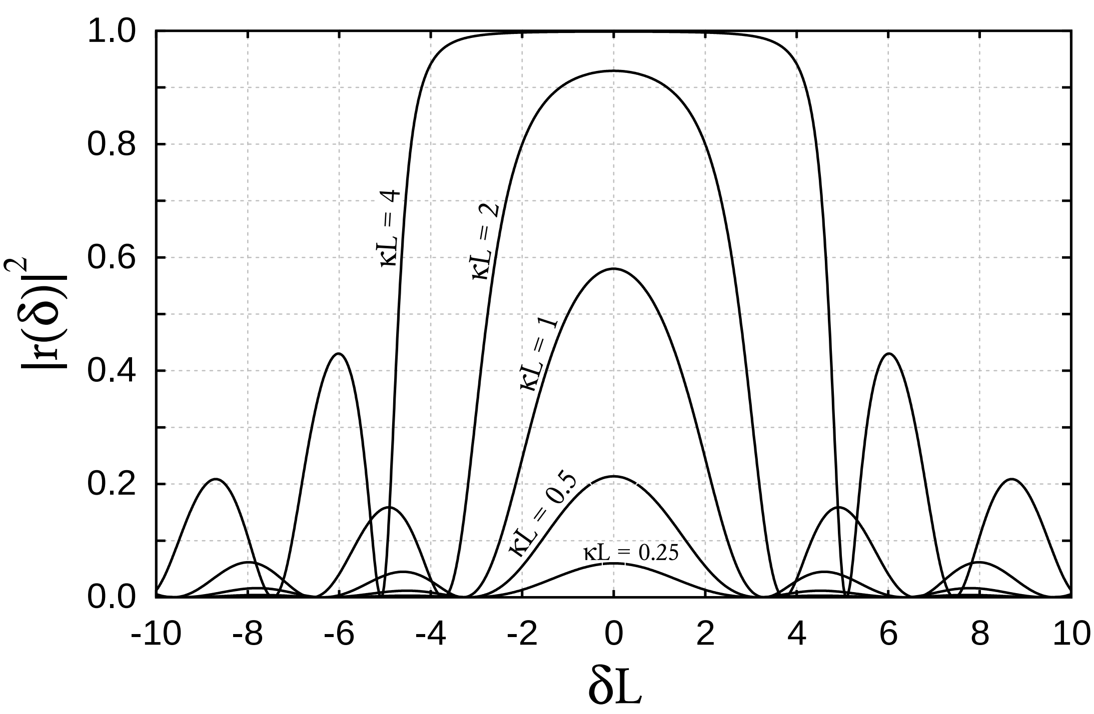


Figure 3.3 Reflection spectra from Bragg gratings with five different κL values. Plotted along the abscissa is the normalized deviation from the Bragg condition. Notice that for small κL values, the spectral response has a sinc-shaped response, whereas for large κL, the response has a plateau shape.

to the exact reflection spectrum over the *entire* spectral range, even when equation 3.19 is not satisfied, as can be seen in the spectra plotted in Fig. 3.3. Moreover, the spectral response can be arbitrarily close to the $\sin(\delta L)/\delta$ shape given in equation 3.18, provided the κL product can be made arbitrarily small. Figure 3.4 is a plot which compares the exact spectral response of a grating (described by equation 3.16) with the approximate spectral response given by equation 3.18, for the case of $\kappa L = 0.4$.

The asymptotic spectral response given in equation 3.18 is identical in form to the spectrum of a binary sequence of data, derived in Chapter 2. Accordingly, a weak Bragg grating can have a reflection spectral response that closely matches the spectrum of an incident binary signal, forming a matched filter. In the weak-grating limit, the width of the spectral response is determined only by the length of the grating. The peak reflectivity of the grating is determined by the product κL. Therefore, to construct a matched filter for a particular communications data rate, one must first select the grating length L so that the spectral width matches that of the incident data signal. The grating strength κ must then be selected so that the product κL is less than 1, otherwise the reflection spectrum will not have the desired $\sin(\delta L)/\delta$ shape over the entire spectral range. Thus, the communications data rate of the signal to be matched uniquely determines the grating length L, whereas the grating strength κ (or more specifically the product κL) determines how closely the spectra are matched.

The relationship between the communications bit rate and the grating length L can be found by comparing the spectrum derived here with that found in Chapter 2. In order for the grating to be

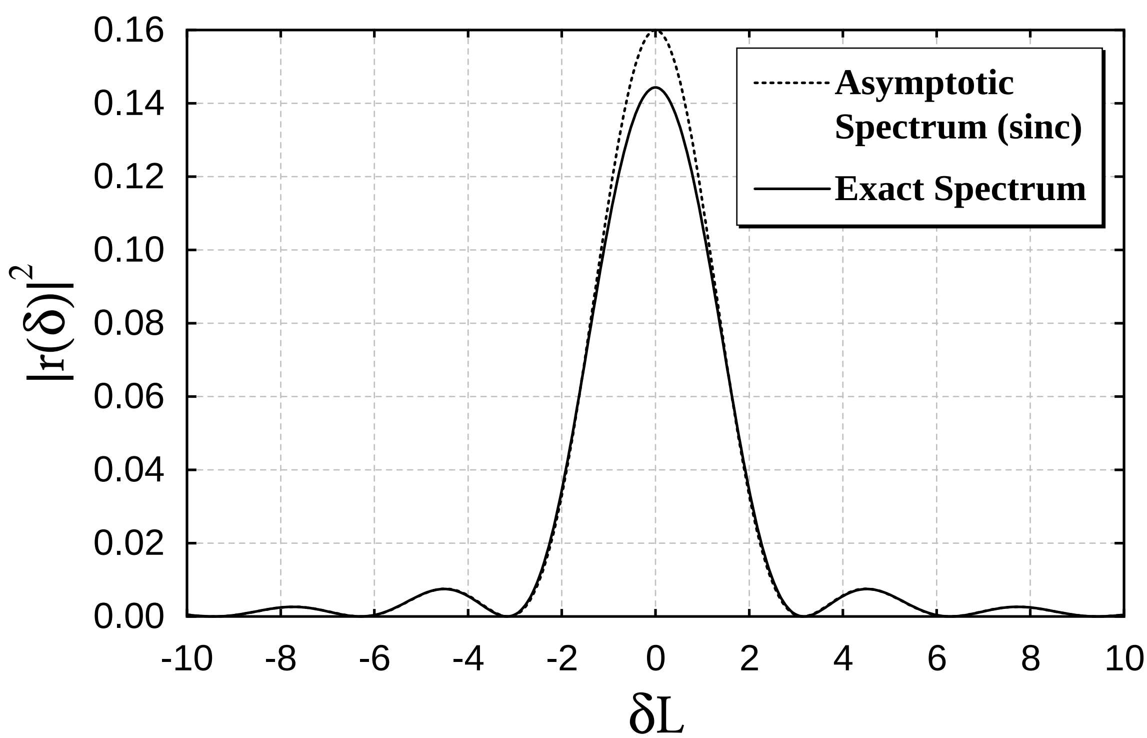


Figure 3.4 A comparison of the exact reflection spectrum from a Bragg grating with $\kappa L = 0.4$ and the asymptotic response obtained using the weak-grating approximation described in this chapter.

matched to the spectrum of an OOK or PSK digital signal, the length of the grating L must be selected to be exactly one half of the length associated with a single bit of information. Another way to state this relationship is that the down and back propagation time for a wave through the grating must be equal to one bit duration.

If the grating strength κ can be made arbitrarily small, the spectral response can be made arbitrarily close to the asymptotic $\sin(\delta L)/\delta$ limit. However, as indicated in equation 3.17, decreasing the κL product also decreases the peak reflectivity of the grating. If the grating is very weak ($\kappa L << 1$), then even though its spectral response is closely matched to the incident signal, only a very small fraction of the signal will be reflected by the grating while most of the signal is transmitted through the grating. The amplitude of the reflection response does not effect whether the filter is matched to the signal; it is only the *shape* of the reflection response which determines the degree of improvement in signal to noise [7]. After all, a weak reflection will attenuate both the signal and the background noise. However, if the resulting filtered signal is *very* weak, it can be susceptible to other sources of noise in the receiver. In designing such a matched filter for an optical receiver, one must consider not only how closely matched the spectra are, but also how much signal loss can be tolerated in the receiver before other sources of noise begin to contribute.

The $\sin(\delta L)/\delta$ response of a weak grating can be understood physically by considering the grating as a single pass filter. If the grating is very weak, then multiple reflections within the grating can be neglected to first order and the grating can be treated as a one-pass distributed reflector. For a single-pass distributed reflector, the reflection spectral response is proportional to the Fourier transform of the grating pattern. Because the grating is uniform for a finite length and absent outside of that length, the Fourier transform of the grating pattern is expected to have a characteristic sinc-shaped response.

3.3 Temporal Response of Bragg Grating

In analyzing the performance an optical filter's performance, it is important to consider not only the spectral response of the filter but also the temporal response. The performance of a filter often depends directly on how the shapes of individual bits of data are altered by the filter. Any filter which does not have a uniform spectral response will necessarily distort the shapes of the pulses in the data stream. If this pulse distortion is severe, it can become difficult for the receiver to distinguish adjacent bits of information, or the data in one bit slot can interfere with the neighboring bits. In this section we examine the temporal response of a Bragg grating filter, and compare its performance with other commonly used filters.

3.3.1 Calculation of Temporal Response by Fourier Transform

In calculating the temporal response of a filter, we will assume that the input and output data signal can be represented as a product of a slowly varying amplitude function which modulates an otherwise uniform rapidly oscillating optical carrier signal:

$$
\begin{aligned}
f_{IN}(z,t) &= e^{i(\beta_0 z - \omega_0 t)} f'_{IN}(z - v_g t) \\
f_{OUT}(z,t) &= e^{i(-\beta_0 z - \omega_0 t)} f'_{OUT}(z + v_g t)
\end{aligned}
\tag{3.20}
$$

In the above equations, the primed quantities denote the slowly varying wave amplitude envelopes, ω_0 represents the optical carrier frequency, β_0 is the corresponding propagation constant inside of the waveguide, and v_g is the group velocity of the signal, evaluated at the carrier frequency. f'_{IN} represents an input signal which is travelling in the positive z direction. f'_{OUT} represents a reflected output signal which is travelling in the negative z direction. In order to study how a matched filter distorts a digital optical signal, we let $f'_{IN}(z)$ represent a single isolated pulse of an On-Off-Keyed data sequence:

$$
f'_{IN}(z) = \begin{cases} A_0 & -L_b < z < 0 \\ 0 & \text{otherwise} \end{cases}
\tag{3.21}
$$

Equations 3.20 and 3.21 describe a single isolated bit of length L_b and amplitude A_0 travelling with the group velocity in the positive z direction. The leading edge of this pulse reaches the position z = 0 at time t = 0. This pulse is represented schematically in Fig. 3.5. In order to analyze the effect of a filter, we must determine the spectrum of the incident pulse, denoted $\Phi(\delta)$.

$$
\begin{aligned}
f'_{IN}(z) &= \int_{-\infty}^{\infty} \Phi(\delta) e^{+i\delta z} d\delta \\
\Phi(\delta) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f'_{IN}(z) e^{-i\delta z} dz
\end{aligned}
\tag{3.22}
$$

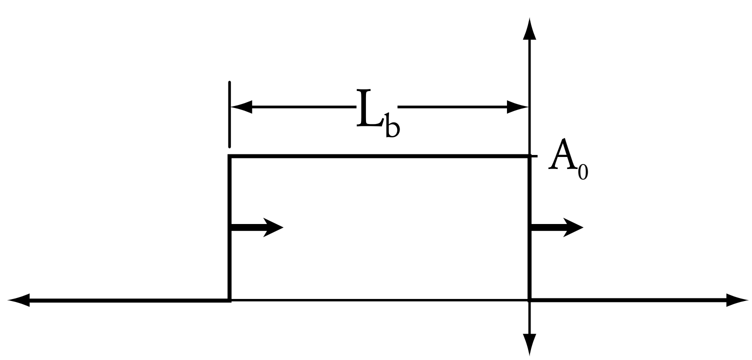


Figure 3.5 Amplitude profile of a square pulse of length L_b and height A_0, travelling to the right. The leading edge of the pulse reaches the position z = 0 at time t = 0.

In the above equation δ represents the deviation from β_0, the wavenumber of the optical carrier signal. In equation 3.22, we have expressed the square pulse amplitude modulation in terms of its spatial frequency components, using the Fourier transform relations. Applying the Fourier transform given in equation 3.22 to the pulse shape given in equation 3.21 yields the following pulse spectrum:

$$\Phi(\delta) = \frac{A_0}{\pi\delta} e^{i\frac{\delta L_b}{2}} \sin\left(\frac{\delta L_b}{2}\right) \tag{3.23}$$

Equation 3.23 (which was earlier derived in Chapter 2) reveals that the isolated square pulse has a characteristic sinc-shaped spectrum, centered about the carrier frequency ($\delta = 0$).

MATCHED FILTER

It is clear by direct comparison of equation 3.23 and equation 3.18 that in order for a Bragg grating to be matched to the spectrum of a pulse of length L_b, the following conditions must hold:

$$\begin{aligned} \beta_0 &= \frac{1}{2}k_g = \frac{\pi}{\Lambda} \\ L_0 &= 2L \end{aligned} \tag{3.24}$$

The first equation states that the Bragg frequency of the grating must be the same as the carrier frequency of the incident optical square pulse. The second equation states that the length of the grating L must be one half of the length of the optical pulse in the waveguide (L_b).

Provided the conditions given in equation 3.24 hold, the pulse spectrum $\Phi(\delta)$ can be rewritten in terms of the grating length L:

$$\Phi(\delta) = \frac{A_0}{\pi\delta} e^{i\delta L} \sin(\delta L) \tag{3.25}$$

When the pulse spectrum is cast into the above form, the only difference between the pulse spectrum and the asymptotic reflection spectrum of the grating given in equation 3.18 is a proportionality constant.

The spectrum of the reflected pulse from the grating is determined by multiplying the spectrum of the incident signal (equation 3.25) by the reflection spectrum of the grating (equation 3.18). The shape of the reflected pulse is then found by computing the inverse Fourier transform:

$$f'_{OUT}(z) = \int_{-\infty}^{\infty} \Phi(\delta) r(\delta) e^{+i\delta z} d\delta \tag{3.26}$$

Substituting equations 3.25 and 3.18 into equation 3.26 and evaluating the integral gives:

$$f'_{OUT}(z) = -\kappa^* L A_0 \hat{T}_{2L}(z - 2L) \tag{3.27}$$

where $\hat{T}_{2L}(z)$ denotes a unit triangle function which has a maximum value of 1 at $z = 0$ and falls linearly to zero at $z = \pm 2L$. Figure 3.6 is a plot of the reflected pulse amplitude from a matched grating filter. Note that this plot shows only the amplitude modulation envelope, and does not attempt to depict the rapid oscillations of the optical carrier frequency. In fact there would be of order 10^4 such optical oscillations per pulse. The reflected signal from the grating has a triangular shape which is twice as wide as the original pulse, centered at $z = 2L$, and travelling in the negative z direction at the group velocity. The peak amplitude of the reflected pulse is reduced from the amplitude of the incident pulse by a factor of κL, which is consistent with the result that for a weak grating the peak reflectivity is κL.

The pulse shape $f'_{OUT}(z)$ given in equation 3.27 and plotted in Fig. 3.6 describes the shape and position of the reflected pulse at time $t = 0$. Of course, at time $t = 0$, the leading edge of the incident square pulse has just reached the left edge of the grating at $z = 0$. Therefore at $t = 0$, a complete reflected pulse has not yet formed, and it is certainly unreasonable for the reflected pulse shape depicted in Fig. 3.6 to instantaneously exist inside of the grating region. The explanation of this apparent contradiction lies in the fact that the calculated reflected pulse shape does not describe the mode amplitude within the grating segment, but only the mode amplitude in the region to the left of the grating segment. Therefore, the calculated pulse response of the grating does not indicate the state of the electromagnetic fields inside of the grating but rather the shape of the pulse which is about to emerge from the left edge of the

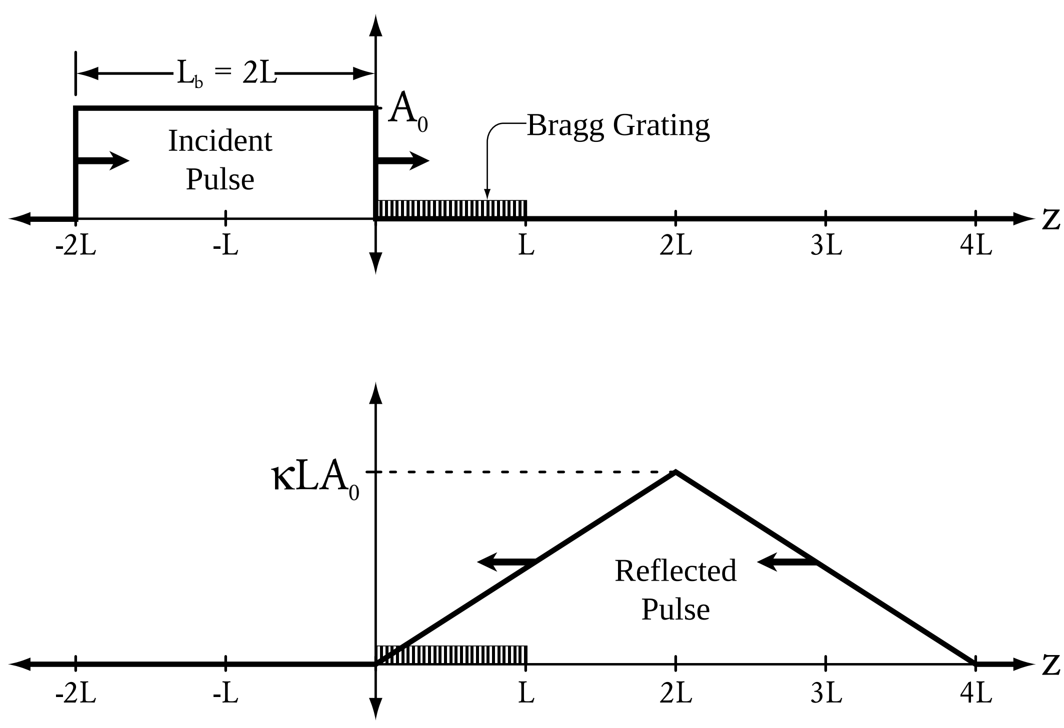


Figure 3.6 Reflected pulse from a weak Bragg grating, assuming the reflection spectrum of the grating is perfectly matched to the spectrum of the incident pulse. The reflected pulse shape is triangular, and twice as wide as the incident pulse.

grating. From the point-of-view of an observer located to the left of the grating, it would appear as if a backward travelling triangular pulse is emerging from the edge of the grating at z = 0. The leading edge of the reflected triangular pulse begins at z = 0, indicating that a reflected signal begins to emerge at the instant that the square pulse begins to strike the grating from the left.

The triangular shape of the reflected pulse can be understood as the convolution of two identical square pulses in space: computing the product of two identical spectra in the frequency domain is equivalent to convolving the signal with itself in the time domain.

WEAK GRATING FILTER

The time response given in the previous section was derived assuming that the spectral response of the Bragg grating is accurately described by the asymptotic limit given in equation 3.18. As described earlier, the validity of equation 3.18 depends on the grating strength, and for very weak gratings with $\kappa L \ll 1$ where equation 3.18 is accurate, the reflection response is impractically small. In order to determine the temporal response of the grating for gratings with intermediate strength, it is necessary to use the more exact expression for the reflection spectrum, given in equation 3.16. Because of the complicated form of equation 3.16, analytically evaluating the inverse transform integral of equation 3.26 is not tractable and numerical techniques must be used instead. The inverse transform can be accurately and efficiently calculated by applying the discrete Fourier transform to a sampled spectrum representing the reflected pulse.

Figure 3.7 plots the reflection response for a somewhat weak grating, with $\kappa L = 0.4$. Again, for the calculation presented in Fig. 3.7 it is assumed that a pulse of length 2L is incident on the grating of length L from the left.

STRONG GRATING FILTER

For a strong grating filter, with $\kappa L > 1$, the spectral response of the grating can no longer be accurately represented by the asymptotic expression given in equation 3.18, and again numerical techniques must be used to compute the shape of the reflected pulse. Figure 3.8 plots the numerically calculated reflected pulse shape for a strong grating with $\kappa L = 4$. Notice that the reflected signal from a strong grating is characterized by oscillations caused by multiple reflections of the signal within the grating.

FABRY-PEROT FILTER

As described in Chapter 2, another commonly used filter in optical communications is the fiber Fabry-Perot filter, which is constructed by creating a small Fabry-Perot cavity between polished and coated optical fiber facets. A Fabry-Perot cavity acts as a transmission filter with narrow, regularly spaced

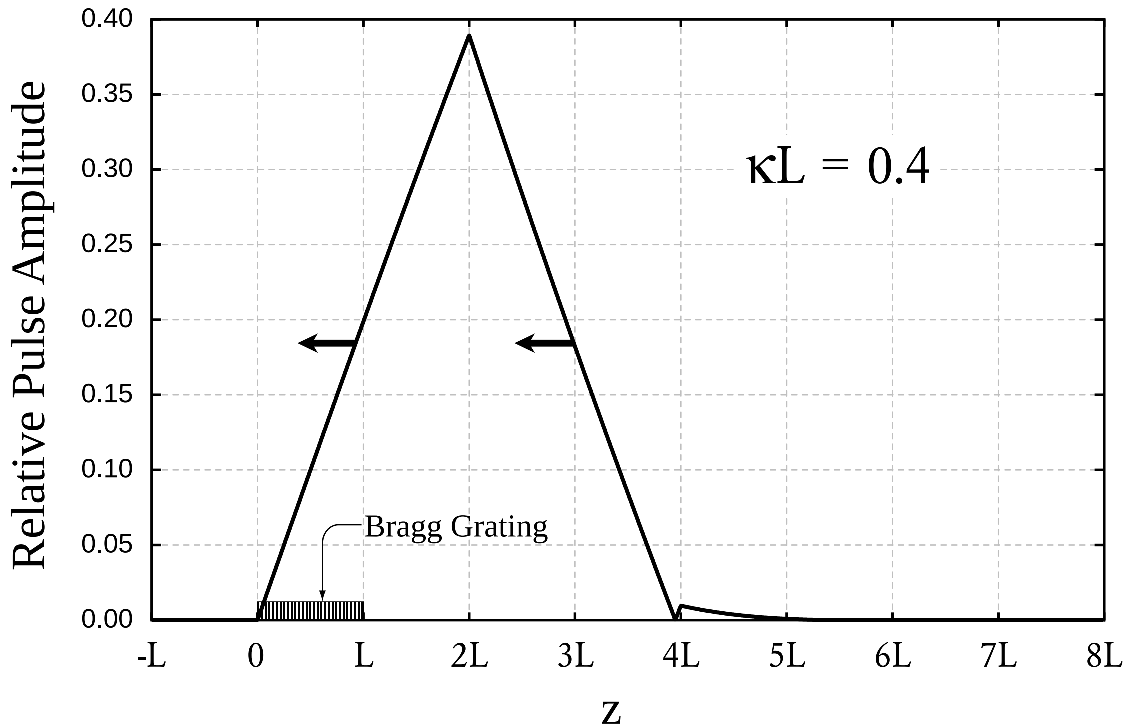


Figure 3.7 Reflected pulse shape from a somewhat weak grating, with κL = 0.4. Note that the reflected pulse closely resembles the triangular shape characteristic of a perfectly matched filter.

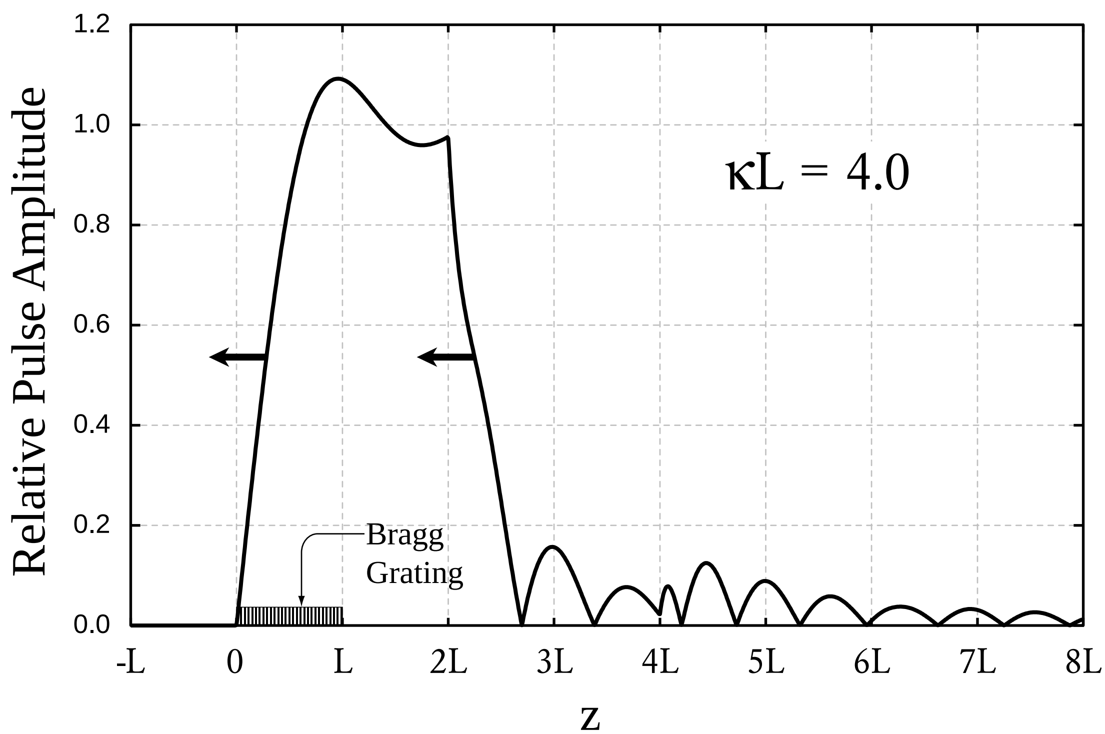


Figure 3.8 Reflected pulse shape from a strong grating with κL = 4.0. The rapid oscillations which extend well into the next bit slot are caused by multiple reflections within the Bragg grating.

transmission peaks which characterize the longitudinal modes of the cavity. The separation between adjacent transmission peaks is known as the free spectral range (FSR) and is determined by the condition that an integral number of half-wavelengths must fit exactly within the cavity. Each transmission peak can be well approximated by a Lorentzian spectrum, whose width is related to the mirror reflectivities [14]. The ratio of the free spectral range to the individual peak bandwidth width is known as the finesse of the cavity. The Lorentzian transmission spectrum for a Fabry-Perot filter can be approximated near resonance as:

$$t(\delta) = \frac{\alpha}{\alpha + i\delta} \tag{3.28}$$

As with the Bragg grating filter, we express the spectral response of the Fabry-Perot filter in terms of the deviation in spatial frequency from the central value, thus δ represents the deviation in spatial frequency from the resonance peak of the cavity. (For the Bragg grating, the central value corresponds to the Bragg frequency of the grating, whereas for the Fabry Perot cavity the central value corresponds to a resonance of the cavity.) The parameter α in equation 3.28 depends only on the parameters of the Fabry-Perot cavity:

$$\alpha \equiv \frac{\pi}{2Fd} \tag{3.29}$$

In the above equation, d represents the length of the Fabry-Perot cavity and F is the finesse of the cavity. Equation 3.28 is the same spectral response which was plotted earlier in Chapter 2, however it is now expressed in terms of its spatial frequency, δ. Figure 3.9 is a plot of the spectral response of a Fabry-Perot filter. Following the same procedure described earlier for a grating filter, the pulse response of the Fabry-Perot filter can be calculated using the Fourier transform, which yields the following resultant pulse shape:

$$f'_{OUT}(z) = \begin{cases} 0 & 0 < z < \infty \\ A_0\left(1 - e^{\alpha z}\right) & -L_b < z < 0 \\ A_0\left(1 - e^{-\alpha l}\right)e^{\alpha(z + L_b)} & -\infty < z < -L_b \end{cases} \tag{3.30}$$

In this equation, $f'_{OUT}(z)$ represents a forward travelling amplitude modulation signal which has been transmitted through a Fabry-Perot cavity. Notice that the pulse has an exponential response which is characterized by a decay rate α. Such a response is characteristic of a single-pole filter such as a Fabry-Perot cavity. Figure 3.10 plots a the transmitted pulse from a Fabry-Perot filter, for the case where αL_b = 1. Because the Fabry-Perot filter is a transmission filter, this pulse should be understood as travelling in the positive z direction

Clearly, the Fabry-Perot filter has some disadvantages when compared with a weak grating filter. The transmission spectrum of the filter is limited to a Lorentzian response, which cannot match the

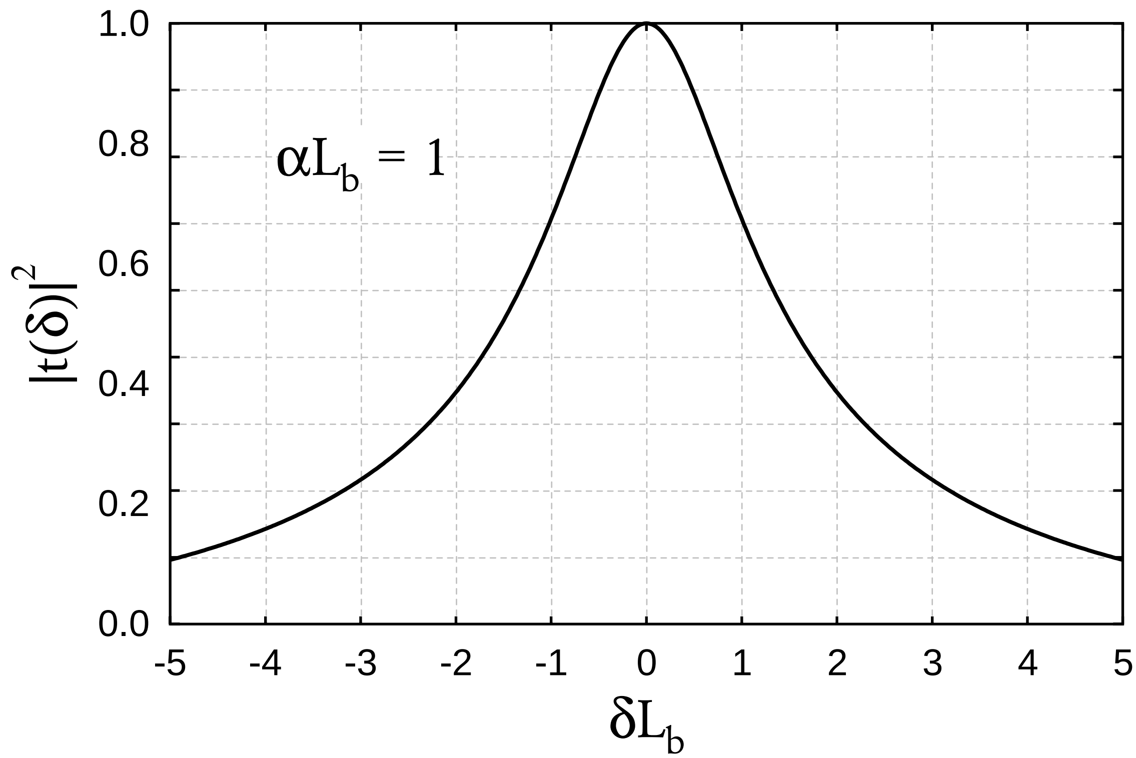


Figure 3.9 Spectral response of a Fabry-Perot filter. The spectrum has a lorentzian shape whose width depends on reflectivity of the cavity facets. The abscissa has been normalized in a way that makes it directly comparable with the grating spectra presented in Fig. 3.3.

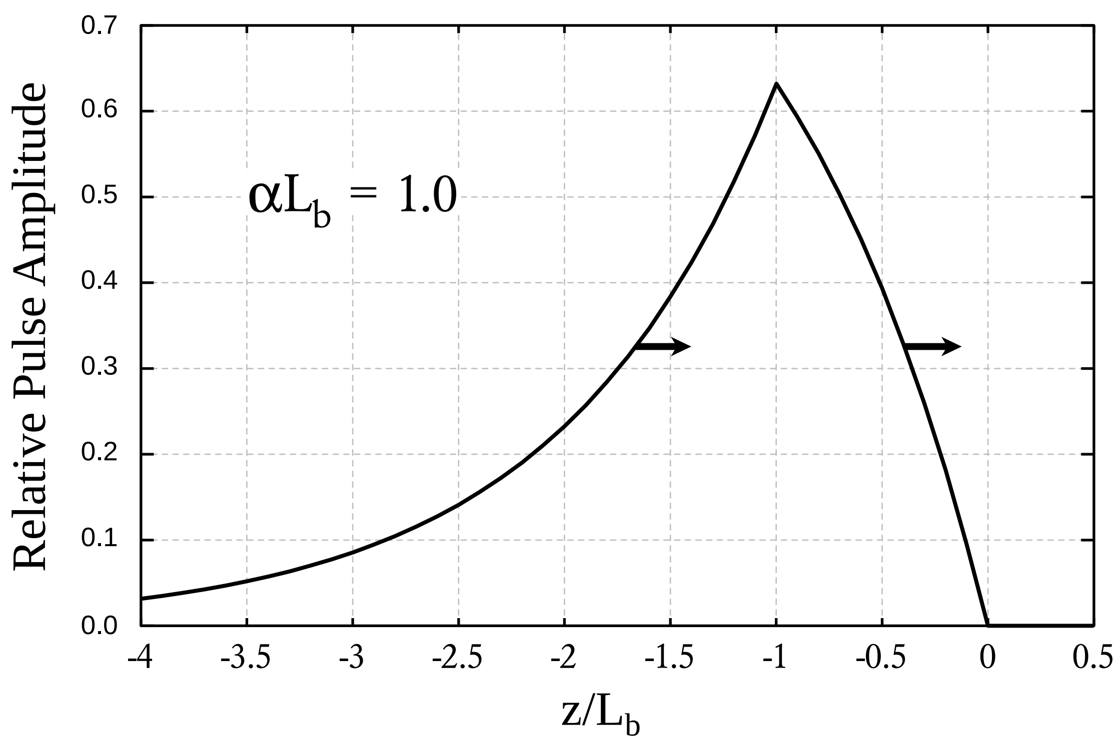


Figure 3.10 Transmitted pulse from a Fabry-Perot filter. The incident pulse of length L_b is illustrated in Fig. 3.5. The long trailing edge of the filtered pulse decays exponentially with decay constant α.

spectrum of a communications signal. Also, the transmission response a Fabry-Perot cavity contains many identical equally spaced transmission peaks corresponding the resonant frequencies of the cavity. This often necessitates the use of a second broad-band filter to remove all transmission peaks other than the one of interest. However, the fiber Fabry-Perot filter is an attractive and commonly used component because it is easily integrated into existing fiber-optic systems with low insertion loss. Additionally, the Fabry-Perot filter can be easily tuned to a specific resonant frequency by changing the length of the cavity.

3.3.2 Intersymbol Interference

The emerging pulse from a Fabry-Perot transmission filter has a long exponentially decaying tail which extends well beyond the timeslot allocated for the bit. In a typical communications stream, this exponential tail would interfere with subsequent bits in the signal, possibly leading to errors in detection. This effect is known as intersymbol interference. Intersymbol interference is a measure of how a given bit of data effects the subsequent bits in the data stream [3].

It would appear that the matched filter described in Section 3.3.1 also suffers from intersymbol interference, because the resulting triangular pulses which emerge from the Bragg grating are twice as wide as the incident square pulses. Thus, there is some degree of overlap between adjacent bits (or symbols) in a data sequence. However, if the detector consists of a device which samples the pulses at their peak values, then at the instant when a pulse reaches its peak value, the previous pulse amplitude has just decayed to zero. Figure 3.11 illustrates two adjacent binary pulses reflected from an ideal matched Bragg

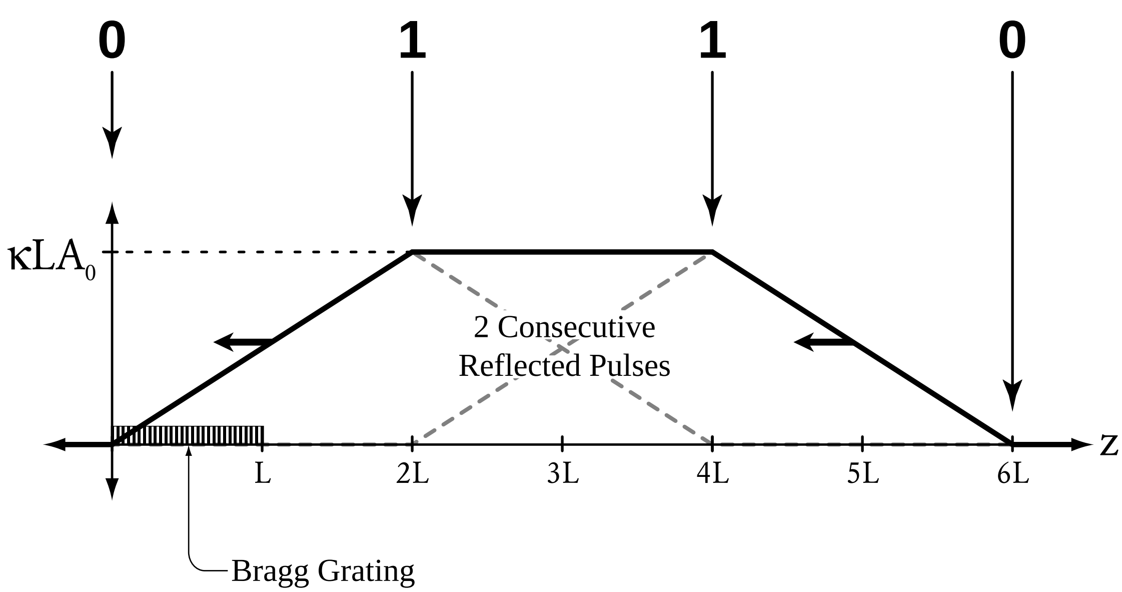


Figure 3.11 Two consecutive reflected pulses from an ideal matched grating filter. The resultant pulse has a triangular shape formed by two superposed triangles as shown. If the filtered signal is sampled at the points indicated in this figure, there is no interference between adjacent pulses.

filter, indicating where the signal would be sampled. The resulting filtered signal is trapezoidal in shape, which when sampled as shown in Fig. 3.11 would give two independent binary 1's. Even though the pulse shape is modified by the filter, it is modified in such a way that if the data is regularly sampled at the peak points, the measurements of individual bits are independent of one another.

Unfortunately, for a practical Bragg grating filter, the spectrum does not exactly match the incident pulse, and as a result the pulse amplitude does not decay to zero before the subsequent pulse is sampled, as can be seen in Fig. 3.4 for a Bragg grating with $\kappa L = 0.4$. Figure 3.12 is a plot of the response of the same Bragg grating filter to two adjacent bits of data, indicating the points at which the resulting signal would be sampled in the receiver. The extent to which the sampled values differ between adjacent bits of information is a measure of the amount of intersymbol interference caused by the non-ideal spectrum of the grating. The intersymbol interference of a grating filter is a consequence of multiple reflections within the grating region. If the grating acted as a one-pass distributed reflector, then the maximum path length that light could travel in the grating would be 2L, corresponding to one down-and-back trip through the grating. For a pulse of length 2L incident on a single-pass reflection filter of length L, it would be impossible for the reflected pulse to be longer than 4L because light from the trailing edge of the incident pulse enters the grating at time $2L/v_g$, and at a time $2L/v_g$ later all of this light has emerged from the grating. The single-pass distributed reflector model accurately predicts the triangular response derived in equation 3.27. For a very weak grating, the reflection generated by the grating is small to begin with, and therefore the effect of multiple reflections can be neglected to lowest order.

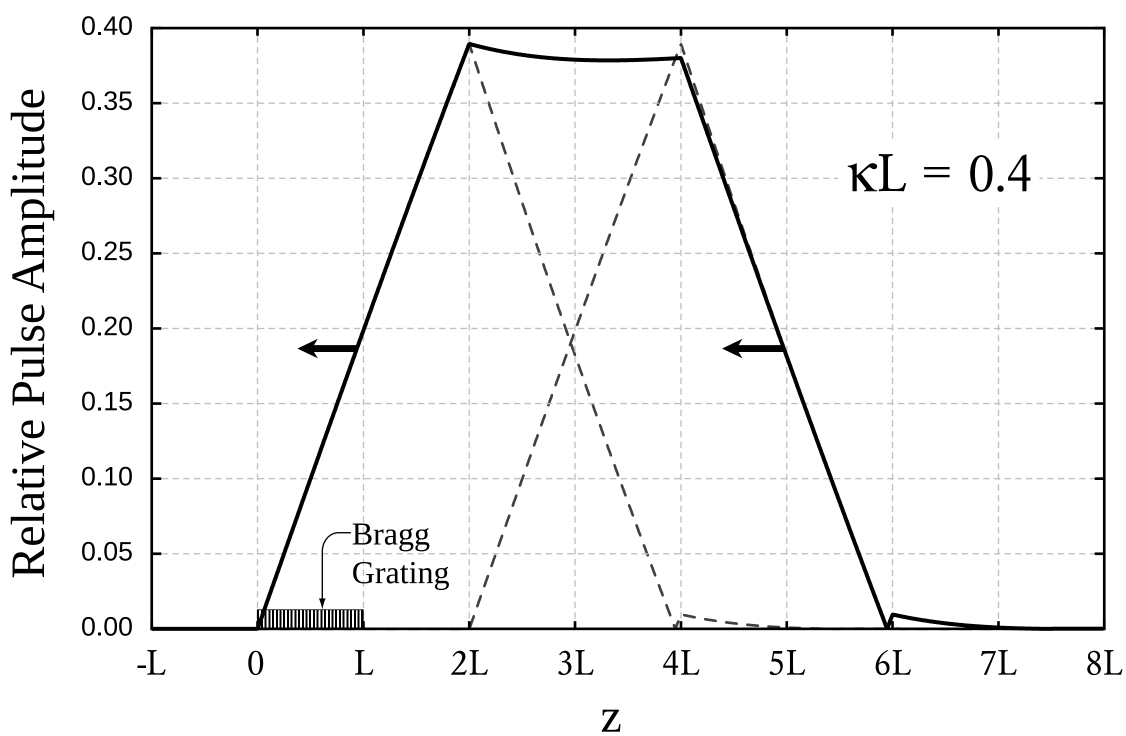


Figure 3.12 Consecutive reflected pulses from a weak Bragg grating. Notice that the peak values of the reflected signal are no longer the same, indicating that there is some interference between adjacent bits.

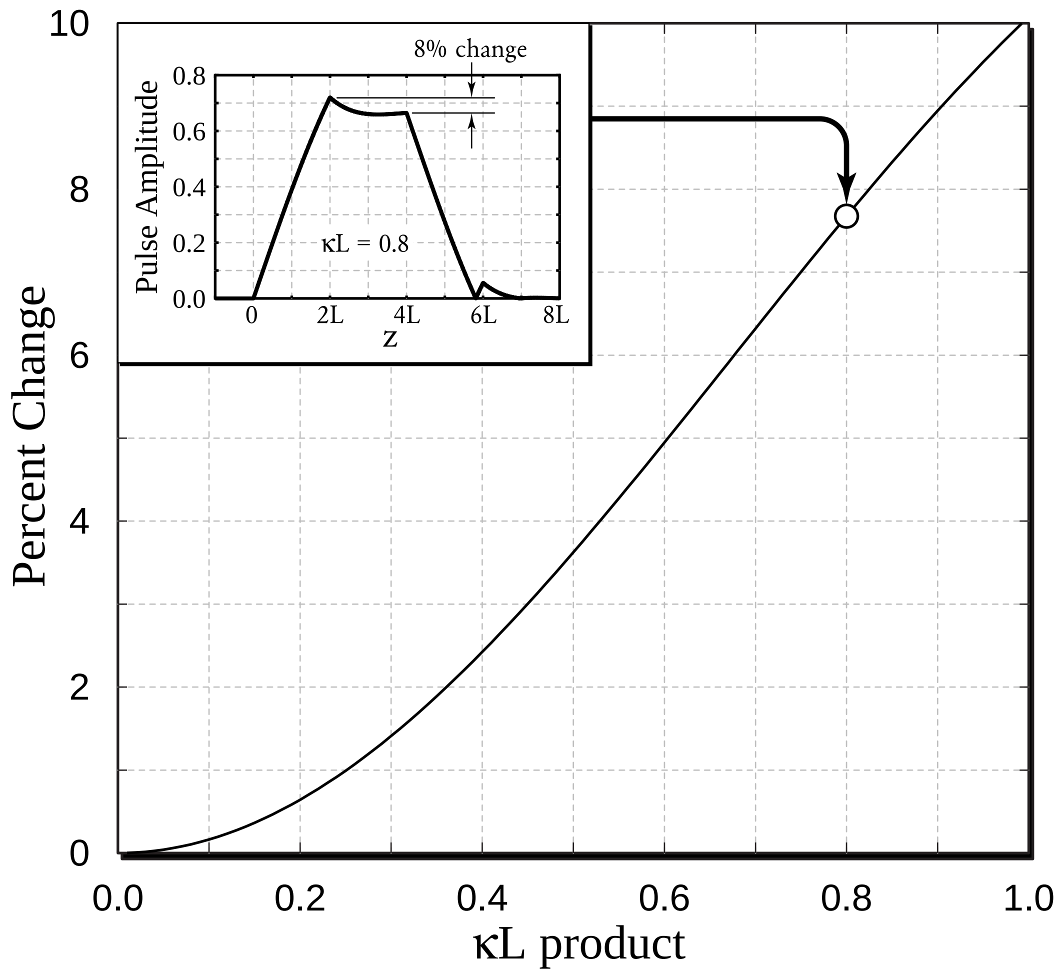


Figure 3.13 Intersymbol interference caused by multiple reflections within the grating. The intersymbol interference plotted here is specified as the percent change in the sampled values of two consecutive positive bits of information in a signal, as illustrated in the inset figure for the case of κL = 0.8.

However for a stronger grating, multiple reflections cause some of the light from the incident pulse to bounce around in the grating, ultimately emerging during a subsequent timeslot, thereby causing intersymbol interference.

Figure 3.13 provides a plot of the amount of intersymbol interference as a function of grating strength (κL). In Fig. 3.13, the intersymbol interference is defined as the percent change in the sampled peak value of a pulse resulting from the presence of an immediately preceding pulse.

Of course, a real communications signal consists of a random sequence of binary 1's and 0's, not an isolated pulse in time. Figure 3.14 presents the time response of various filters to a random sequence of binary pulses. Four separate cases are plotted: an ideal matched filter for which the reflected pulses are triangles with no intersymbol interference, a commonly used fiber Fabry-Perot (Lorentzian) filter, a strong Bragg grating filter, and finally a weak Bragg grating filter which has a

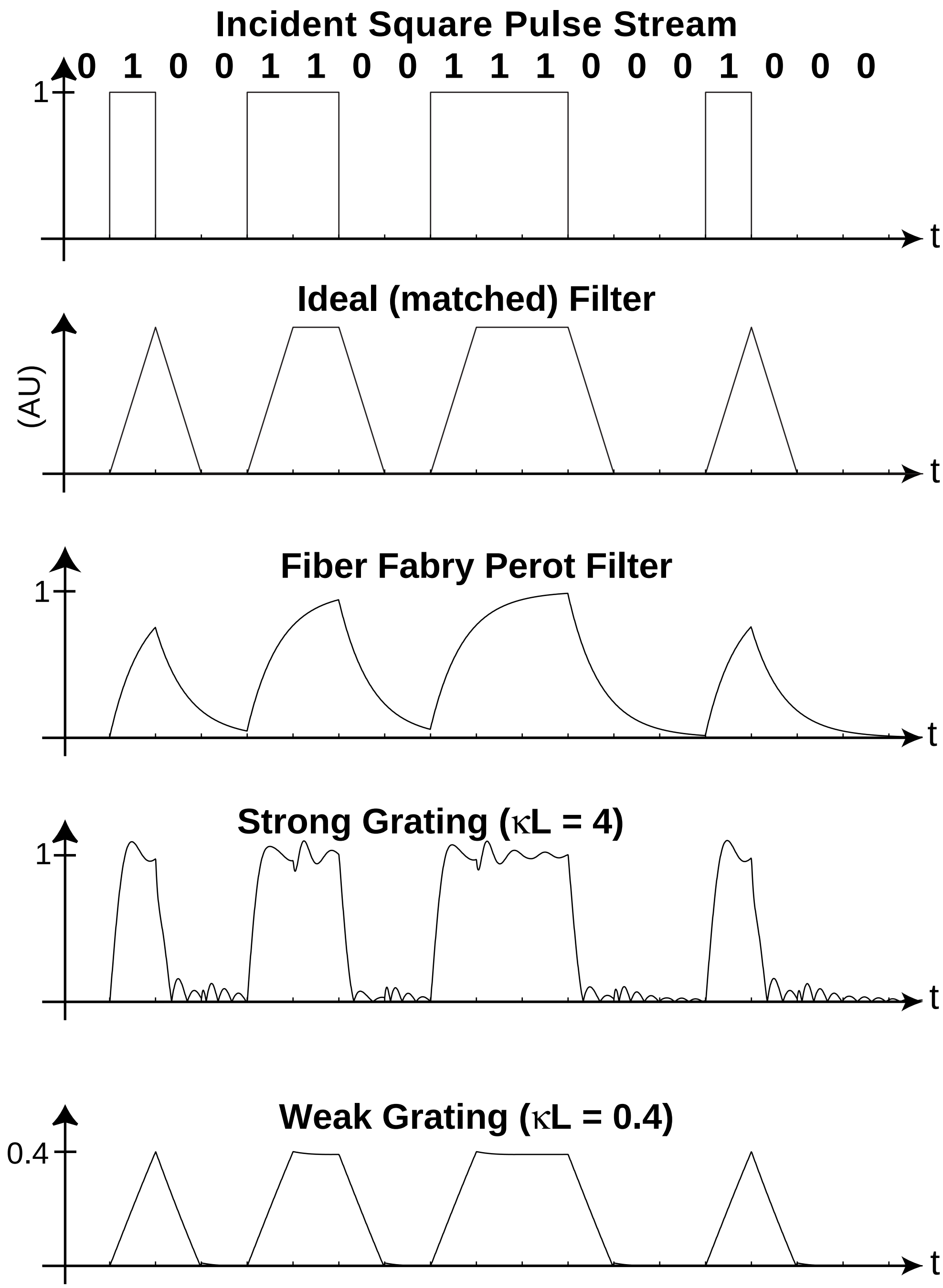


Figure 3.14 The effect of a filter on a sequence of On-Off Keyed data. The first plot shows the incident pulse sequence, represented as a series of square pulses. The following four plots illustrate the shape of the filtered pulse sequence for the case of a matched filter, a Fabry-Perot filter, a strong grating filter, and for a weak grating filter.

response close to the ideal matched filter. The results presented in Fig. 3.14 show how a weak Bragg grating filter can have a temporal response that closely resembles that of the ideal matched filter.

3.4 Summary

This chapter described in some detail how a Bragg grating device can act as a filter for optical communications signals. The reflection spectral response of a Bragg grating was calculated using coupled mode theory, which treats the grating as a periodic perturbation of an otherwise normal waveguide. We showed that if the parameters of the grating are properly chosen, the reflection spectrum can match the spectrum of a digital optical signal. Having derived the spectral response of the grating filter, the time response was then examined. The undesirable effect of a non-ideal spectral response was quantified in terms of the amount of intersymbol interference.

Chapter 4: Dielectric Waveguide Structures

The goal of this chapter is to provide a concise description of some important properties of dielectric waveguides. There are several excellent references texts such as [15, 16, 17, 18] which are devoted entirely to the subject of guided wave optics, and the reader is referred to them for an unabridged explanation of the physics of dielectric waveguides. This chapter seeks to cover as completely as possible those concepts of dielectric waveguides which are necessary for an understanding of grating-based filter devices, including the basic physics of optical waveguides, an overview of waveguide materials and fabrication techniques, and a derivation of the coupled mode equations which describe how oppositely travelling modes interact in a periodic waveguide.

In the Section 4.1, we present an overview of the theory of dielectric waveguides. We begin with a general discussion of the concepts of waveguides, and describe the characteristics of bound modes in optical waveguides, using the simple case of a planar slab waveguide as an illustrative example.

Section 4.2 describes the geometry of some common waveguide structures, including a brief overview of the fabrication techniques and materials used to construct these devices. Included is a description of the waveguide materials and geometry which we plan to use to fabricate the matched optical filters described in this work.

Having introduced the central ideas of optical guiding structures, we present a derivation of the coupled mode equations which describe the interaction between a grating and waveguide. This derivation explains how the grating strength, κ, is related to the properties of the bound mode of the waveguide and the geometry of the grating. The coupled mode equations were used as the basis for much of the material presented in Chapter 3, which was devoted to their solutions.

4.1 Electromagnetic Waves in Dielectric Media

Fundamentally, the propagation of light in a waveguide is described by Maxwell's equations, which govern all electromagnetic phenomena. There are many excellent references which completely describe the historical origins of Maxwell's equations and the meanings of all of the terms and quantities [19, 20, 21]. This thesis is not intended to be a treatise on electromagnetic theory, and we will therefore assume that the reader is reasonably familiar with the basic concepts of electromagnetism and dielectric materials. When describing electromagnetic propagation in lossless dielectric media, it is convenient to express Maxwell's equations in the following form:

$$\begin{aligned}
&\nabla \times \mathbf{E} = ik\sqrt{\frac{\mu_0}{\varepsilon_0}}\mathbf{H} \\
&\nabla \cdot (n^2\mathbf{E}) = 0 \\
&\nabla \times \mathbf{H} = -ikn^2\sqrt{\frac{\varepsilon_0}{\mu_0}}\mathbf{E} \\
&\nabla \cdot \mathbf{H} = 0
\end{aligned} \tag{4.1}$$

where **E** and **H** are the electric and magnetic fields respectively, and $n(\mathbf{r})$ is the dimensionless index of refraction which for a waveguide is a function of position. In lossless materials, $n(\mathbf{r})$ is a positive real quantity. ε_0 and μ_0 are the permittivity and permeability of free space, respectively. In this thesis, we will treat only non-magnetic materials, for which $\mu = \mu_0$. A time dependence of $e^{-i\omega t}$ is assumed for all of the field quantities in equation 4.1, therefore, the fields **E** and **H** described in equation 4.1 are explicitly functions of position only. The quantity k in equation 4.1 is defined by:

$$k \equiv \frac{\omega}{c} = \frac{2\pi}{\lambda} \tag{4.2}$$

k, which has dimensions of inverse length, is the free-space propagation constant of radiation with frequency ω. Some authors use the symbol k to denote the propagation constant of the wave in the waveguide. We have instead adopted the notation where k is defined in terms of the free-space wavelength as described in equation 4.2, and β will denote the propagation constant in the waveguide. The form of Maxwell's equations given in equation 4.1 applies only in source-free regions of space, where there is no free current or charge ($\mathbf{J} = 0$ and $\sigma = 0$.)

4.1.1 General Properties of Waveguides

A dielectric waveguide structure is completely specified by the refractive index profile $n(\mathbf{r})$. For a uniform straight waveguide, $n(\mathbf{r})$ is invariant to translations in one direction. Typically, the Cartesian axes are chosen such that the waveguide points along the z direction, meaning that $n(\mathbf{r})$ depends only on the Cartesian coordinates x and y, but not on z:

$$n(\mathbf{r}) = n(x, y) \tag{4.3}$$

The geometry of a uniform waveguide is described completely by the two dimensional index profile n(x,y). Figure 4.1 illustrates one example of an optical waveguide structure. In this figure, the waveguide is composed of a *core* region with index of refraction n_{core} surrounded by a *cladding* region with index n_{clad}. The index profile depicted in Fig. 4.1 is a piecewise constant function in the x-y plane, however, in general n(x,y) can vary continuously in the x-y plane.

In solving for the electric and magnetic fields in a waveguide structure, we first assume solutions which represent travelling waves in the z direction:

$$\begin{aligned} \mathbf{E}(x, y, z) &= \mathbf{e}(x, y)e^{i\beta z} \\ \mathbf{H}(x, y, z) &= \mathbf{h}(x, y)e^{i\beta z} \end{aligned} \tag{4.4}$$

where β describes the propagation constant in the z direction. Positive values of β correspond to light propagating in the positive z direction, while negative values of β correspond to light travelling in the reverse direction. When these equations are substituted into Maxwell's equations (4.1), the following vector wave equations are obtained after some algebra:

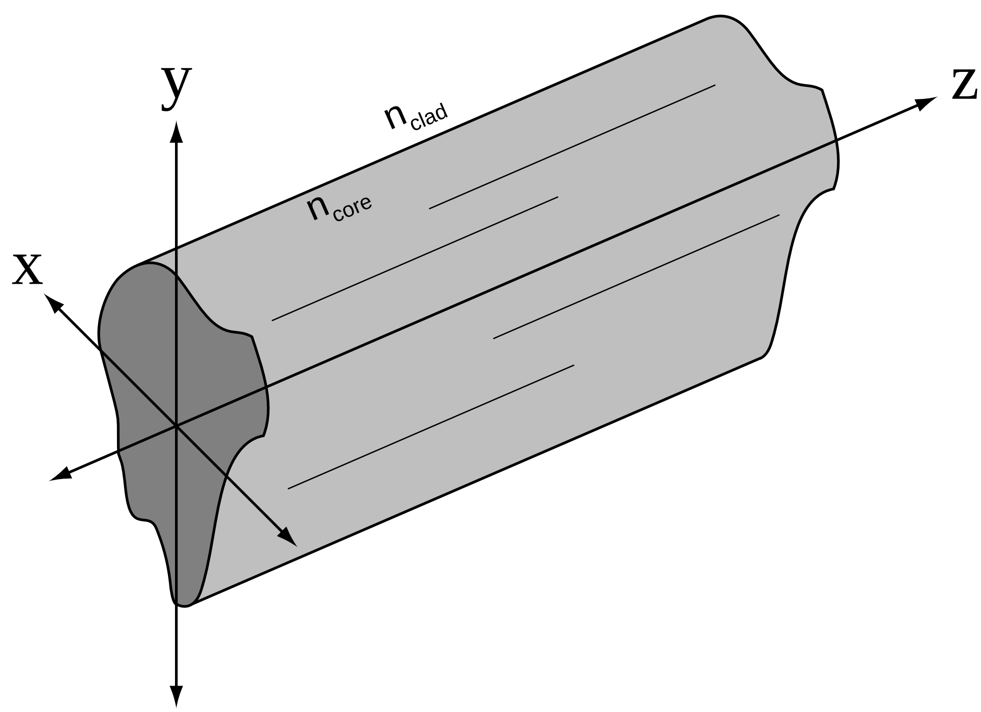


Figure 4.1 An optical waveguide structure. The axes have been oriented so that the index of refraction profile $n(\mathbf{r})$ is invariant to translations in the z direction. In the above figure, the index profile has a core region with index of refraction n_{core} surrounded by an infinite cladding region with index of refraction n_{clad}.

$$\left(\nabla_{xy}^2 + n^2k^2 - \beta^2\right)\mathbf{e} = -\left(\nabla_{xy} + i\beta\hat{\mathbf{z}}\right)\left(\mathbf{e}\cdot\nabla_{xy}\ln(n^2)\right)$$
$$\left(\nabla_{xy}^2 + n^2k^2 - \beta^2\right)\mathbf{h} = \left(\left(\nabla_{xy} + i\beta\hat{\mathbf{z}}\right)\times\mathbf{h}\right)\times\left(\nabla_{xy}\ln(n^2)\right) \tag{4.5}$$

[22, 15, 23] Equation 4.5 is the vector wave equation for light propagation in a guiding dielectric structure. It is important to realize that the six field components of **e** and **h** in equation 4.5 are not independent. If the two longitudinal components e_z and h_z are known, the transverse components $\mathbf{e_{xy}}$ and $\mathbf{h_{xy}}$ can be expressed in terms of the known field components using Maxwell's equations [15]. Likewise, if either $\mathbf{e_{xy}}$ or $\mathbf{h_{xy}}$ is known, the remaining 4 field components can be derived from the two known transverse field components using Maxwell's equations [22].

Often the index profile n(x,y) is a piecewise constant function, meaning that when viewed in cross section, the waveguide is comprised of a finite number of regions, each of which has a uniform index of refraction. One such structure, illustrated in Fig. 4.1, is a uniform core region surrounded completely by a cladding region with a lower index of refraction. Even for continuously varying index profiles such as graded index fibers or thermally diffused waveguides, it is often a good approximation to treat the index of refraction profile as a series of piecewise constant regions. For cases where the index of refraction can be expressed as a piecewise constant function of x and y, the complicated left-hand side of equation 4.5 is zero everywhere *except* on the boundary between regions. Within a given region i, the electric and magnetic fields must obey the simplified equations:

$$\left(\nabla_{xy}^2 + n_i^2k^2 - \beta^2\right)\mathbf{e} = 0$$
$$\left(\nabla_{xy}^2 + n_i^2k^2 - \beta^2\right)\mathbf{h} = 0 \tag{4.6}$$

where n_i is the index of refraction of the i^{th} region. These equations have well known solutions, which can be written in terms of sine's, cosine's exponential's, or other harmonic functions. Once the form of the solutions in each region are known, the entire electromagnetic field profile can be found by piecing together the solutions in each region in such a way that the electromagnetic boundary conditions are satisfied at the interface between regions. As long as the boundary conditions are satisfied, the $\nabla_{xy}\ln(n^2)$ terms on the right-hand side of equation 4.5 may be neglected. The electromagnetic boundary conditions at the interface between two dielectrics can be derived from Maxwell's equations:

$$\begin{aligned}
(\mathbf{e_2} - \mathbf{e_1})\times\hat{\mathbf{n}} &= \mathbf{0} && \text{Parallel component of } \mathbf{e} \text{ is continuous}\\
(\mathbf{h_2} - \mathbf{h_1})\times\hat{\mathbf{n}} &= \mathbf{0} && \text{Parallel component of } \mathbf{h} \text{ is continuous}\\
(n_2^2\mathbf{e_2} - n_1^2\mathbf{e_1})\cdot\hat{\mathbf{n}} &= 0 && \text{Normal component of } n^2\mathbf{e} \text{ is continuous}\\
(\mathbf{h_2} - \mathbf{h_1})\cdot\hat{\mathbf{n}} &= 0 && \text{Normal component of } \mathbf{h} \text{ is continuous}
\end{aligned} \tag{4.7}$$

[19] The first two equations state that the components of the electric and magnetic field which are parallel to the boundary surface must be continuous. The last two equations state that the normal component of the magnetic field must be continuous across the interface, as must the normal component of

$n^2\mathbf{e}$. As mentioned before, the six field components are interrelated through Maxwell's equations, so that in solving equation 4.6 in a given region only two of the field components need to be found; the others can then be derived from Maxwell's equations.

We can classify the solutions to equation 4.6 in a given region as either oscillatory or exponential in nature, depending on the value of the propagation constant β. When $\beta < kn_i$, the fields will be oscillatory in nature, expressed for example in terms of sine's, cosine's, Bessel functions of the first kind, or some other appropriate set of functions. When $\beta > kn_i$ the fields will be growing or decaying in nature, expressible for example in terms of exponential functions, hyperbolic functions or Bessel functions of the second kind. Strictly speaking, it is also possible for the mode to exhibit oscillatory characteristics in one direction while growing or decaying in the orthogonal direction. Generally, however, the fields in waveguide structures are oscillatory over a finite region near the center (or core) of the waveguide and decay to zero outside of this region where the surrounding index of refraction is lower. For this reason, all practical waveguide structures consist of a core region surrounded by a cladding region which has a lower index of refraction. In order for the light to be confined or bound in the guiding structure, the propagation constant β must satisfy:

$$n_{clad}k < \beta < n_{core}k \tag{4.8}$$

where n_{core} and n_{clad} represent the minimum and maximum values of the index of refraction profile $n(x,y)$. Under these conditions, it is found that the wave equation 4.5 only admits solutions for certain discrete values of the propagation constant β. In other words, equation 4.5 is an eigenvalue equation, which yields a series of discrete eigenvalues β_m describing the allowable propagation constants for bound modes. The corresponding electromagnetic field solutions (eigenmodes) are commonly referred to as *bound modes* of the waveguide.

When β lies outside of the range defined in equation 4.8, the solutions to equation 4.5 are no longer bound modes. When $\beta < n_{clad}k$, the fields correspond to unbound modes, also called radiation modes. In this case there is no longer a discrete set of allowable propagation constants, rather any value of β admits an unbound solution. The other case, $\beta > n_{core}k$, yields an unphysical solution in which the electromagnetic fields grow exponentially away from the waveguide core.

It can be shown that because of the symmetry properties of the wave equations, the transverse components of the electromagnetic field are 90 degrees out of phase with the longitudinal components [16]. Most authors therefore choose to treat the longitudinal field components e_z and h_z as pure imaginary quantities, making the transverse field components $\mathbf{e_{xy}}$ and $\mathbf{h_{xy}}$ pure real quantities:

$$\begin{array}{ll} \mathbf{e_{xy}}, \mathbf{h_{xy}} & \text{real quantities} \\ e_z, h_z & \text{imaginary quantities} \end{array} \tag{4.9}$$

One special type of waveguide is the metallic waveguide, which consists of a metal tube of arbitrary cross section extending in the z direction, possibly filled with a dielectric material. This type of structure is often used as a waveguide for microwave devices. Inside of the core region of this metallic waveguide, the electromagnetic fields again obey the wave equations 4.6. However, for the metallic waveguide, the boundary conditions are particularly simple: the parallel components of the electric field must go to zero as must the perpendicular components of the magnetic field [19]. With these simple boundary conditions, it can be shown that the modes of a metallic waveguide can be divided into two classes: TE modes, in which $e_z = 0$, and TM modes in which $h_z = 0$. The wave equation can be solved separately for these two types of modes. In dielectric waveguides, however, the fields do not approach zero at the core boundary, but rather they must be matched to the decaying fields outside of the core. In general, it is impossible to satisfy the boundary conditions everywhere unless the fields have both e_z and h_z components. Thus the nature of the boundary conditions at dielectric interfaces prohibits pure TE and TM solutions for most dielectric waveguide structures. For this reason, the modes of dielectric waveguides are often referred to as "hybrid modes." [24] Exceptions to this are the planar slab waveguide and the circular core waveguide, for which pure TE and TM modes *do* exist. Nonetheless, for many waveguide structures, one field component is found to be substantially larger than the others and the mode is described as quasi-TE or quasi-TM. [25]

The process of wave guidance and bound modes is analogous to the quantum mechanical phenomenon of quantized eigenstates associated with potential energy wells. For example, in quantum mechanics, an electron can be bound by the potential energy well generated by a positively charged nucleus. The quantum mechanical solution to this well known problem is a series of eigenstates, characterized by discrete energy levels [26]. In addition to the discrete spectrum there is also a continuous spectrum associated with unbound solutions in which the electron's wavefunction is not confined to the region surrounding the nucleus. The transverse index profile of a dielectric waveguide structure can be thought of as analogous to a potential well in quantum mechanics. Just as a negative potential well can yield bound mode solutions in quantum mechanics, a region of relatively higher index of refraction can give rise to bound *optical* mode solutions in a dielectric waveguide. Of course, one key difference between these phenomena is that the eigenstates of a quantum mechanical system are described by a scalar wavefunction*, whereas the bound modes of a dielectric waveguide are described by the full electric and magnetic vector field profiles. The boundary conditions 4.7 for the case of a dielectric waveguide are correspondingly more complex than for a quantum mechanical wavefunction. However, as we will describe later in Section 4.1.3, when the waveguide structure is a so-called weakly guiding structure, the wave equations and boundary conditions can be simplified to a form which is mathematically equivalent to the Schrodinger equation for a two-dimensional potential well.

---

* For particles with spin, the quantum mechanical wavefunction is in fact a 4-vector which obeys the Dirac equation. However in many cases the quantum mechanical behavior of particles can be adequately treated using the conventional scalar wavefunction which obeys the Schrodinger equation.

4.1.2 Planar Waveguide

There are very few dielectric waveguide structures that yield analytical solutions. Often numerical methods such as the finite difference method [27, 28], or finite elements method [29, 30, 31] are used to obtain numerical solutions to the partial differential equations describing the waveguide. One of simplest cases which can be solved analytically is a slab dielectric waveguide, illustrated in Fig. 4.2. Another dielectric structure which can be solved analytically is the circular fiber structure. In order to better illustrate some of the concepts of dielectric waveguides, we will present here a brief solution of the dielectric slab waveguide. The slab waveguide consists of an infinite slab of core material of thickness d, surrounded by two regions of cladding material. Because of its simple geometry, the planar waveguide structure has solutions which are either TE or TM modes. This result can be explained by the reflection of plane waves from a dielectric stack structure. When a plane wave with TE polarization is incident on a dielectric interface, the reflected and transmitted waves maintain the TE polarization [24]. Likewise, TM polarization is maintained in reflections from planar dielectric boundaries. Therefore, because of the simple geometry of the dielectric slab waveguide, there is no mixing between TE and TM components. In more complicated waveguide geometries such as channel waveguides with rectangular cores, the modes are neither TE nor TM in general.

For the slab waveguide, it is convenient to seek solutions for the transverse components e_y or h_y of the electromagnetic fields, and then derive the longitudinal components using Maxwell's equations. The transverse field components must be solutions to the wave equation 4.6. Table 4.1 summarizes the field components and boundary conditions for the TE and TM modes of the slab dielectric structure.

The quantity $\phi(x)$ represents e_y for the case of TE modes or h_y for the case of TM modes. In both cases, $\phi(x)$ must be a solution to the differential equation:

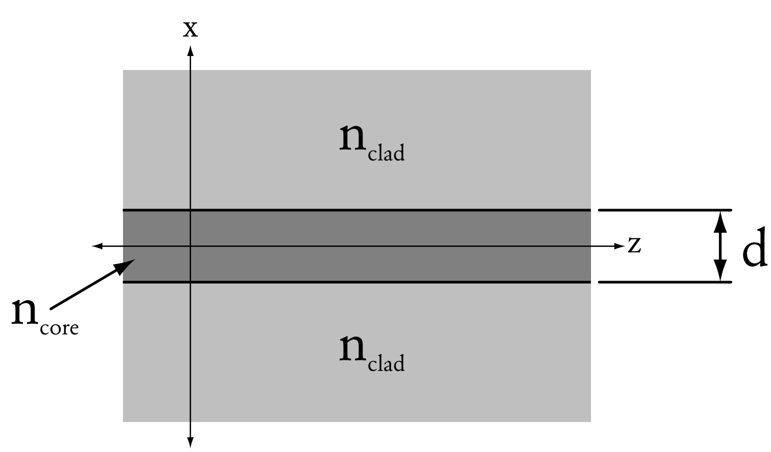


Figure 4.2 A slab dielectric waveguide. Light is confined in a planar slab of core material, which is surrounded on both sides by infinite cladding regions.

Table 4.1: Summary of field components for slab waveguides

TE Modes	**TM Modes**
$e_z = 0$	$h_z = 0$
$e_x = 0$	$h_x = 0$
$h_y = 0$	$e_y = 0$
$e_y = \phi(x)$	$h_y = \phi(x)$
$h_x = -\frac{\beta}{k}\sqrt{\frac{\varepsilon_0}{\mu_0}}\phi(x)$	$e_x = \frac{\beta}{k}\frac{1}{n^2}\sqrt{\frac{\mu_0}{\varepsilon_0}}\phi(x)$
$h_z = -\frac{i}{k}\sqrt{\frac{\varepsilon_0}{\mu_0}}\phi'(x)$	$e_z = \frac{i}{k}\frac{1}{n^2}\sqrt{\frac{\mu_0}{\varepsilon_0}}\phi'(x)$
$\phi(x)$ continuous at $x = \pm\frac{d}{2}$	$\phi(x)$ continuous at $x = \pm\frac{d}{2}$
$\phi'(x)$ continuous at $x = \pm\frac{d}{2}$	$\frac{1}{n^2}\phi'(x)$ continuous at $x = \pm\frac{d}{2}$

$$\left\{\frac{d^2}{dx^2} + n^2k^2 - \beta^2\right\}\phi(x) = 0 \tag{4.10}$$

The solution for $\phi(x)$ has the following form:

$$\phi(x) = \begin{cases} Be^{-\alpha x} & \frac{d}{2} < x \\ A\,{\cos \atop \sin}(k_x x) & -\frac{d}{2} \le x \le \frac{d}{2} \\ \pm Be^{\alpha x} & x < -\frac{d}{2} \end{cases} \tag{4.11}$$

where the upper symbols are used for modes with even symmetry and the lower symbols are used for modes with odd symmetry. The quantities α and k_x are defined by:

$$\alpha \equiv \sqrt{\beta^2 - n_{clad}^2 k^2} \tag{4.12}$$

$$k_x \equiv \sqrt{n_{core}^2 k^2 - \beta^2} \tag{4.13}$$

The differential equation for $\phi(x)$ and the form of the solution 4.11 is identical for TE and TM modes, however the boundary conditions differ for these two cases. Applying the boundary conditions given in table 4.1 gives two linear equations relating the coefficients A and B in equation 4.11. In order for a non-trivial solution for A and B to exist, the following secular equation must be satisfied:

$$k_x d = m\pi + 2\tan^{-1}\left(\frac{\alpha}{k_x}\right), \text{ for TE modes} \tag{4.14}$$

$$k_x d = m\pi + 2\tan^{-1}\left(\frac{\alpha}{k_x}\left(\frac{n_{core}}{n_{clad}}\right)^2\right), \text{ for TM modes} \tag{4.15}$$

where m is a non-negative integer which is even for modes with even symmetry and odd for modes with odd symmetry. α and k_x are related to β through the relationships given in equations 4.12 and 4.13. Therefore, equations 4.14 and 4.15 are transcendental equations which when solved give the values of β corresponding the bound modes of the slab waveguide.

In general, equations 4.14 and 4.15 can have one or more solutions, depending on the parameters of the waveguide and the optical frequency. Figure 4.3 depicts graphically the solution to equation 4.14 for one particular set of waveguide parameters. First, equation 4.14 is rewritten as a function of the parameter $k_x d$, using equations 4.12 and 4.13:

$$\tan\left(\frac{1}{2}k_x d - m\frac{\pi}{2}\right) = \frac{\sqrt{\left(n_{core}^2 - n_{clad}^2\right)k^2 d^2 - k_x^2 d^2}}{k_x d} \tag{4.16}$$

The solution to equation 4.16 can be found by separately plotting the right-hand side and left hand side of the equation as a function of the dimensionless parameter $k_x d$, as shown in Fig. 4.3. The points of intersection correspond to the solutions of the equation. The eigenvalues β can then be found from $k_x d$, using equation 4.13. Figure 4.3 illustrates the solution to equation 4.16 for the case where $kd\sqrt{n_{core}^2 - n_{clad}^2} = 8$. It is clear by inspection of the graphical solution presented in Fig. 4.3 that the number of bound modes of the system is determined entirely by the parameter $kd\sqrt{n_{core}^2 - n_{clad}^2}$. This parameter is often referred to as the normalized frequency, denoted V:

$$V = kd\sqrt{n_{core}^2 - n_{clad}^2} \tag{4.17}$$

For the case of V = 8, there are three bound TE modes, but often waveguides are designed so that there is only one bound mode for each polarization. From Fig. 4.3 and equation 4.16 it is easy to see that when the parameter V falls below the threshold value of 2π, there will no longer be a solution for the m = 2 mode. This means that if the optical frequency (which is proportional to k, as described in equation 4.2) drops below a certain value, the m = 2 mode will no longer exist. The mode is then said to be *cut-*

off. Every mode of the waveguide, except for the fundamental (m = 0) mode has a characteristic cut-off frequency.

The value of the transverse electric field component $e_y = \phi(x)$ for each of the bound modes is plotted in Fig. 4.4. The TM modes of the structure can be found in a similar manner, by adding an appropriate factor to equation 4.16. Note that for the TM case, the boundary conditions specify that the first derivative of $\phi(x)$ must be discontinuous at the boundary between adjacent layers. Therefore, the field profiles and eigenvalues would be somewhat different for the case of TM modes.

One feature of waveguide modes which is illustrated by the slab waveguide example is mode orthogonality. The wavefunctions $\phi_m(x)$ plotted in Fig. 4.4 are orthogonal in that the overlap integral of any two different wavefunctions will be zero. The mode orthogonality relationship can be expressed more generally as [15, 16]:

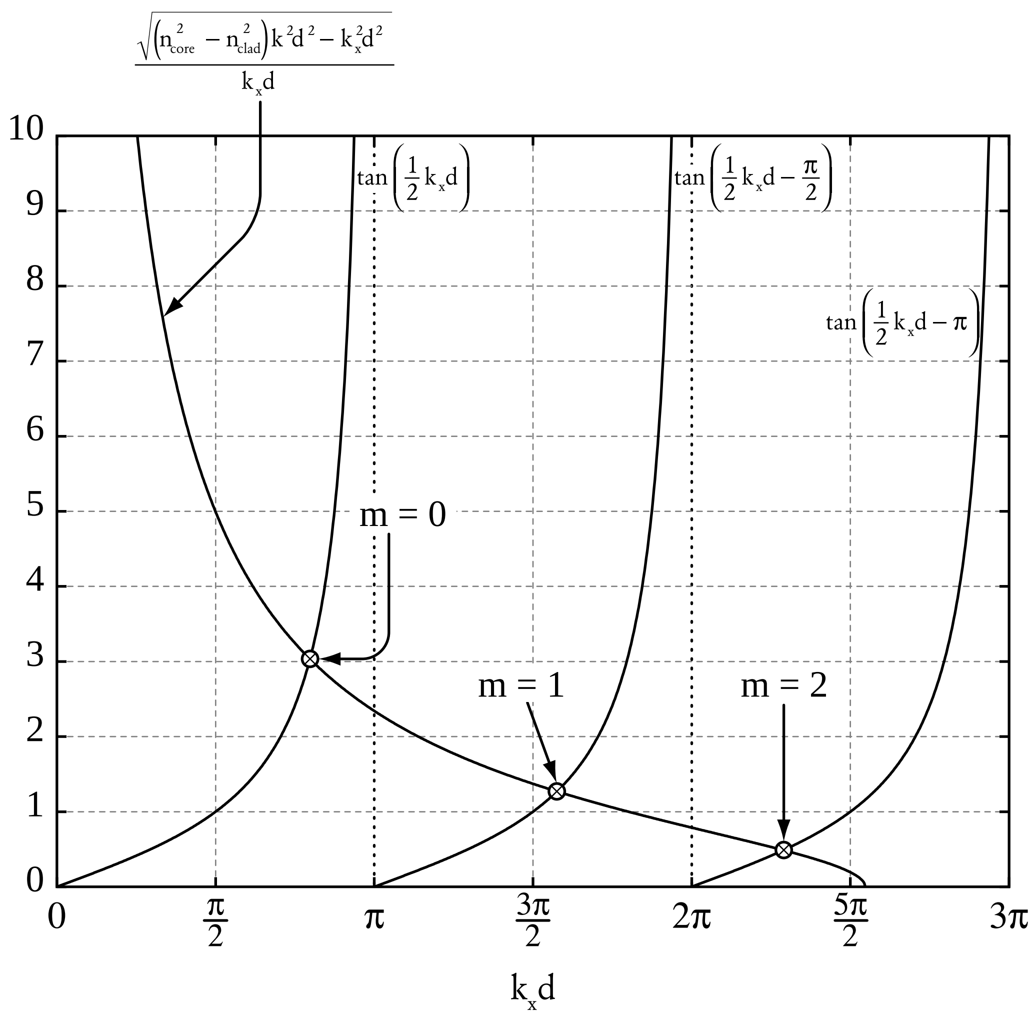


Figure 4.3 A graphical solution for the TE modes of a planar dielectric waveguide. In this example, there are three bound mode solutions, corresponding to the points of intersection. Even values of m correspond to symmetric modes, whereas odd values of m correspond to antisymmetric modes.

$$\iint \left[\mathbf{e}_n \times \mathbf{h}_m^* + \mathbf{e}_m^* \times \mathbf{h}_n \right] dA = \pm 4 P_n \, \delta_{n,m} \tag{4.18}$$

where $\mathbf{e_n}$ and $\mathbf{h_n}$ represent the modal fields associated with the n^{th} bound mode and P_n is taken to be a positive quantity representing the power contained in the n^{th} mode. The positive sign in equation 4.18 is used for forward travelling modes, whereas the negative sign is used for backward travelling modes. Often the modes are normalized so that each mode carries unity power, thus $P_n = 1$.

4.1.3 Weakly Guiding Waveguides

Section 4.1.2 provides a complete description of the solution for the TE modes of a planar slab waveguide. The differential equation (4.10) satisfied by the transverse field component is identical for TE and TM modes, and the only difference between these two cases is that the boundary condition on the

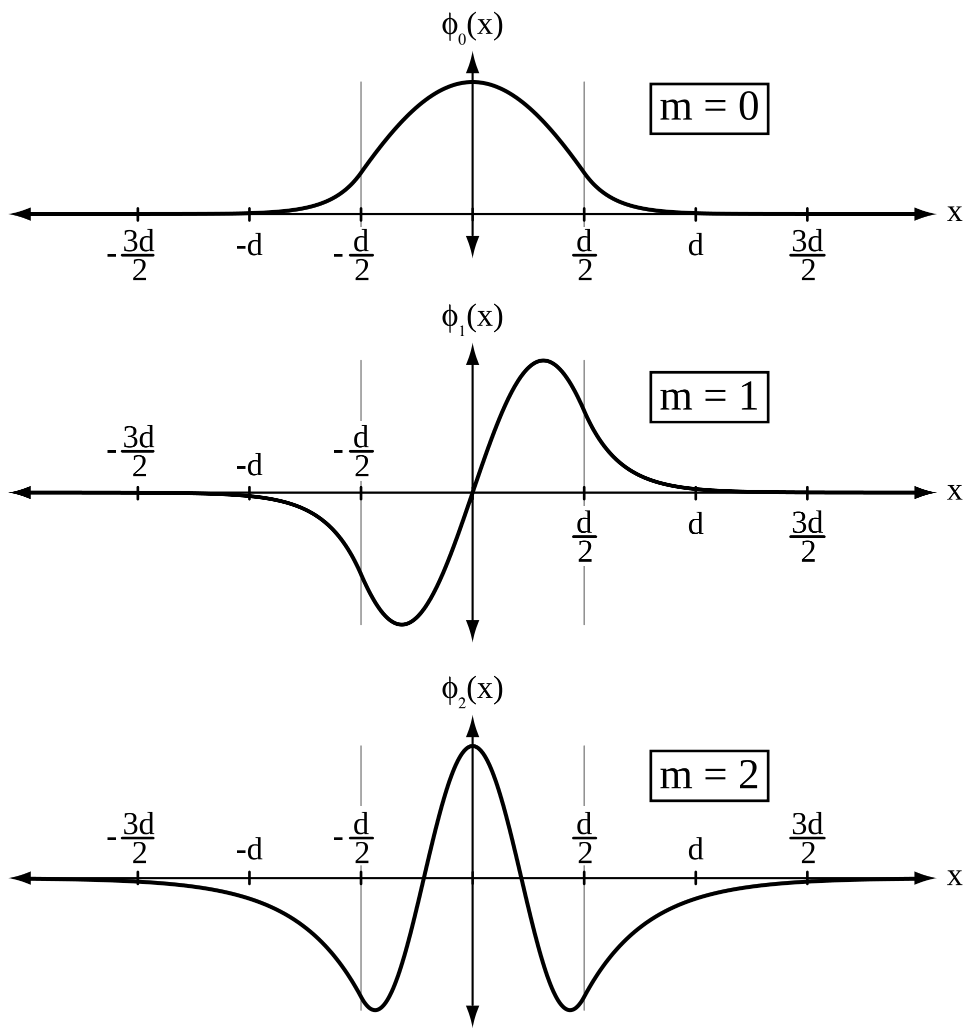


Figure 4.4 Transverse field profile for TE modes of a planar slab waveguide. There are three bound mode solutions, as illustrated graphically in Fig. 4.3.

derivative of $\phi(x)$ is different for the case of TM modes. This difference manifests itself through an extra factor of $(n_{core}/n_{clad})^2$ which appears in the dispersion relation for TM modes. For many waveguide structures of interest, the difference between the core and cladding index of refraction is small. For example, in commonly used single-mode optical fibers, the core has an index of refraction which is only 0.3% higher than the surrounding cladding region [32]. If we consider a planar slab waveguide with very similar core and cladding indices, then it is clear from the preceding analysis that the TE and TM cases would be almost identical. We would therefore expect both types of modes to have almost identical mode profiles (although the polarizations are orthogonal.)

The argument given above suggests that when the index contrast in a waveguide is small, the effects of polarization can be neglected. The resulting electromagnetic fields are almost TEM in character, with very small longitudinal components. The bound light rays inside of the core may be thought of as reflecting at grazing incidence from the core-clad boundary in such a way that total internal reflection occurs and the field outside of the core is evanescent.

The condition that the core and cladding have similar indices is often referred to as the “weak guiding” condition. The term “weak guiding” is perhaps misleading because it suggests that the optical mode is poorly confined within the waveguide core, or that the light leaks out of the waveguide. In fact, it is possible to construct a waveguide with a low core-clad index contrast which confines almost all of the optical power within the waveguide core. The term “weakly guiding” should be understood to mean that there is a very small fractional difference between the minimum and maximum indices of refraction. [15]

For waveguides that are weakly guiding, the wave equation can be simplified considerably. In a weakly guiding waveguide, it is sufficient to solve for only one scalar field quantity Φ, which represents one of the transverse field components. Φ must obey the scalar wave equation:

$$\left(\nabla_{xy}^2 + n^2(x,y)k^2 - \beta^2\right)\Phi(x,y) = 0 \tag{4.19}$$

The above equation seems deceptively similar to equation 4.6. However, equation 4.6 is only valid in a particular region of a piecewise constant index profile, and the field solutions must be pieced together in a way that satisfies the complicated electromagnetic boundary conditions. The scalar wave equation 4.19 describes the behavior of the scalar field Φ over the entire index transverse plane, even for the case of continuously varying profiles. Moreover, the quantity $\Phi(x,y)$ and its first derivative are both taken to be continuous at all points in the x-y plane.

The scalar wave equation greatly simplifies the analysis of weakly guiding waveguide structures. The form of the scalar wave equation and its simple boundary conditions is identical to the Schrodinger equation for two dimensional potential wells, completing the analogy described earlier in Section 4.1.1.

4.2 *Waveguide Structures*

Undoubtedly the most ubiquitous optical waveguide structure is the optical fiber. The type of optical fiber most commonly used in long distance communication links is single-mode optical fiber, which is engineered so that the structure supports only one bound optical mode for each polarization. The core and cladding regions are both constructed out of glass materials: primarily SiO_2, with other oxides such as GeO_2, B_2O_2, and Al_2O_3 added to modify the optical properties. [32] The index change between the core and cladding layers is achieved by modifying the amount of impurities or dopants in the glass. The fiber is fabricated by a process of drawing a thin tube of molten glass in a vertical furnace. Single mode optical fibers typically have core diameters ranging from 3 – 10 μm, and cladding diameters ranging from 50 – 125 μm. It is important to recognize that because of the circular symmetry of the optical fiber, there is no preferred polarization direction for the fundamental mode. Typically the waveguide core has an index of refraction which is only a fraction of a percent higher than the cladding region.

Integrated waveguides differ from fiber-optic cables in that the integrated waveguide is fabricated on a planar substrate using lithographic techniques. This geometry has some advantages over the fiber configuration, the greatest potential advantage being that the optical waveguide could in principle be integrated with other electronic or optical components in the communications system. In addition, the planar geometry allows for good control of the waveguide dimensions. However, the optical attenuation per unit length of even the best integrated waveguides is still an order of magnitude higher than that of optical fiber.

Because of the planar techniques used to build integrated waveguides, most integrated waveguides have a rectangular shape. Figure 4.5 illustrates some of the more common waveguide geometries. One of the simplest waveguide geometries is the channel waveguide, which consists of a rectangular core region, surrounded on all sides by a cladding region. Another common waveguide geometry is the ridge waveguide, which can be thought of as a planar slab waveguide in which the core has a ridge which confines the light in the horizontal direction. A variant of the ridge waveguide is the rib waveguide, in which the core region is a planar layer, and a rib is patterned on the upper surface of the cladding layer.

Semiconductor materials are often attractive for building integrated optical waveguides because it is easier to integrate the optical device with other semiconductor optical or electronic components of the system such as photodetectors, modulators, and diode lasers. Additionally, by injecting electrons or holes into a semiconductor waveguide device, it is possible to add optical gain to the device. The materials most commonly used for constructing such integrated waveguides are $Al_{1-x}Ga_xAs$ and $In_{1-x}Ga_xAs_{1-y}P_y$. The refractive index of the waveguide can be controlled by varying the mole fraction (x,y) of the atomic constituents. The index of refraction for these types of materials is typically around 3.1 – 3.5.

Various techniques can be used to deposit single-crystal semiconductor materials, including liquid phase epitaxy, vapor phase epitaxy, and molecular beam epitaxy.

Glass is often used as the raw material for fabricating passive optical waveguides. [33] One advantage of glass is that it is identical to the materials used to construct common optical fibers. Because of the close match in index of refraction, it is possible to couple efficiently from a fiber into a glass waveguide. As evidenced by the very low propagation loss of optical fibers, glass is an intrinsically low loss material, compared to other waveguide materials. Although the attenuation in planar glass waveguides is not as low as for optical fiber, it is still lower than in most semiconductor materials. When used for constructing integrated Bragg grating filters, another advantage of using glass waveguides is that they places less of a demand on the nanolithography technology. The index of refraction of glass is approximately 1.5, meaning that light with a free-space wavelength of 1.55 μm (commonly used in optical communications) has a wavelength of about 1 μm inside of a glass waveguide. Therefore, in order to construct a Bragg grating device in a glass waveguide, the grating period must be ~500 nm in order to meet the Bragg condition, whereas for common semiconductor materials with higher index of refraction, the grating period must be ~250 nm. Glass waveguides are typically fabricated over the surface of silicon substrates, which are very readily available in large sizes and are less expensive than semiconductor substrates.

Researchers at LETI, in France, have fabricated SiO_2 channel waveguides using plasma-enhanced chemical vapor deposition (PECVD) and reactive ion etching (RIE) [34]. Initially, a SiO_2 lower cladding is deposited onto a substrate, after which a phosphorus doped core region is deposited. The core region is patterned with photolithography to define a waveguide structure, and the waveguide is then etched

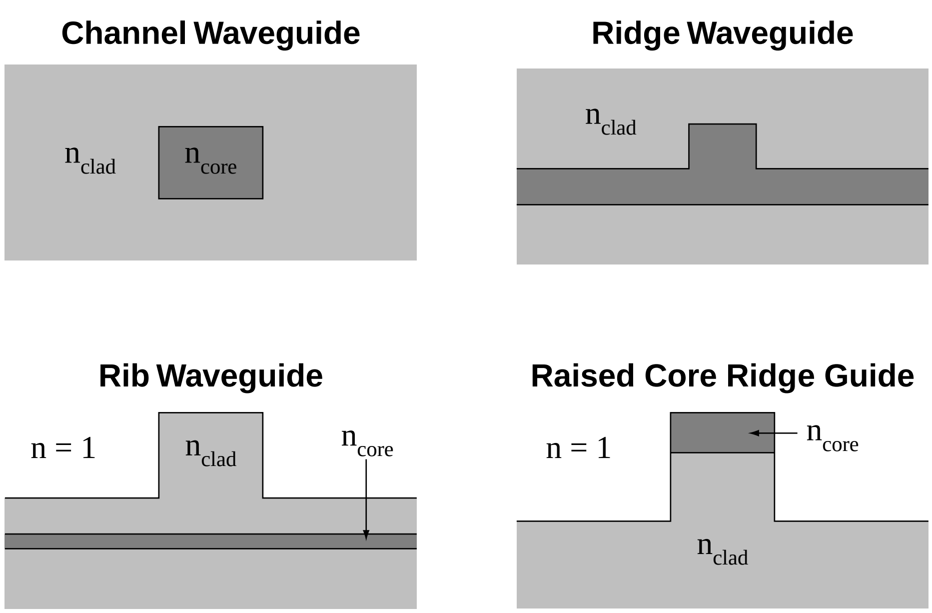


Figure 4.5 Common integrated waveguide configurations.

using RIE to form a rectangular core. The final step is to deposit an upper cladding layer of undoped SiO_2, thereby burying the rectangular core. The core dimensions are comparable in size to optical fibers, and the core-clad index contrast is 0.4%. The propagation loss for these waveguides ranged from 0.1–0.2 dB/cm. Researchers at Lucent Technologies have employed a similar procedure to construct glass optical waveguides [35]. In their work, the upper cladding region is doped with both phosphorus and boron in order to both match the index of the SiO_2 lower cladding layer while at the same time achieve a lower flow temperature. Because the upper cladding flows at a lower temperature than the underlying layers, it is possible to bury reasonably high aspect ratio waveguides without melting the underlying structures. Using these techniques, waveguides with propagation losses as low as 0.05 dB/cm were obtained. [36]

Workers at NTT construct glass waveguides using a process of flame hydrolysis deposition, in which fine glass particles are deposited onto a substrate with an oxygen/hydrogen torch [37, 38]. The process of flame hydrolysis is also used to generate the glass *preforms* from which optical fiber is drawn. In flame hydrolysis deposition, gaseous $SiCl_4$ is sent into an oxygen/hydrogen flame which produces porous layer of fine SiO_2 particles on the substrate. When this layer is heated in a furnace, it flows to form a solid high-quality optical layer. The index of refraction and flow temperature of the deposited layer can be adjusted by introducing $TiCl_4$ or $GeCl_4$ into the gas mixture. Using this technique, researchers at NTT were able to construct channel waveguides with propagation losses as low as 0.01 dB/cm [39, 40], and fiber coupling losses of only 0.05 dB [41].

Lithium niobate is an attractive material for fabrication of optical waveguides because of its acousto-optic and electro-optic properties. [42] Single crystal $LiNbO_3$ has the property that mechanical stresses in the material can change the optical index of refraction, and likewise an applied electric field can induce a change in index. These properties make $LiNbO_3$ an attractive waveguide material for constructing optical modulators and switches. However, forming a waveguide structure in crystalline $LiNbO_3$ is not as straightforward as for semiconductor materials or for doped glasses. Usually, the core region is formed by a process of proton exchange, or metal diffusion [42]. These methods typically yield a graded-index waveguide profile which cannot be controlled to the same degree as an etched waveguide structure.

There are a few other materials used for constructing optical waveguides which are not discussed in this work, including polymer film waveguides [43] and silicon nitride/oxide structures [33]. For the waveguide filter devices which are the subject of this thesis, we have selected doped glass waveguides as a particularly suitable material. This material system has several advantages over competing technologies, including low material loss, good index matching to optical fiber, and favorable grating dimensions for nanolithography. Planar waveguide layers, deposited by the flame hydrolysis deposition method, are now commercially available with a variety of core-clad index contrasts and core thicknesses. Furthermore, similar SiO_2 materials are used commonly in integrated circuit fabrication, and therefore the fabrication of glass waveguides can utilize the already mature processing technology developed for the IC industry.

Figure 4.6 is a diagram of the proposed waveguide structure for this work. The lower cladding region is nominally undoped SiO_2. The core layer is comprised of Ge doped SiO_2, with the doping level adjusted to achieve an index contrast of 0.3%. After patterning the waveguide and etching a grating in the surface of the core, a final layer of cladding oxide is deposited over the top of the structure. The channel waveguide geometry offers some advantages over ridge or rib waveguides: the square shape allows efficient coupling to an optical fiber, and relatively low polarization dependence. As described later in Chapter 5, the size of the waveguide and index contrast have been selected to achieve optimal coupling to an optical fiber. The raw material and flame hydrolysis deposition is to be supplied by a commercial vendor. Therefore, our research focuses on designing the waveguide and grating structures and developing the lithography for patterning the waveguide and grating onto supplied planar substrate materials.

4.3 Coupling of Modes with a Grating

We now consider the problem of a periodic waveguide structure. The modal analysis of waveguides presented earlier in this chapter assumes that the dielectric waveguide is described by an index profile n(x,y) that is independent of z. However, for a grating structure the index of refraction is a periodic function of z. Often, the grating can be treated as a small perturbation on top of an otherwise z-independent waveguide, in which case the effect of the grating is to couple the otherwise independent

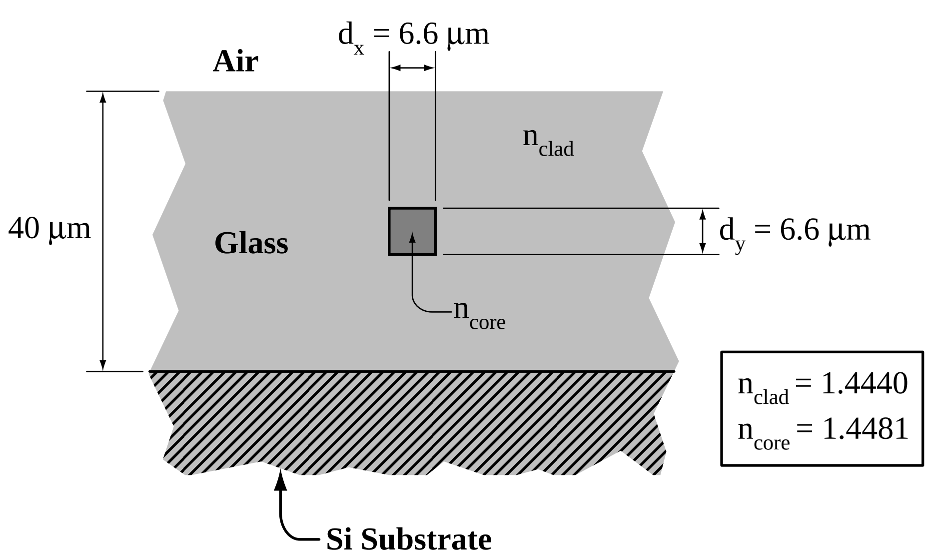


Figure 4.6 Geometry of integrated glass channel waveguide. The glass layers are to be deposited by flame hydrolysis. The core region is doped with GeO_2 which causes it to have an index of refraction which is 0.3% higher than the surrounding cladding.

forward and backwards travelling waves of the waveguide [44]. Coupled mode theory describes how this coupling occurs, and relates the coupling parameters to the geometry of the waveguide and grating.

Figure 4.7 illustrates how a grating in the surface of a waveguide can be treated as a perturbation of an otherwise uniform waveguide. The unperturbed index of refraction is denoted n(x,y). Generally, the unperturbed waveguide specified by n(x,y) has a series of forward and backward travelling modes:

$$n^2(x,y) \rightarrow \mathbf{e}_m(x,y)e^{i\beta_m z} \quad , \quad \mathbf{h}_m(x,y)e^{i\beta_m z} \tag{4.20}$$

The perturbed waveguide is described by a modified index of refraction $n^2(x,y) + \delta\varepsilon(x,y,z)$, where $\delta\varepsilon(x,y,z)$ is a perturbation of the waveguide structure. The field solutions for the perturbed waveguide are denoted **E** and **H**:

$$n^2(x,y) + \delta\varepsilon(x,y,z) \rightarrow \mathbf{E}\ ,\ \mathbf{H} \tag{4.21}$$

In coupled mode theory, we treat **E** and **H** as a linear superposition of the unperturbed waveguide modes:

$$\begin{aligned} \mathbf{E}(x,y,z) &= \sum_n a_n(z)\mathbf{e}_n(x,y) \\ \mathbf{H}(x,y,z) &= \sum_n a_n(z)\mathbf{h}_n(x,y) \end{aligned} \tag{4.22}$$

The summation index n extends over all of the forward and backward travelling bound modes of the system. For this analysis, we shall take n to be a non-zero integer which numbers the bound modes of the waveguide. Negative values of n correspond to backward travelling modes, and positive values of n cor-

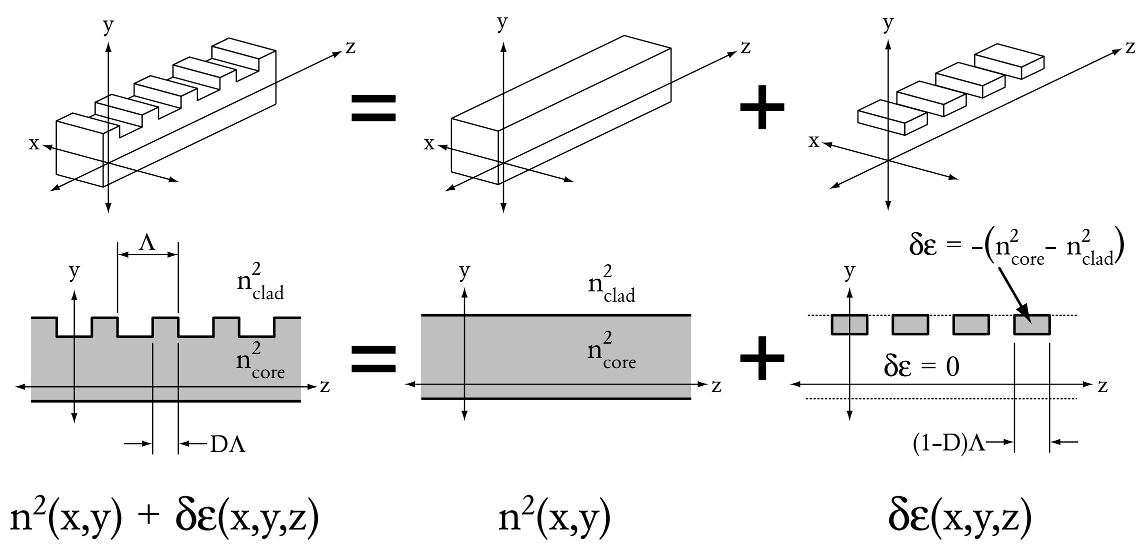


Figure 4.7 When a grating is introduced in a waveguide structure, the grating can be treated as a perturbation of an otherwise uniform waveguide. In this diagram $\delta\varepsilon(x,y,z)$ represents the perturbation function and n(x,y) represents the unperturbed uniform waveguide.

respond to forward travelling modes. The forward and backward modes are conventionally related to each other in the following way:

$$\begin{aligned} &\beta_{-n} = -\beta_n \\ &\mathbf{e}_{-n\perp} = \mathbf{e}_{n\perp} \quad \mathbf{h}_{-n\perp} = \mathbf{h}_{n\perp} \\ &e_{-n||} = -e_{n||} \quad h_{-n||} = h_{n||} \end{aligned} \tag{4.23}$$

[15] The magnetic field of the backward travelling (–n) mode is the same as for the forward travelling mode. The electric field of the backward travelling mode is the same as the corresponding forward travelling mode, with the exception that the perpendicular (x-y) components of **e** are reversed in direction.

In order to completely describe the electromagnetic fields of the waveguide, the superposition given in equation 4.22 should include an integration over the continuous spectrum of radiation modes in addition to a summation over bound modes. However, since we are interested only in the interactions between guided modes we have included only the bound modes. If the perturbation $\delta\varepsilon$ is sufficiently large, it is no longer valid to consider only bound modes, and coupling to radiation modes must also be included in the analysis [45].

Coupled mode theory seeks to replace Maxwell's equations with a set of coupled ordinary differential equations describing the coefficients $a_n(z)$. Consider the vector quantity **F**, defined as:

$$\mathbf{F} = \left(\mathbf{E} \times \mathbf{h}_m^* + \mathbf{e}_m^* \times \mathbf{H}\right) e^{-i\beta_m z} \tag{4.24}$$

$\mathbf{e_m}$ and $\mathbf{h_m}$ are the modal fields of the m^{th} eigenmode of the unperturbed waveguide. These eigenmodes are solutions to Maxwell's equations with the unperturbed index profile n(x,y). Similarly, **E** and **H** are the electromagnetic fields of the perturbed waveguide. If we apply the divergence theorem to the vector **F** over an infinitesimally thin slab which spans the x-y plane at a location z, we arrive at the identity:

$$\frac{\partial}{\partial z} \iint \mathbf{F} \cdot \hat{\mathbf{z}}\, dA = \iint \nabla \cdot \mathbf{F}\, dA \tag{4.25}$$

The integration in equation 4.25 is taken over the entire x-y plane at a location z.

The divergence term $\nabla \cdot \mathbf{F}$ in the left-hand side of equation 4.25 can be calculated using equation 4.24, noting that the fields satisfy Maxwell's equations 4.1:

$$\nabla \cdot \mathbf{F} = ik\sqrt{\frac{\varepsilon_0}{\mu_0}}\, \delta\varepsilon\, \mathbf{e}_m^* \cdot \mathbf{E}\, e^{-i\beta_m z} \tag{4.26}$$

In obtaining the previous result we have made use of the vector identity:

$$\nabla \cdot \left(\mathbf{A} \times \mathbf{B}\right) = \mathbf{B} \cdot \left(\nabla \times \mathbf{A}\right) - \mathbf{A} \cdot \left(\nabla \times \mathbf{B}\right) \tag{4.27}$$

Equation 4.25 then becomes:

$$\frac{\partial}{\partial z}\iint \mathbf{F}\cdot\hat{\mathbf{z}}\,dA = ik\sqrt{\frac{\varepsilon_0}{\mu_0}}\,e^{-i\beta_m z}\iint \delta\varepsilon\,\mathbf{e}_m^*\cdot\mathbf{E}\,dA \tag{4.28}$$

This relationship is sometimes referred to as the Lorentz reciprocity relation [15, 16, 46]. Note that the orthogonality relationship given earlier in equation 4.18 can be derived from equation 4.28 by taking $\delta\varepsilon = 0$ thus eliminating the right-hand side, and letting **E** and **H** be the n^{th} eigenmode of the waveguide.

The coupled mode equations are found by substituting the mode expansion 4.22 into equation 4.28, making use of the orthogonality relationship given in equation 4.18. This yields:

$$\left(\frac{d}{dz}-i\beta_m\right)a_m(z) = \pm\frac{ik}{4P_m}\sqrt{\frac{\varepsilon_0}{\mu_0}}\sum_n a_n(z)\iint \delta\varepsilon\,\mathbf{e}_n\cdot\mathbf{e}_m^*dA \tag{4.29}$$

where the upper sign is used when m represents a forward travelling mode and the lower sign is used when m is negative. Equation 4.29 describes a system of coupled first order ordinary differential equations which govern the expansion coefficients $a_m(z)$. Coupled mode theory effectively replaces Maxwell's equations which describe the electromagnetic fields with a series of coupled differential equations which describe how the expansion coefficients evolve. If we let $\delta\varepsilon = 0$ in equation 4.29, the right-hand side vanishes and the equations of motion become:

$$\left(\frac{d}{dz}-i\beta_m\right)a_m(z) = 0 \tag{4.30}$$

Thus, in the absence of a perturbation, the expansion coefficients are decoupled, and the solution for $a_m(z)$ is a travelling wave in the z-direction, as expected.

For a waveguide that supports only one bound mode, equation 4.29 reduces to a pair of coupled differential equations which relate the forward and backward travelling mode amplitudes, denoted $a_+(z)$ and $a_-(z)$:

$$\begin{aligned}\left(\frac{d}{dz}-i\beta\right)a_+(z) &= +ig(z)\big(a_+(z)+a_-(z)\big)\\ \left(\frac{d}{dz}+i\beta\right)a_-(z) &= -ig(z)\big(a_+(z)+a_-(z)\big)\end{aligned} \tag{4.31}$$

where β is the propagation constant of the forward travelling mode and g(z) is a real-valued, periodic function of z, given by:

$$g(z) = \frac{k}{4P}\sqrt{\frac{\varepsilon_0}{\mu_0}}\iint \delta\varepsilon(x,y,z)\left|\mathbf{e}(x,y)\right|^2 dA \tag{4.32}$$

For a Bragg grating, the perturbation $\delta\varepsilon(x,y,z)$ is a periodic function of z. For integrated Bragg gratings, this perturbation usually has a rectangular profile, as illustrated in Fig. 4.7, because the reactive

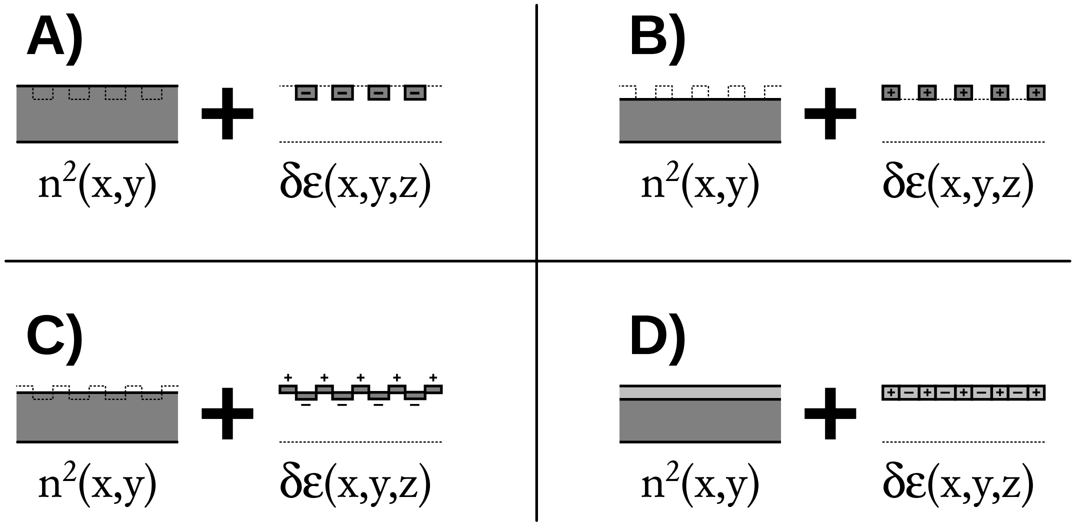


Figure 4.8 Four different ways to select the unperturbed waveguide geometry. The + and – symbols indicate where the perturbation function is positive and negative.

ion etching process typically used to form the grating is a directional etch which tends to form rectangular teeth. There is some ambiguity in how the unperturbed waveguide geometry is selected. Figure 4.8 illustrates four different ways to choose the unperturbed geometry. In Fig. 4.8A, the unperturbed waveguide is taken to be an un-etched guide, and therefore the perturbation function $\delta\varepsilon(x,y,z)$ is zero except within the etched groove regions where it is negative. It is equally valid to treat the unperturbed waveguide as a completely etched structure where $\delta\varepsilon(x,y,z)$ is zero everywhere except within the grating teeth, where it is positive, as shown in Fig. 4.8B. Some authors instead choose to place the unperturbed waveguide boundary somewhere in the center of the teeth, as shown in Fig. 4.8C [47]. Another reasonable choice of the unperturbed geometry is to treat the entire grating region as a uniform layer with an index of refraction somewhere between n_{clad} and n_{core} [48, 45, 49], as depicted in Fig. 4.8D. In this case, the perturbation function $\delta\varepsilon(x,y,z)$ alternates between positive and negative values. For shallow gratings the resulting coupled mode equations and coupling coefficients are independent of the choice of the unperturbed waveguide geometry. In the analysis which follows, the unperturbed waveguide is chosen to be an un-etched waveguide as in Fig. 4.8A. Figure 4.7 illustrates the function $\delta\varepsilon(x,y,z)$ for rectangular teeth etched into a channel waveguide. For convenience, we have chosen the axes so that the grating is symmetric about the z = 0 plane. The symbol D denotes the duty cycle of the grating, which is the ratio of the tooth width to the grating period Λ. Therefore, the function g(z) in equation 4.32 can be expressed as:

$$g(z) = \begin{cases} K & |z| < \frac{\Lambda}{2}(1-D) \\ 0 & \frac{\Lambda}{2}(1-D) < |z| < \frac{\Lambda}{2} \\ g(z-n\Lambda) & \text{elsewhere} \end{cases} \tag{4.33}$$

where the quantity K is defined as:

$$K = -\frac{k}{4P}\left(n_{core}^2 - n_{clad}^2\right)\sqrt{\frac{\varepsilon_0}{\mu_0}} \iint_{grating} \left|\mathbf{e}(x,y)\right|^2 dA \tag{4.34}$$

The integral in equation 4.34 is carried out only over the cross sectional region of the waveguide where the grating is located. Outside of this region, the function $\delta\varepsilon(x,y,z)$ is zero.

To simplify the coupled mode equations, it is useful to expand the coupling coefficient g(z) in terms of its Fourier components:

$$g(z) = \sum_{n=-\infty}^{\infty} g_n e^{in2\pi\frac{z}{\Lambda}} \tag{4.35}$$

$$g_n = \frac{1}{\Lambda}\int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} g(z) e^{-in2\pi\frac{z}{\Lambda}}\, dz \tag{4.36}$$

Evaluating the Fourier coefficients in equation 4.36 gives:

$$g_n = \begin{cases} K(1-D) & n = 0 \\ -\dfrac{K}{n\pi}(-1)^n \sin(n\pi D) & n \neq 0 \end{cases} \tag{4.37}$$

Substituting the Fourier expansion of equation 4.35 into the coupled mode equations 4.31 gives:

$$\begin{aligned} \left(\frac{d}{dz} - i\beta\right) a_+(z) &= +ig_0 a_+(z) + ig_1 a_-(z) e^{+i2\pi\frac{z}{\Lambda}} + \ldots \\ \left(\frac{d}{dz} + i\beta\right) a_-(z) &= -ig_0 a_-(z) - ig_{-1} a_+(z) e^{-i2\pi\frac{z}{\Lambda}} + \ldots \end{aligned} \tag{4.38}$$

In the above equation, only the significant Fourier terms have been included in each of the coupled mode equations. The effect of the g_0 terms in equation 4.38 is to slightly modify the propagation constant of the waveguide, so that the perturbed waveguide behaves as if it had a propagation constant β_{new} given by:

$$\beta_{new} = \beta_{old} + g_0 \tag{4.39}$$

Often the unperturbed waveguide geometry is deliberately chosen so that the 0^{th} Fourier coefficient is zero, in which case there is no first-order change in the propagation constant of the waveguides. In either case, equation 4.38 reduces to the form:

$$\left(\frac{d}{dz} - i\beta\right)a_+(z) = +ig_1 a_-(z)e^{+i2\pi\frac{z}{\Lambda}}$$
$$\left(\frac{d}{dz} + i\beta\right)a_-(z) = -ig_{-1} a_+(z)e^{-i2\pi\frac{z}{\Lambda}} \quad (4.40)$$

In the absence of a grating, the solution for $a_+(z)$ is $e^{+i\beta z}$ and the solution for $a_-(z)$ is $e^{-i\beta z}$. If the grating period is selected so that $\pi/\Lambda \approx \beta$, the grating can cause coupling between the forward and backward modes. The condition that $\pi/\Lambda \approx \beta$ is known as the Bragg condition, as described in Chapter 3. When the amplitude function $a_-(z)$ is multiplied by the $n = 1$ term from the Fourier expansion of $g(z)$, the resulting product has roughly the same periodicity as $a_+(z)$, provided the Bragg condition is satisfied. Likewise, when the function $a_+(z)$ is multiplied by the $n = -1$ Fourier component, the resulting product has roughly the same periodicity as $a_-(z)$. All of the terms which have been left out of equation 4.38 have a spatial periodicity which is much different from the accompanying terms in the differential equation. The omitted Fourier expansion terms do not meet the Bragg condition, meaning they oscillate with spatial frequencies which are much different from $e^{\pm i\beta z}$. When integrated over z, the omitted terms would average to zero, whereas the retained terms which meet the Bragg condition would not. Therefore, we are justified in neglecting all of the extra perturbation terms in the right hand side of equation 4.38 which are not synchronous.

To complete the analysis, we define the coupling constant κ in the following way:

$$\kappa = ig_1 = -\frac{ik}{4\pi}\left(n_{core}^2 - n_{clad}^2\right)\sin(\pi D)\sqrt{\frac{\varepsilon_0}{\mu_0}}\,\frac{1}{P}\iint_{grating}\left|\mathbf{e}(x,y)\right|^2 dA \quad (4.41)$$

Substituting equation 4.41 into equation 4.40 gives:

$$\left(\frac{d}{dz} - i\beta\right)a_+(z) = \kappa a_-(z)e^{+i2\pi\frac{z}{\Lambda}}$$
$$\left(\frac{d}{dz} + i\beta\right)a_-(z) = \kappa^* a_+(z)e^{-i2\pi\frac{z}{\Lambda}} \quad (4.42)$$

The phase of the coupling constant κ in equation 4.41 depends upon how the grating is placed relative to the $z = 0$ origin. In Fig. 4.7, the origin has been chosen so that the grating perturbation is symmetric about $z = 0$. If, for example, the origin were instead placed at the leading edge of a tooth, the value of κ would have the same magnitude as in equation 4.41, but would no longer be purely imaginary. However, provided the perturbation $g(z)$ is a real valued periodic function of z, the Fourier coefficients will be related by $g_{-n} = (g_n)^*$, and therefore the coupled mode equations will still be of the form given in equation 4.42. Henceforth, we will simplify the analysis by treating κ as a real quantity, recognizing that the phase of κ is arbitrarily determined by the relative position of the $z = 0$ origin. If the grating includes a change in the optical gain or loss of the material, the perturbation $\delta\varepsilon$ can no longer be adequately modeled by a real-valued function, and the coupled mode equations will be different from those given in equation 4.42 [49, 50].

As discussed in Section 4.1.3, in weakly guiding structures the bound modes can be sufficiently described with a single scalar quantity $\Phi(x,y)$, representing one of the transverse field components of the mode. If $\Phi(x,y)$ represents the transverse electric field of a mode, then the bound power P of the mode is given by:

$$P = \frac{\beta}{2k}\sqrt{\frac{\varepsilon_0}{\mu_0}}\iint \left|\Phi(x,y)\right|^2 dA \tag{4.43}$$

By substituting the above equation into equation 4.41 (and suppressing the arbitrary phase) we obtain the following expression for the grating strength κ in terms of the scalar field $\Phi(x,y)$:

$$\kappa = \frac{k^2}{2\pi\beta}\left(n_{core}^2 - n_{clad}^2\right)\sin(\pi D)\Gamma \tag{4.44}$$

where the quantity Γ is a dimensionless overlap integral given by:

$$\Gamma = \frac{\iint\limits_{grating} \left|\Phi(x,y)\right|^2 dA}{\iint \left|\Phi(x,y)\right|^2 dA} \tag{4.45}$$

The factor Γ defined in equation 4.45 is a measure of how much the optical mode overlaps with the grating region. The integral in the numerator of equation 4.45 is carried out only over the grating region whereas the integral in the denominator extends over the entire cross-sectional plane.

It can be shown that the four different choices for an unperturbed waveguide geometry illustrated in Fig. 4.8 *all* yield the same expression for the grating strength, given by equations 4.44 and 4.45 [45, 47, 48]. The differences between these formulations reside completely within the unperturbed optical mode $\Phi(x,y)$, which effects the overlap factor Γ through equation 4.45. For small perturbations, the waveguide modes are not modified significantly and therefore Γ depends only weakly on the choice of an unperturbed waveguide geometry.

Figure 4.9 plots the calculated grating strength as a function of grating etch-depth for the glass channel waveguide structure depicted in Fig. 4.6. These data were calculated using the mode overlap technique described above. The scalar mode profile was numerically calculated using a finite difference modesolver, and the grating overlap factor Γ was approximated by a discrete summation over the grating cross-section. The duty cycle of the grating was taken to be 50%; the effect of a different duty cycle can easily be calculated using equation 4.44. The shaded region in Fig. 4.9 indicates the range of grating strengths and etch depths which will be used to construct the matched filters. This range of values was selected to yield a κL product which is between 0.5 and 0.7, assuming a grating length of 1 cm. As

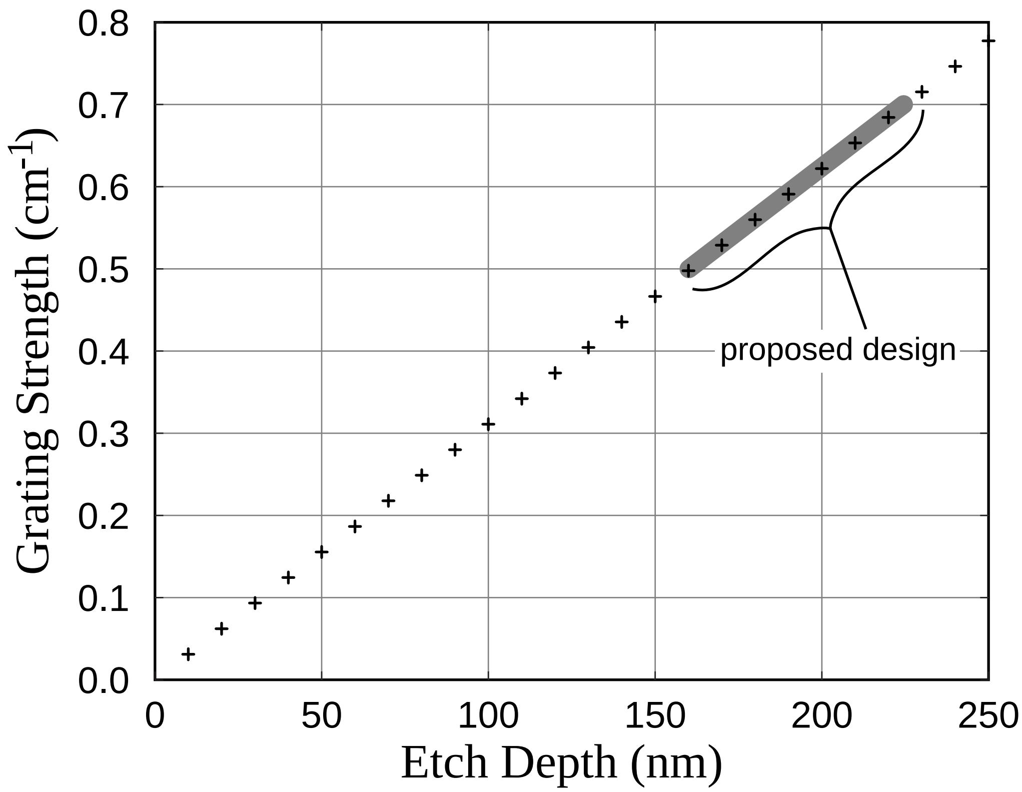


Figure 4.9 Calculated grating strength as a function of etch depth for the channel waveguide structure illustrated in Fig. 4.6 on page 58. In these calculations, the grating profile was assumed to be rectangular with a 50% duty cycle.

described in Section 3.2.2, this should yield a filter which is reasonably well matched to a binary communications signal operating at 10 Gb/s.

4.4 Summary

A waveguide is a dielectric structure which is used to confine and direct light. The guiding properties of a waveguide are usually described by the bound modes of the structure. Waveguides are characterized by a series of discrete bound modes which are eigenmodes of the wave equation, each with a characteristic eigenvalue or propagation constant. The phenomenon of bound optical modes in a waveguide is analogous to the well known phenomenon of quantized energy eigenstates in a quantum mechanical potential well. In Section 4.1.2 we used the planar slab waveguide as a simple, tractable example to illustrate some of the important features that are common to all waveguide structures.

Waveguides come in many shapes and sizes. The most common optical waveguide is the optical fiber which has already found widespread use in long distance communications. Planar-based waveguides differ from optical fiber in that they are fabricated on planar substrates using lithography techniques. Such planar-based devices are important because they offer the potential for integration with other electronic or optical components of the communications system on a single optoelectronic integrated chip. There are many candidate materials for building integrated optical waveguides, including compound semiconductor materials (InP, GaAs), glass materials (SiO_2, GeO_2), polymers, and other single-crystal dielectric materials ($LiNbO_3$). Each of these materials has advantages and disadvantages in terms of the intrinsic material loss, cost, ease of fabrication, availability of deposition or growth technologies, and integrability. We have chosen to use SiO_2 planar waveguides with GeO_2 doped cores because these materials are commercially available, have low intrinsic material loss, and can provide good coupling to optical fiber.

In a uniform waveguide, the bound modes propagate independently without intermixing. When the waveguide structure is slightly perturbed, the result is that the previously decoupled modes of the waveguide become coupled. Coupled mode theory is a mathematical tool that is used to describe the coupling between modes as a result of a waveguide perturbation. In Section 4.3, we described how a periodic perturbation in a waveguide structure can couple the forward and backward travelling modes of the waveguide, and presented a mathematical derivation of the coupling coefficient κ for the case of a rectangular grating profile.

CHAPTER 5: CODIRECTIONAL WAVEGUIDE COUPLING

5.1 *Redirecting Filtered Signal*

In Chapter 3, we described how the reflection spectral response of a weak Bragg grating can be made to match the spectrum of a signal carrying binary information. A uniform Bragg grating is inherently a reflection filter, reflecting a range of frequencies about the Bragg frequency and transmitting frequencies outside that range. In a practical communications system, such a reflection filter is not desirable because the filtered signal emerges from the device in the same port that is used for the input signal. One reason why reflection in the input port is unacceptable is because unwanted reflections could cause feedback in the erbium doped fiber amplifiers which are often used in communications systems prior to detection. Such feedback could cause the fiber amplifier to lase. When a grating is used in a reflection configuration, another device is needed to redirect the reflection signal, preferably without disturbing the input signal or losing optical power.

The simplest device which allows access to the filtered signal is a Y-branch/isolator configuration, depicted in Fig. 5.1A. The reflected signal from the grating is split equally between two branches. An isolator is an optical device which allows light to pass through in one direction while blocking light in the reverse direction. An isolator placed in one branch of the device allows light to be sent into the grating while blocking the reflected signal. The output signal is then collected in the other branch. One obvious drawback of using the Y-branch/isolator configuration is that only 50% of the filtered signal is collected, while the remaining 50% is discarded in the isolator, assuming that the Y-branch operates under ideal conditions.

A more sophisticated device which can be used to separate the reflected signal from the incident signal is the optical circulator, depicted schematically in Fig. 5.1B. The circulator is a bulk-optical device

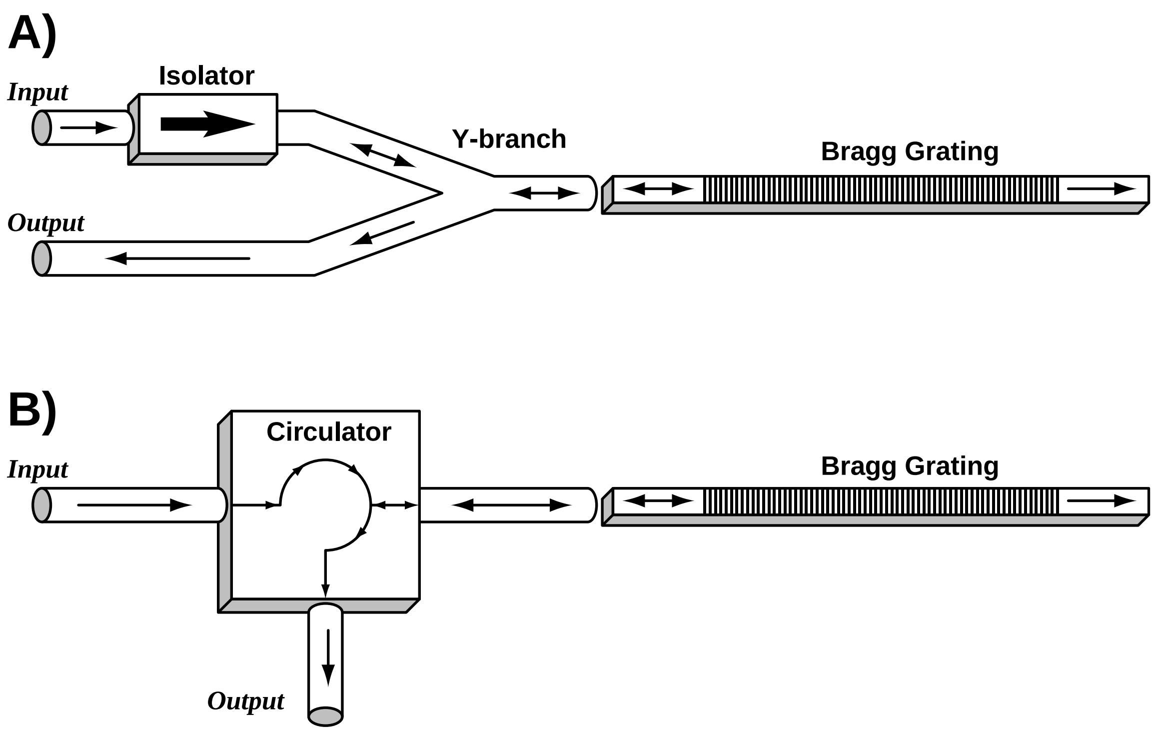


Figure 5.1 Two configurations which allow access to a reflected signal from a Bragg grating. A) The reflected signal from the grating is split with a beamsplitter, and 50% of the reflected signal is therefore discarded by the isolator. B) A circulator effectively redirects all of the reflected signal.

which redirects the reflected light to a separate port, thereby separating the reflected signal from the input signal without sacrificing any significant amount of the output signal. [51]

Circulators and isolators operate on the principle of wave propagation in birefringent and non-reciprocal media. Accordingly, they are not well suited for on-chip integration with other integrated optical components, such as waveguides and gratings, which are typically constructed from linear, isotropic dielectric materials. Commercially available isolators and circulators are designed to be discrete optical components which are connected directly to optical fibers. [52]

One approach which allows for optical integration is to use an integrated Michelson interferometer configuration, as depicted in Fig. 5.2A. In the Michelson interferometer, the input signal in the upper waveguide is split equally between the upper and lower arms by an integrated coupler. The signals in each arm reflect from two identical Bragg grating filters. If the arms of the interferometer are precisely the same length, the reflected signals will be recombined in the lower waveguide, as depicted [53, 54]. Figure 5.2B depicts the analogous free-space propagating version of the Michelson interferometer. In this device, the incident signal is split between two paths with a loss-less, partially reflective mirror (often termed a beamsplitter). When the two arm lengths in the interferometer are precisely matched (or mismatched by an integral number of half-wavelengths) the reflected signals recombine in a different port of the device* [55, 24]. In the integrated optical device depicted in Fig. 5.2A, the coupler serves the same purpose as the beamsplitter in Fig. 5.2B, and the integrated Bragg grating is analogous to the mirrors.

Chapters 3 and 4 describe the physics of optical waveguides, and the principle of Bragg reflection from an integrated grating. This chapter focuses on the design of the codirectional coupler, which is used for splitting light between two waveguides. The codirectional coupler enables the output signal to be separated from the input signal by connecting the gratings in a Michelson interferometer configuration as shown in Fig. 5.2A.

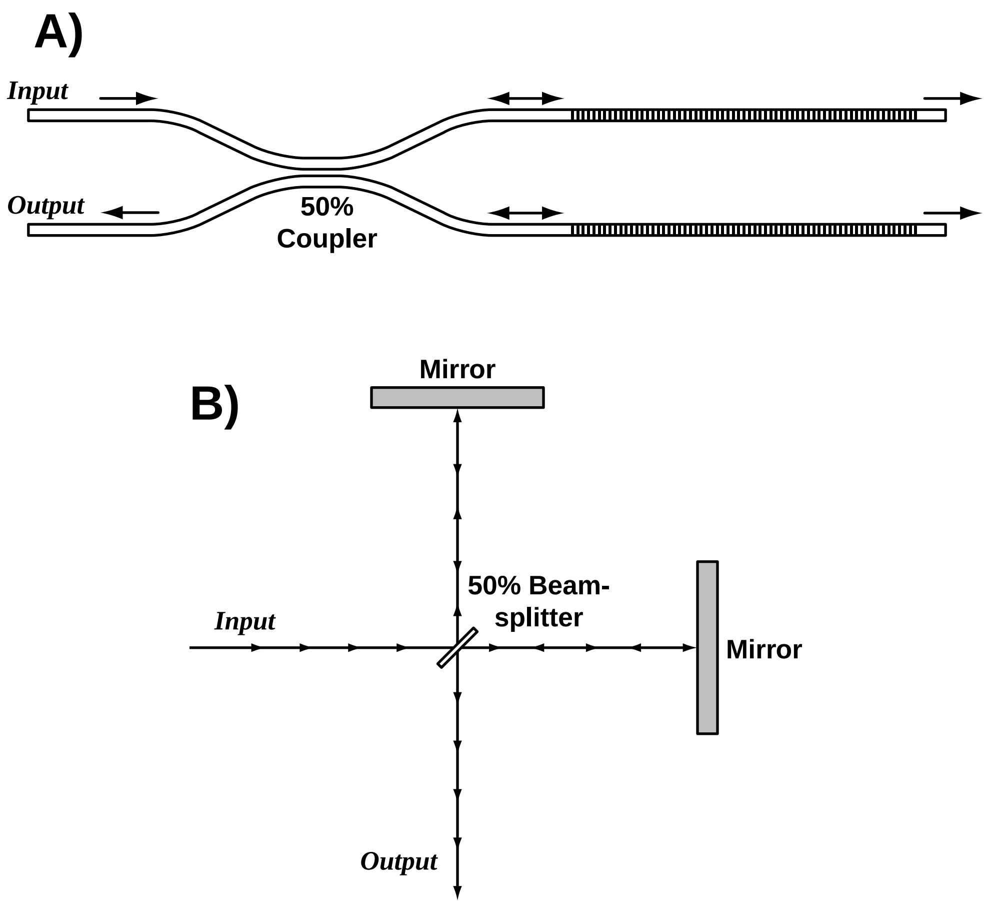


Figure 5.2 The Michelson interferometer configuration depicted in A provides a way of redirecting the reflected filtered output signal to a different port of the device, in a way that can be integrated on a photonic chip with the waveguides. The analogous free-space Michelson interferometer is depicted in B.

* The analogy between the integrated Michelson interferometer and its free-space equivalent is valid only if we assume that the input signal in Fig. 5.2B is a perfect plane wave. If the input signal is not a plane wave, the reflected output will form an interference patter consisting of concentric circular fringes.

5.2 *Codirectional Coupled Mode Equations for Parallel Waveguides*

When two waveguides are brought close together, coupling occurs between the waveguides, and power can be transferred from one guide to the other. The phenomenon of codirectional waveguide coupling is analogous to the motion of two coupled oscillators, such as the two coupled pendulums illustrated in Fig. 5.3. When there is no spring connecting the two pendulums in Fig. 5.3, they oscillate independently of each other. If the masses and lengths are identical, the characteristic oscillation frequency is the same for both pendulums. However, when a spring is added between the two, the pendulums are no longer decoupled. If, for example, at time $t = 0$, the right pendulum is at rest in its relaxed position while the left pendulum is in motion, the energy will be observed to slosh back and forth

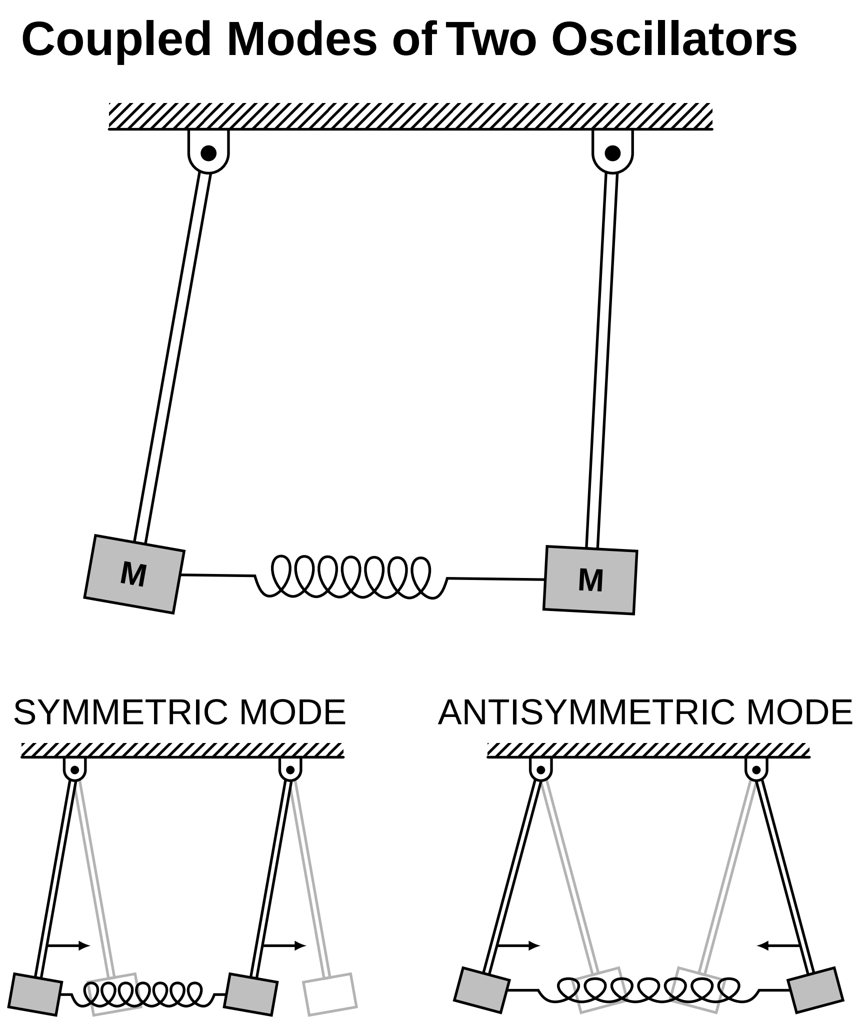


Figure 5.3 The phenomenon of light coupling between proximate waveguides is analogous to the coupling of two harmonic oscillators, such as the connected pendulums depicted here. The motion of the coupled pendulums can be analyzed by finding the symmetric and antisymmetric modes of oscillation of the structure. The motion of the system can be described as a superposition of the symmetric and antisymmetric oscillation modes.

between the pendulums, gradually moving from left to right and back so that at some later time T, the left pendulum will be at rest while the right one is in motion. After a time of 2T, all of the energy will again be transferred back to the left pendulum.

The motion of the coupled pendulums can be analyzed in terms of the normal modes of oscillation of the system. As depicted in the lower portion of Fig. 5.3, there are two normal modes of oscillation for the coupled pendulums: the symmetric mode in which both pendulums oscillate in the same direction, and the antisymmetric mode in which the pendulums oscillate in opposite directions. (Note that if the pendulums were dissimilar, the normal modes of the system would no longer be strictly symmetric or antisymmetric, but the system would nonetheless have two normal modes.) When the system is in one of its normal modes, all parts of the system oscillate in synchronism with the same frequency. When the spring connecting the pendulum is very weak, the oscillation frequency for the antisymmetric mode is only slightly higher than for the symmetric mode, and both mode frequencies are close to the oscillation frequency of the decoupled pendulums. In this case, the spring can be treated as a small perturbation which introduces coupling between the motion of two otherwise independent pendulums. The most general description of the motion of the coupled pendulums would consist of a superposition of the two normal modes of the system. If the frequencies of these two modes are similar, as is the case for a weak spring, the resulting motion of the system will be almost periodic motion at the fundamental frequency. However, energy will gradually slosh back and forth between the two pendulums. The frequency of this gradual sloshing, called the beat frequency, is simply the difference between the symmetric and antisymmetric oscillation frequencies. To design a device which would transfer part of the energy from the left pendulum to the right pendulum (or vice versa) we could place a spring between the two pendulums at $t = 0$ and abruptly disconnect the spring at some later time after the desired amount of energy had been transferred to the second pendulum.

The coupling between two optical waveguides is analogous to the coupling between two harmonic oscillators. Whereas for coupled pendulums the amount of coupling depends upon the parameters of the connecting spring, for parallel waveguides the amount of coupling depends upon the proximity of the waveguides to one another and the electromagnetic mode profiles of the waveguides. The most rigorous way to analyze two parallel waveguides is to solve for the symmetric and antisymmetric modes of the coupled guides, using the same techniques that are employed to solve for the electromagnetic modes of an isolated waveguide. Figure 5.4 is a gray-scale plot of the field amplitudes for the symmetric and antisymmetric modes of two parallel identical channel waveguides, calculated using Finite Difference software [27]. (As with the coupled pendulums, if the two waveguides are not identical, the two bound modes of the structure will not have definite symmetry.) The symmetric mode has a propagation constant β_s which is slightly greater than that of an isolated waveguide, while the antisymmetric mode has a propagation constant β_a which is slightly less than that of an isolated waveguide. The difference between the symmetric and antisymmetric propagation constants describes the beat frequency of the structure, or the rate at which power gradually sloshes from one guide to the other. Consider, for

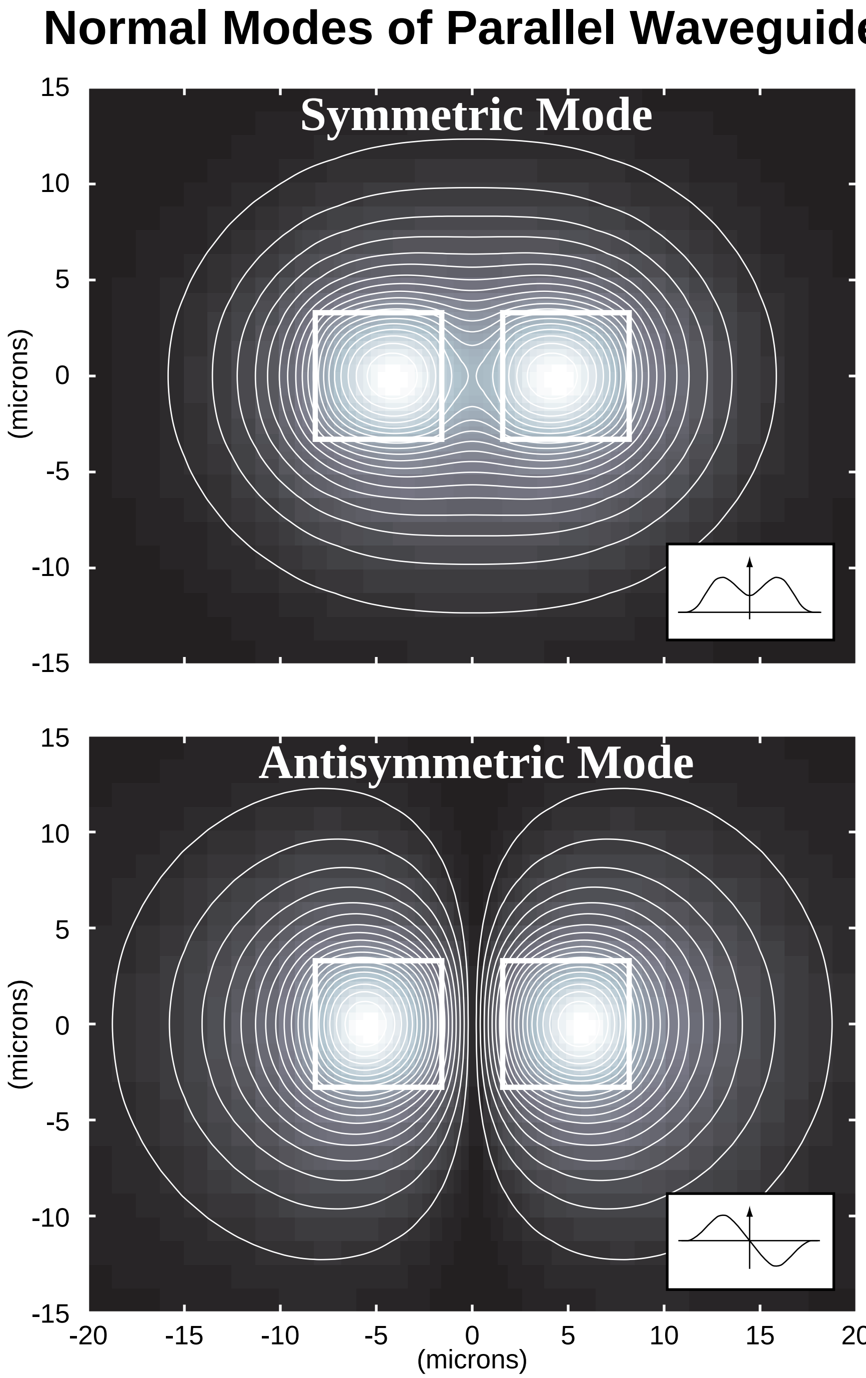


Figure 5.4 Mode plots of the symmetric (upper plot) and antisymmetric (lower plot) modes of two closely spaced channel waveguides. The grayscale level is proportional to the field amplitude; field contours are also superposed for reference. The core regions are outlined in white. As with the coupled pendulums, the optical coupling between two waveguides can be described in terms of the symmetric and antisymmetric modes of the structure. The symmetric mode has a propagation constant that is slightly higher than the antisymmetric mode, and the difference between these propagation constants describes the rate at which power is transferred between the guides.

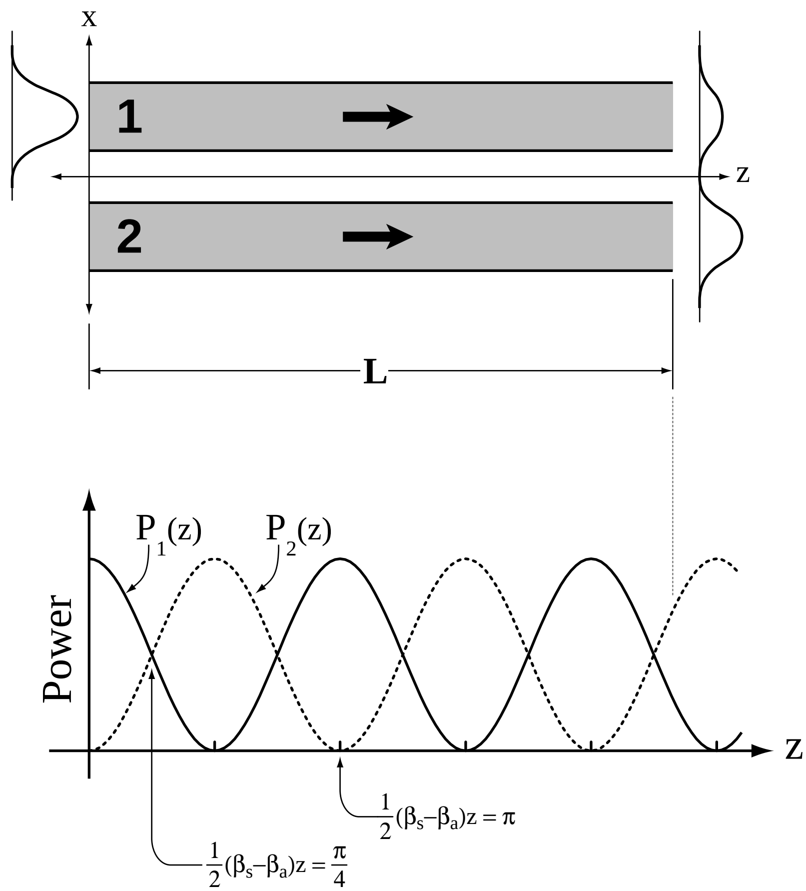


Figure 5.5 A simple codirectional coupler consisting of two parallel waveguides. Waveguide 1 is illuminated at z = 0, exciting an even superposition of the symmetric and antisymmetric modes. At some distance L, the power is divided between waveguides 1 and 2. The lower portion of the figure plots the power in each guide as a function of distance. The periodicity of the oscillations is given by the difference between the symmetric and antisymmetric propagation constants.

example, the two parallel waveguides illustrated in Fig. 5.5, where at z = 0 light is launched into the upper waveguide. The initial mode excitation of the upper guide (denoted guide 1) at the z = 0 plane can be described as an equal superposition of the symmetric and antisymmetric modes of the structure:

$$\phi(x, z = 0) = \phi_1(x) = \frac{1}{\sqrt{2}}(\phi_s(x) + \phi_a(x)) \tag{5.1}$$

where $\phi_s(x)$ represents the symmetric mode and $\phi_a(x)$ represents the antisymmetric mode. The symmetric and antisymmetric modes propagate in the positive z direction, with slightly different propagation constants β_s and β_a respectively, so that at some later position z, initial mode excitation has evolved to:

$$\phi(x,z) = \frac{1}{\sqrt{2}}(\phi_s(x)e^{i\beta_s z} + \phi_a(x)e^{i\beta_a z}) \tag{5.2}$$

If we then re-express the symmetric and antisymmetric modes in terms of the modes of the isolated waveguides $\phi_1(x)$ and $\phi_2(x)$, equation 5.2 becomes*:

$$\phi(x,z) = e^{+\frac{i}{2}(\beta_s+\beta_a)z}\left[\cos\left(\frac{1}{2}(\beta_s-\beta_a)z\right)\phi_1(x) + i\sin\left(\frac{1}{2}(\beta_s-\beta_a)z\right)\phi_2(x)\right] \tag{5.3}$$

The amount of power contained in the lower waveguide is then given directly by the square of the coefficient of $\phi_2(x)$ in the above equation:

$$P_2(z) = \sin^2\left(\frac{1}{2}(\beta_s-\beta_a)z\right) \tag{5.4}$$

The procedure described above for calculating the behavior of two parallel waveguides is straightforward, but can be computationally difficult, especially when the waveguides are separated by a large distance. At large separations, the symmetric and antisymmetric modes have almost identical propagation constants. Therefore, obtaining an accurate measure of the beat frequency requires that the propagation constants be calculated to high precision. Beyond a certain waveguide separation, the accuracy of the difference in propagation constant is fundamentally limited by the machine precision used to store the symmetric and antisymmetric propagation constants. In this regime, it is more accurate and efficient to treat the coupled waveguides using a perturbation approach, in which the presence of a second waveguide is treated as a perturbation of the mode of an isolated waveguide. The same coupled mode formulation derived in Chapter 4 to describe an integrated Bragg grating can also be employed to describe the coupling of parallel waveguides.

In order treat the interaction between parallel waveguides using coupled mode theory, it is necessary to treat the unperturbed structure as an isolated guide, and the second waveguide as a perturbation $\delta\varepsilon(x,y)$, as illustrated in Fig. 5.6. We shall denote the electromagnetic fields of waveguide 1 (in the absence of waveguide 2) as $\mathbf{e_1}$ and $\mathbf{h_1}$, and those of waveguide 2 (in the absence of waveguide 1) as $\mathbf{e_2}$ and $\mathbf{h_2}$. The electric field of the parallel waveguides is then written as a linear superposition of these two mode profiles:

$$\begin{aligned}\mathbf{E}(z) &= a_1(z)\mathbf{e}_1(x,y) + a_2(z)\mathbf{e}_2(x,y)\\ \mathbf{H}(z) &= a_1(z)\mathbf{h}_1(x,y) + a_2(z)\mathbf{h}_2(x,y)\end{aligned} \tag{5.5}$$

* Note that it is not strictly correct to write the symmetric and antisymmetric modes as symmetric and antisymmetric combinations of the isolated waveguide modes. However, when the waveguides are sufficiently separated that the evanescent field overlap is small, the rigorously calculated symmetric and antisymmetric modes become indistinguishable from symmetric and antisymmetric combinations of the isolated waveguide modes.

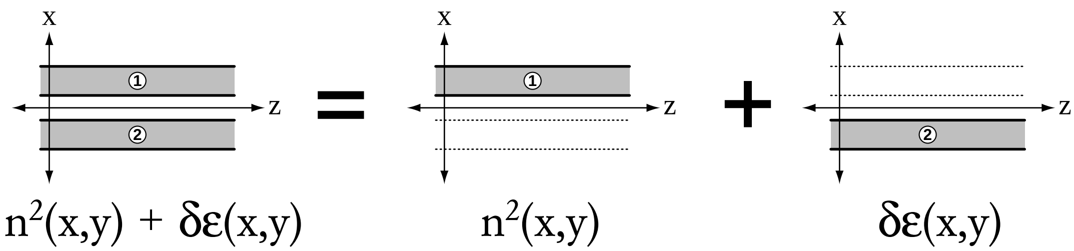


Figure 5.6 In analyzing two parallel waveguides with coupled mode theory, the presence of the second waveguide is treated as a perturbation of the mode of the first waveguide.

As in Chapter 4, the goal of coupled mode theory is to derive a set of differential equations describing the evolution of the expansion coefficients $a_1(z)$ and $a_2(z)$. The coupled mode equations follows directly from equation 4.29, however for this case the perturbation function $\delta\varepsilon(x,y)$ is independent of z.

$$\left(\frac{d}{dz} - i\beta_m\right)a_m(z) = \frac{ik}{4P_m}\sqrt{\frac{\varepsilon_0}{\mu_0}}\sum_n a_n(z)\iint \delta\varepsilon_m\, \mathbf{e}_n \cdot \mathbf{e}^*_m dA \tag{5.6}$$

The above equation is identical to equation 4.29 with the following provisos:

1. The subscripts n and m are understood to number the waveguides in the structure, rather the orthogonal modes of the structure. Therefore, the field $\mathbf{e_m}$ denotes the optical mode of the m^{th} waveguide, in the absence of all other waveguides.
2. The perturbation function $\delta\varepsilon(x,y)$ is given a subscript according to which guide is being considered. For example, in describing the evolution of $a_1(z)$, waveguide 2 is treated as a perturbation whereas in the differential equation for $a_2(z)$, waveguide 1 is treated as the perturbation.
3. The ± symbol has been suppressed because for now we are considering only the codirectional coupling of forward travelling modes.

For the case of two identical parallel waveguides, the coupled mode equations 5.6 simplify to:

$$\begin{aligned}\left(\frac{d}{dz} - i\beta\right)a_1(z) &= i\mu a_2(z)\\ \left(\frac{d}{dz} - i\beta\right)a_2(z) &= i\mu a_1(z)\end{aligned} \tag{5.7}$$

Where the coupling coefficient μ is given by [16]:

$$\mu = \frac{k}{4P}\sqrt{\frac{\varepsilon_0}{\mu_0}}\left(n^2_{core} - n^2_{clad}\right)\iint_{\text{guide 1}} \mathbf{e}_1 \cdot \mathbf{e}^*_2\, dA \tag{5.8}$$

and the propagation constant β is given by:

$$\beta = \beta_0 + \frac{k}{4P}\sqrt{\frac{\varepsilon_0}{\mu_0}}\left(n_{core}^2 - n_{clad}^2\right)\iint\limits_{\text{guide 1}} \mathbf{e}_2 \cdot \mathbf{e}_2^* \, dA \tag{5.9}$$

where β_0 is the propagation constant of the isolated waveguide. The integration in equations 5.8 and 5.9 is performed only over the core region of waveguide 1, although it should be pointed out that because of the symmetry of the problem, each equation could be expressed as an equivalent integral over guide 2. Notice that the presence of a second waveguide introduces a coupling between the mode amplitudes $a_1(z)$ and $a_2(z)$, and it also slightly changes the propagation constant β of the guides. For the case of weakly guiding waveguides, the coupling coefficient μ defined in equation 5.8 can be simplified to:

$$\mu = \frac{k^2}{2\beta}\left(n_{core}^2 - n_{clad}^2\right)\tilde{\mu} \tag{5.10}$$

where $\tilde{\mu}(z)$ is the normalized dimensionless coupling coefficient given by:

$$\tilde{\mu} \equiv \frac{\displaystyle\iint\limits_{\text{guide 1/2}} \Phi\left(x - \frac{d}{2}, y\right)\Phi\left(x + \frac{d}{2}, y\right) dA}{\displaystyle\iint \left|\Phi(x,y)\right|^2 dA} \tag{5.11}$$

where $\Phi(x,y)$ is the scalar mode profile of a single isolated guide and d is the center to center waveguide separation. The integral in the numerator of equation 5.11 is performed over the cross-section of either waveguide core, and the integral in the denominator is performed over the entire x–y plane. Figure 5.7 depicts a contour plot of the isolated waveguide modes for two waveguides separated by a distance d. The coupling coefficient μ is determined by calculating the normalized overlap integral of these two modes over the shaded core region of either guide.

The coupled mode equations 5.7 can be cast into the matrix form:

$$\frac{d}{dz}\begin{bmatrix} a_1(z) \\ a_2(z) \end{bmatrix} = \begin{bmatrix} i\beta & i\mu \\ i\mu & i\beta \end{bmatrix}\begin{bmatrix} a_1(z) \\ a_2(z) \end{bmatrix} \tag{5.12}$$

Which can be solved by eigenvalue decomposition. The eigenvalues and associated eigenvectors of the system are:

$$\begin{aligned} i\beta_s &= i(\beta + \mu) \quad & \mathbf{v}_s &= \frac{1}{\sqrt{2}}\begin{bmatrix} +1 \\ +1 \end{bmatrix} \\ i\beta_a &= i(\beta - \mu) \quad & \mathbf{v}_a &= \frac{1}{\sqrt{2}}\begin{bmatrix} +1 \\ -1 \end{bmatrix} \end{aligned} \tag{5.13}$$

We have used the subscripts s and a to label the symmetric and antisymmetric eigenvalues and eigenvectors. As described earlier, it is also possible to solve directly for the symmetric and antisymmetric modes

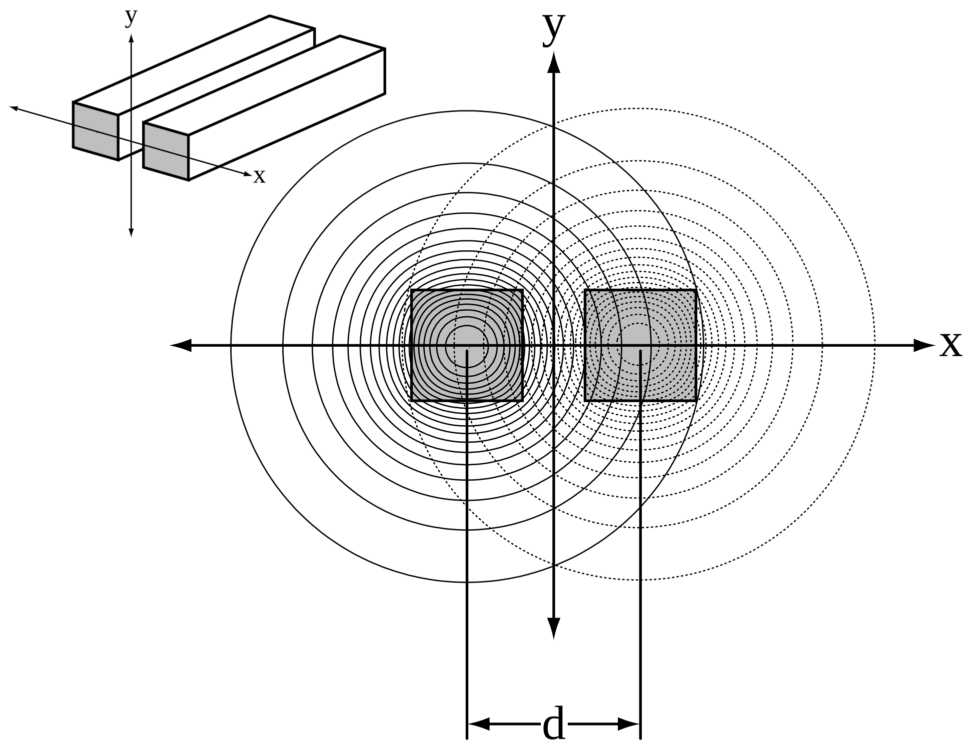


Figure 5.7 Mode contours of two parallel waveguides, separated by a center-to-center distance d. The mode contours have been calculated for the isolated waveguides. The coupling between the two can be found by calculating the overlap integral of the product of the isolated mode profiles over the cross section of either waveguide.

of the entire structure. Coupled mode theory gives the expected result that the eigenmodes of two identical parallel waveguides can be described to lowest order by the symmetric and antisymmetric combination of the modes of the isolated waveguides. The coupling coefficient μ can therefore be associated with the difference between the symmetric and antisymmetric propagation constants:

$$\mu = \frac{1}{2}(\beta_s - \beta_a) \tag{5.14}$$

The solution to the coupled mode equations 5.12 for $a_1(z)$ and $a_2(z)$ is:

$$\begin{bmatrix} a_1(z) \\ a_2(z) \end{bmatrix} = e^{+i\beta z} \begin{bmatrix} \cos(\mu z) & i\sin(\mu z) \\ i\sin(\mu z) & \cos(\mu z) \end{bmatrix} \begin{bmatrix} a_1(0) \\ a_2(0) \end{bmatrix} \tag{5.15}$$

Therefore, if initially (at $z = 0$) all of the power resides in waveguide 1, then the relative power in waveguide 2 as a function of position is:

$$\frac{P_2(z)}{P_1(0)} = \left| \frac{a_2(z)}{a_1(0)} \right|^2 = \sin^2(\mu z) \tag{5.16}$$

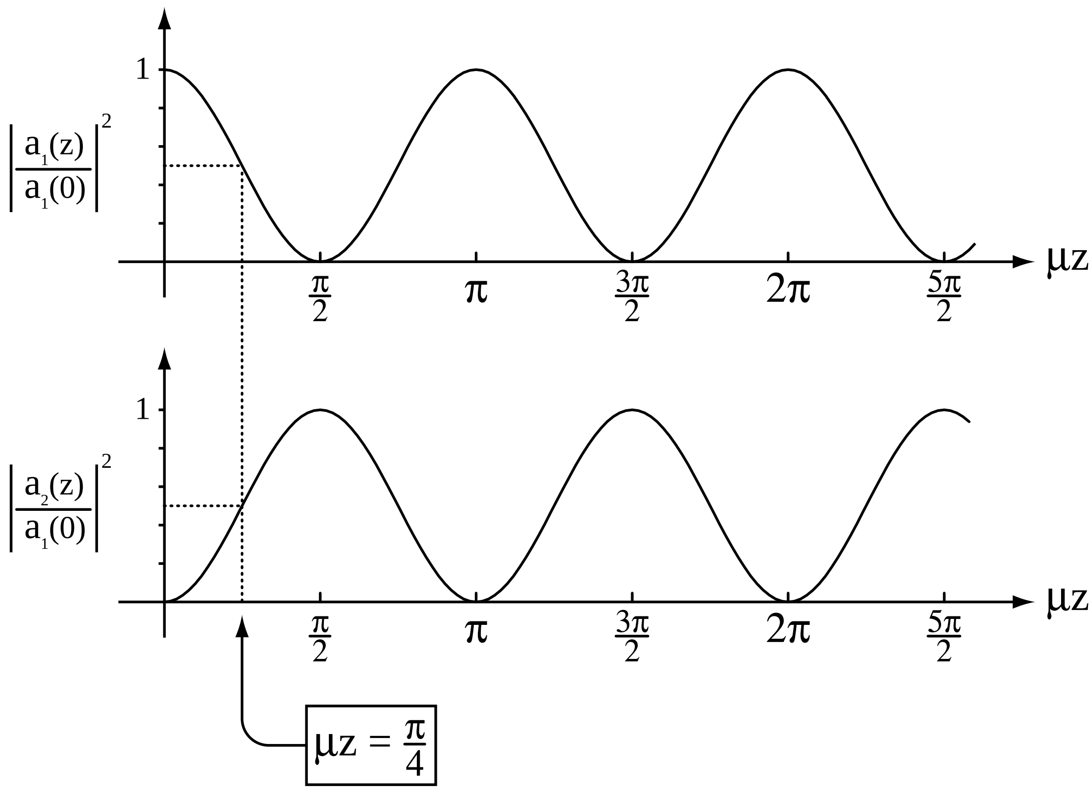


Figure 5.8 The power in each waveguide as a function of distance. This plot assumes that initially (at z = 0) all of the power is launched into waveguide 1. If the length of the coupler is selected so that $\mu z = \pi/4$, the power at the end of the coupler will be evenly split between the two waveguides.

Figure 5.8 plots the power in both waveguides as a function of z for the case where power is launched into waveguide 1 at z = 0. The sinusoidal oscillations describe the gradual sloshing of energy from one waveguide to another. If the length of the coupling region L and the coupling strength z are chosen so that

$$\mu L = \frac{\pi}{4} \tag{5.17}$$

then the power at the end of the coupler will be evenly split between the two waveguides.

5.3 Non-Orthogonal Coupled Mode Theory

The analysis presented in Section 5.2 describes how the coupled mode theory originally presented in Chapter 4 can be applied to the case of parallel waveguides. When we derived the coupled mode equations in Section 4.3, we made use of the orthogonality between waveguide modes, described by equation 4.18. However, for the case of parallel waveguides, the electromagnetic fields are written as a superposition of the isolated waveguide modes, as in equation 5.5. The mode of one waveguide is not orthogonal to the mode of the other waveguide located a distance d away. Strictly speaking, the coupled mode equations derived in Chapter 4 are based upon expanding the electromagnetic fields of the perturbed system in terms of a set of orthogonal bound modes of a *single* waveguide. Therefore, applying this formalism to the case of *two* waveguides whose modes are not orthogonal requires some qualification [56, 57].

The concept of orthogonality generally involves an integral of the product of two different modes over the cross sectional plane, as described in equation 4.18. For two orthogonal modes, the resultant integration yields 0. For the case of two parallel waveguide modes, if the integration of equation 4.18 were carried out, the result would not be zero:

$$\iint \left[\mathbf{e}_1 \times \mathbf{h}_2^* + \mathbf{e}_2^* \times \mathbf{h}_1 \right] dA \neq 0 \tag{5.18}$$

However, because the waveguide modes decay exponentially outside of the core region, the result of the overlap integral in equation 5.18 will be small compared with the power associated with either mode separately:

$$\iint \left[\mathbf{e}_1 \times \mathbf{h}_2^* + \mathbf{e}_2^* \times \mathbf{h}_1 \right] dA << P \tag{5.19}$$

Accordingly, it is often valid to treat the modes of the isolated waveguides as approximately orthogonal. Stated qualitatively, the integrated overlap of the waveguide modes is nonzero, but small, and it can therefore often be neglected. The non-orthogonality of the waveguide modes becomes more significant when the waveguides are close together.

Expanding the electromagnetic fields in terms of orthogonal modes is an attractive approach because in principle the modes of an optical waveguide comprise a complete basis set: any transverse electromagnetic field profile can be expressed in terms of a superposition of the orthogonal bound modes and radiation modes of the waveguide. However, for parallel waveguides it is more convenient to expand the electromagnetic fields in terms of the modes of the constituent waveguides. Haus et al. have developed an improved coupled mode treatment which does not rely on expanding the electromagnetic fields in terms of orthogonal modes [58, 59]. This treatment is known as non-orthogonal coupled mode theory (NCMT). The approach presented in [58] demonstrates that NCMT is inherently a variational theory, wherein the electromagnetic modes of a perturbed geometry are written as a superposition of several not-necessarily orthogonal trial functions. The variational approach begins with an integral expression for the propagation constant of a waveguide in terms of the electromagnetic fields. The expansion coefficients are treated as variational parameters of the trial solution, which are adjusted to extremize the propagation constant. This optimization yields the following coupled mode equations [58]:

$$\sum_n P_{mn}\frac{d}{dz}a_n = i\sum_n H_{mn}a_n \tag{5.20}$$

Where the matrix elements P_{mn} and H_{mn} are defined by:

$$P_{mn} \equiv \frac{1}{4}\iint\left[\mathbf{e_n}\times\mathbf{h_m^*} + \mathbf{e_m^*}\times\mathbf{h}_n\right]\cdot\mathbf{dA} \tag{5.21}$$

$$H_{mn} \equiv P_{mn}\beta_n + \frac{1}{4}k\sqrt{\frac{\varepsilon_0}{\mu_0}}\iint\delta\varepsilon_n\mathbf{e_n}\cdot\mathbf{e_m^*}dA \tag{5.22}$$

Equation 5.20 can be seen to be quite similar to the coupled mode equations given earlier in equation 5.6. In fact, if the modes are orthogonal, the matrix P is diagonal ($P_{mn} = 0$ for $m \neq n$), and the nonorthogonal coupled mode equations simplify to the result presented in equation 5.6. Thus, the conventional coupled mode equations are a special case of the more general result given in equation 5.20, applicable when the expansion modes are orthogonal. For the case of two identical parallel waveguides, the coupled mode equations 5.20 reduce to the form:

$$\begin{bmatrix}1 & x\\ x & 1\end{bmatrix}\frac{d}{dz}\begin{bmatrix}a_1\\ a_2\end{bmatrix} = i\left\{\begin{bmatrix}1 & x\\ x & 1\end{bmatrix}\begin{bmatrix}\beta & 0\\ 0 & \beta\end{bmatrix}+\begin{bmatrix}\Delta & \mu\\ \mu & \Delta\end{bmatrix}\right\}\begin{bmatrix}a_1\\ a_2\end{bmatrix} \tag{5.23}$$

Where the constants x, Δ, and μ are defined by:

$$x \equiv \frac{1}{4P}\iint\left[\mathbf{e}_1\times\mathbf{h}_2^* + \mathbf{e}_2^*\times\mathbf{h}_1\right]\cdot\mathbf{dA} \tag{5.24}$$

$$\Delta \equiv \frac{1}{4P}k\sqrt{\frac{\varepsilon_0}{\mu_0}}\left(n_{core}^2 - n_{clad}^2\right)\iint_{\text{guide 2}}\mathbf{e}_1\cdot\mathbf{e}_1^*dA \tag{5.25}$$

$$\mu \equiv \frac{1}{4P}k\sqrt{\frac{\varepsilon_0}{\mu_0}}\left(n_{core}^2 - n_{clad}^2\right)\iint_{\text{guide 1/2}}\mathbf{e}_1\cdot\mathbf{e}_2^*dA \tag{5.26}$$

The dimensionless quantity x is a measure of the non-orthogonality of the waveguide modes, the quantity Δ is the lowest order change in the waveguide propagation constant as a result of the presence of a second guide, and μ is the coupling constant as defined in Section 5.2. The quantity x is always less than one, and it typically decreases exponentially with waveguide separation. The non-orthogonal coupled mode equations of equation 5.23 can be cast into a form which is identical to that of the orthogonal coupled mode equations:

$$\frac{d}{dz}\begin{bmatrix}a_1\\ a_2\end{bmatrix} = \frac{i}{1-x^2}\begin{bmatrix}1 & -x\\ -x & 1\end{bmatrix}\begin{bmatrix}\beta+\Delta & \beta x+\mu\\ \beta x+\mu & \beta+\Delta\end{bmatrix}\begin{bmatrix}a_1\\ a_2\end{bmatrix} \tag{5.27}$$

$$\frac{d}{dz}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = i\begin{bmatrix} \beta' & \mu' \\ \mu' & \beta' \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{5.28}$$

The primed quantities in the above equation denote the effective propagation constant and coupling constant derived from non-orthogonal coupled mode theory, given by:

$$\beta' \equiv \beta + \frac{\Delta - \mu x}{1 - x^2} \tag{5.29}$$

$$\mu' \equiv \frac{\mu - x\Delta}{1 - x^2} \tag{5.30}$$

Clearly, if x is taken to be zero, corresponding to orthogonal modes, equations 5.29 and 5.30 reduce to the previously derived expressions for the propagation constant and coupling constant. Thus, non-orthogonal coupled mode theory provides a more complete description of the coupling between non-orthogonal waveguide modes. The form of the resulting non-orthogonal coupled mode equations is identical to that derived from orthogonal coupled mode theory. The non-orthogonality of the waveguides can be taken into account by using a slightly modified expression for the coupling constant, given in equation 5.30.

For weakly guiding waveguides, the three integrals describing μ, Δ and x can be written in terms of the scalar mode profile $\Phi(x,y)$ of the waveguides:

$$x \equiv \frac{\iint \Phi\left(x - \frac{d}{2}, y\right)\Phi\left(x + \frac{d}{2}, y\right)dA}{\iint \left|\Phi(x,y)\right|^2 dA} \tag{5.31}$$

$$\Delta \equiv \frac{k^2}{2\beta}\left(n_{core}^2 - n_{clad}^2\right)\frac{\iint_{\text{guide 2}} \Phi(x,y)\Phi(x,y)dA}{\iint \left|\Phi(x,y)\right|^2 dA} \tag{5.32}$$

$$\mu \equiv \frac{k^2}{2\beta}\left(n_{core}^2 - n_{clad}^2\right)\frac{\iint_{\text{guide 1/2}} \Phi\left(x - \frac{d}{2}, y\right)\Phi\left(x + \frac{d}{2}, y\right)dA}{\iint \left|\Phi(x,y)\right|^2 dA} \tag{5.33}$$

In order to account for the non-orthogonality of the waveguide modes, it is necessary to compute the three overlap integrals given in equations 5.31–5.33.

5.4 NonParallel Waveguides

The analysis in Section 5.2 describes how power transfers between two parallel waveguides. In a real waveguide device, the coupling section cannot be abruptly terminated after some length L. Instead the waveguides are gradually brought together, remain parallel for some length and then gradually separate as illustrated in Fig. 5.9. At either end of the coupler, the waveguides are sufficiently separated that the coupling between the two is negligible. As the guides are brought closer together, the coupling constant gradually increases, reaching its maximum value when the waveguides are closest. In order to properly design a coupler with the desired splitting ratio, it is necessary to account for the coupling in the tapered regions where the waveguides are not parallel [56, 60, 61].

When the waveguide separation changes slowly, it is valid to treat the coupling between waveguides by merely letting the coupling constant μ be a slowly varying function of z [62], in which case, the modified coupled mode equations are:

$$\frac{d}{dz}\begin{bmatrix} a_1(z) \\ a_2(z) \end{bmatrix} = \begin{bmatrix} i\beta & i\mu(z) \\ i\mu(z) & i\beta \end{bmatrix}\begin{bmatrix} a_1(z) \\ a_2(z) \end{bmatrix} \tag{5.34}$$

This equation can again be solved by eigenvector decomposition and integration. Because of the symmetry of the problem, the eigenvectors are still the symmetric and antisymmetric combination of isolated waveguide modes at any location z. The eigenvalues are slowly varying functions of z. The solution to the z-dependent coupled mode equation is:

$$\begin{bmatrix} a_1(z) \\ a_2(z) \end{bmatrix} = e^{+i\beta z}\begin{bmatrix} \cos\big(\phi(z)\big) & i\sin\big(\phi(z)\big) \\ i\sin\big(\phi(z)\big) & \cos\big(\phi(z)\big) \end{bmatrix}\begin{bmatrix} a_1(0) \\ a_2(0) \end{bmatrix} \tag{5.35}$$

Where the quantity $\phi(z)$ is defined by:

$$\phi(z) \equiv \int_0^z \mu(z')dz' \tag{5.36}$$

This approach is only valid when the separation between the waveguides changes slowly. This condition is often called the adiabatic condition [62].

The non-orthogonal coupled mode analysis described in Section 5.3 can also be extended to account for slowly tapered structures, as described in [63]. The non-orthogonal coupled mode equations for a slowly varying structure are:

$$\mathbf{P}\frac{d\mathbf{a}}{dz} + \frac{1}{2}\frac{d\mathbf{P}}{dz}\mathbf{a} = i\mathbf{H}\mathbf{a} \tag{5.37}$$

Where the matrices **P** and **H** are defined in equations 5.21 and 5.22, except that they are now treated as z-dependent quantities. For the case of two identical waveguides, equation 5.37 can again be solved by diagonalizing the coupled mode equations and integrating. For structures such as the one illustrated in Fig. 5.9, the waveguides are symmetric with respect to reflection about the z = 0 plane. Therefore the d**P**/dz terms in equation 5.37 are antisymmetric about the z = 0 plane, so that when integrated over the

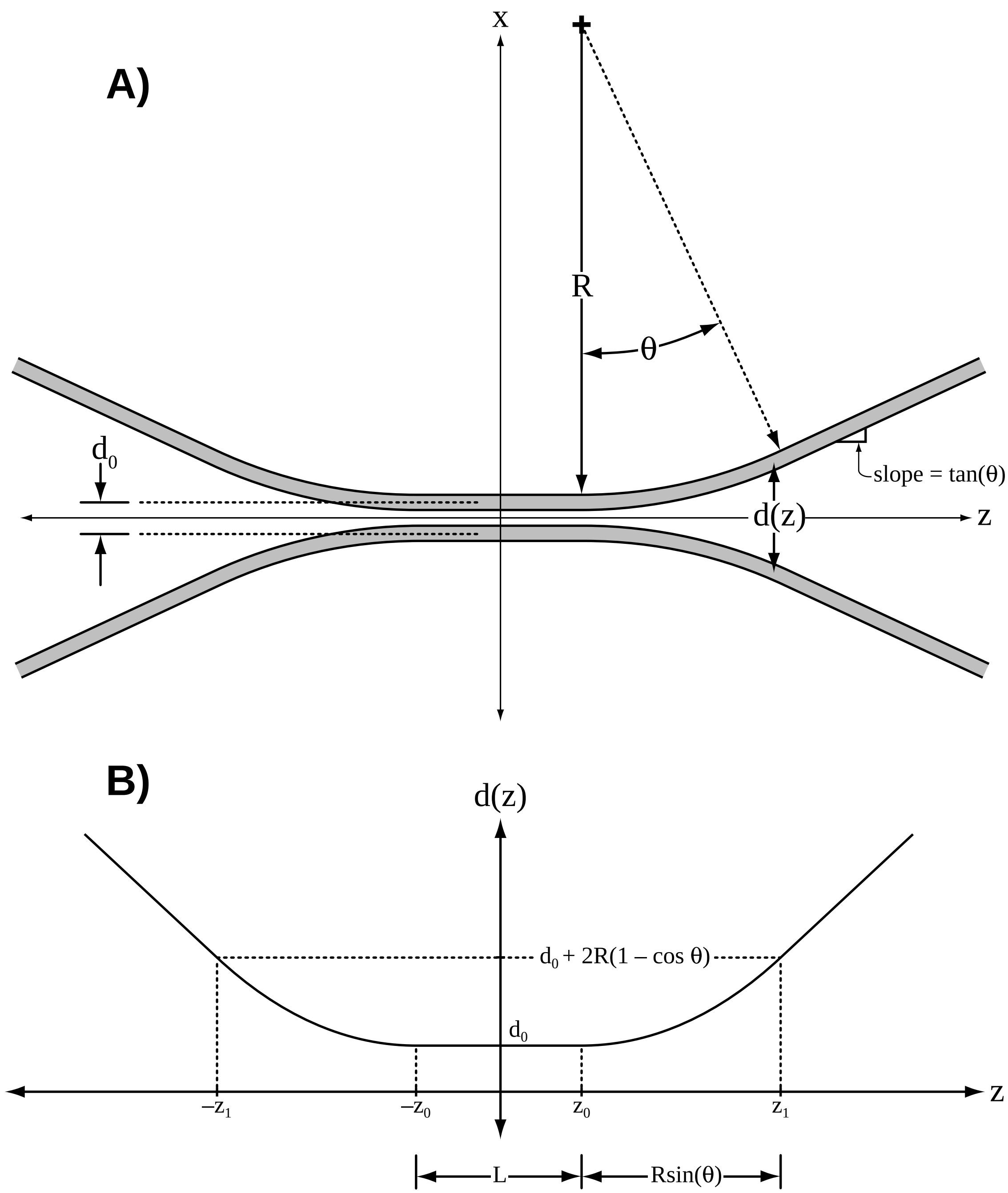


Figure 5.9 In a practical directional coupler, the waveguides are not parallel over the entire coupling region. Depicted in (A) is a configuration where the guides gradually come together and apart in a curved path. The lower portion of the figure (B) plots the waveguide separation d(z) for the coupler geometry depicted in (A).

entire length of the coupler, the contributions from these terms evaluate to zero. The solution of the non-orthogonal coupled mode equations for a tapered coupler is identical to that given in equations 5.35 and 5.36, with $\mu(z)$ replaced by the more accurate expression given in equation 5.30.

Figure 5.9A illustrates one common waveguide coupler configuration. The two waveguides approach one another each following a circular arc path. The curved waveguide section is characterized by an angle θ and a radius of curvature R. Having reached their nearest separation d_0, the two waveguides remain parallel for a length L, and then separate again following a curved path. The waveguide separation d(z) is described by:

$$d(z) = \begin{cases} d_0 & |z| < \dfrac{L}{2} \\ d_0 + 2R - 2\sqrt{R^2 - \left(|z| - \dfrac{L}{2}\right)^2} & \dfrac{L}{2} < |z| < \dfrac{L}{2} + R\sin(\theta) \\ d_0 + 2R\big(1 - \cos(\theta)\big) + 2\left(|z| - \dfrac{L}{2} - R\sin(\theta)\right)\tan(\theta) & \dfrac{L}{2} + R\sin(\theta) < |z| \end{cases} \tag{5.38}$$

The function d(z) is plotted in Fig. 5.9B. In principle, the waveguides need not separate to infinity away from the coupler; once the waveguides have attained a sufficient separation the waveguide coupling becomes negligible and therefore the exact functional form of d(z) outside of the primary coupling region is not important.

In Chapter 4, we demonstrated that the for the case of a slab waveguide, the electromagnetic fields decay exponentially in the region outside of the core (cf. equation 4.11), and we further argued that for most waveguides of interest, the mode amplitude falls exponentially outside of the core region. Because of the exponential decay of the electromagnetic mode away from the core, the coupling coefficient μ also falls exponentially as the waveguide separation d is increased*. Section 5.5 will discuss different ways in which the coupling coefficients can numerically evaluated for the case of a channel waveguide, but it is usually a good approximation to treat μ as an exponentially decaying function of d(z), the slowly varying waveguide separation:

$$\mu(z) = A\exp\left(-\frac{d(z)}{\bar{d}}\right) \tag{5.39}$$

Where the parameters A and $\bar{d}$ are empirically determined parameters which can be found by numerically fitting an exponential relationship to calculated values of μ for selected waveguide separations. Under the assumption that the coupling can be accurately described by equation 5.39, the total integrated coupling for the structure illustrated in Fig. 5.9 can be calculated from equations 5.36, 5.38 and 5.39. As illustrated in Fig. 5.10, the coupling can be separated into three contributions: $\Delta\phi_0$, corre-

* It is for this reason that the codirectional coupling between parallel waveguides is sometimes referred to as "evanescent coupling".

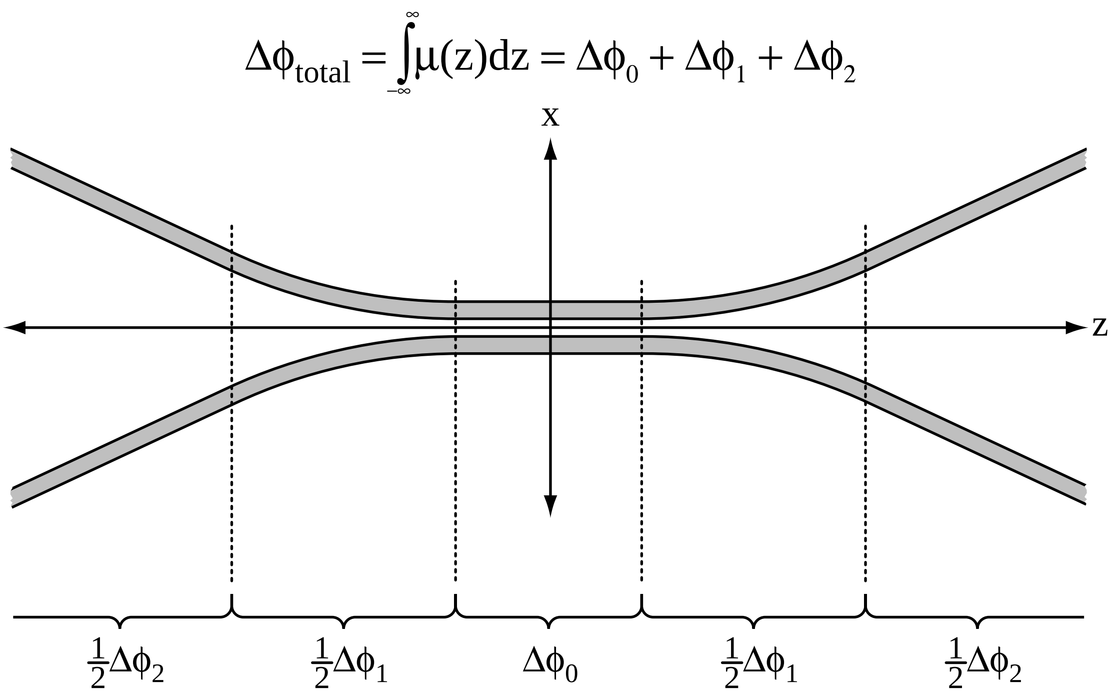


Figure 5.10 For the coupler depicted in Fig. 5.9, the total integrated coupling can be divided into three portions, associated with the parallel, curved, and sloped regions of the coupler.

sponding to the coupling in the parallel region, $\Delta\phi_1$, the coupling in the curved regions, and $\Delta\phi_2$, the coupling associated with the linearly sloped regions.

$$\Delta\phi_{total} = \int_{-\infty}^{\infty}\mu(z)dz = \Delta\phi_0 + \Delta\phi_1 + \Delta\phi_2 \tag{5.40}$$

In the region $-z_0 < z < z_0$, the waveguides are parallel and the integrated coupling $\Delta\phi_0$ is:

$$\Delta\phi_0 = \mu(d_0)L = AL\exp\left(-\frac{d_0}{\bar{d}}\right) \tag{5.41}$$

In the region $z_0 < |z| < z_1$, the waveguide separation d(z) can be approximated as a quadratic function of z, provided the arc angle θ is small:

$$d(z) = d_0 + 2R - 2R\sqrt{1-\left(\frac{z-z_0}{R}\right)^2} \approx d_0 + \frac{(z-z_0)^2}{R} \tag{5.42}$$

Using the above quadratic expression for d(z), the contribution to the total integrated coupling in the curved regions is:

$$
\begin{aligned}
\Delta\phi_1 &= 2\int_{z_0}^{z_1}\mu(d(z))dz = 2A\exp\left(-\frac{d_0}{\bar{d}}\right)\int_0^{z_1-z_0}\exp\left(-\frac{z^2}{\bar{d}R}\right)dz \\
\Delta\phi_1 &= A\exp\left(-\frac{d_0}{\bar{d}}\right)\sqrt{\pi\bar{d}R}\,\mathrm{erf}\left(\sqrt{\frac{R}{\bar{d}}}\sin\theta\right)
\end{aligned}
\tag{5.43}
$$

Finally, within the linearly sloped region ($z_1 < |z|$), the remaining contribution to the waveguide coupling is:

$$
\begin{aligned}
\Delta\phi_2 &= 2\int_{z_1}^{\infty}\mu(d(z))dz = 2A\exp\left(-\frac{d_0}{\bar{d}}\right)\exp\left(-\frac{2R}{\bar{d}}(1-\cos\theta)\right)\int_0^{\infty}\exp\left(-\frac{2z}{\bar{d}}\tan\theta\right)dz \\
\Delta\phi_2 &= A\exp\left(-\frac{d_0}{\bar{d}}\right)\frac{\bar{d}}{\tan\theta}\exp\left(-\frac{2R}{\bar{d}}(1-\cos\theta)\right)
\end{aligned}
\tag{5.44}
$$

Collecting the results from equations 5.42, 5.43 and 5.44, the total waveguide coupling for the structure illustrated in Fig. 5.9 is:

$$
\Delta\phi_{total} = \mu(d_0)\left[L + \sqrt{\pi\bar{d}R}\,\mathrm{erf}\left(\sqrt{\frac{R}{\bar{d}}}\sin\theta\right) + \frac{\bar{d}}{\tan\theta}\exp\left(-\frac{2R}{\bar{d}}(1-\cos\theta)\right)\right]
\tag{5.45}
$$

For many cases of interest, the waveguide parameters are such that all but the first order correction in the above equation can be neglected. Thus, when

$$
\frac{R}{d}\sin^2(\theta) >> 1\,,
\tag{5.46}
$$

the error function in equation 5.45 evaluates to unity, and the contribution from the region ($z_1 < |z|$) can be neglected entirely, so that the total coupling becomes:

$$
\Delta\phi_{total} \approx \mu(d_0)\left[L + \sqrt{\pi\bar{d}R}\right]
\tag{5.47}
$$

This remarkably simple result indicates that under certain conditions, the coupling between two non-parallel waveguides can be treated as if the coupler were comprised of two parallel waveguides separated by d_0, with an effective coupling length of L_{eff} given by:

$$
L_{eff} = L + \sqrt{\pi\bar{d}R}
\tag{5.48}
$$

A more accurate calculation of the total waveguide coupling can be computed by numerically performing the integration in equation 5.36 using the exact expression for d(z). However, such numerical integration is time consuming, and the simple result given in equation 5.47 can usually be used as a quick, intuitive, and often accurate estimate of the integrated coupling.

5.5 Calculation of Coupling Constants

Designing a 50% directional coupler requires that the coupling coefficient μ be known accurately for a range of waveguide separations. We have described two methods for determining the coupling coefficient μ for a specified waveguide separation. The most direct method involves rigorously solving for the symmetric and antisymmetric modes of the parallel waveguide system. As discussed in Section 5.2, when the inter-waveguide spacing is large, this method becomes numerically problematic because of the small difference between the symmetric and antisymmetric propagation constants. However, for closely spaced waveguides this approach is the most accurate way to model the behavior of the system.

The second way to evaluate the coupling coefficients is to first solve for the mode of an isolated waveguide and then treat the presence of the second parallel waveguide as a perturbation. The coupling constant μ can then be written in terms of a simple overlap integral involving the isolated waveguide mode, as described in equation 5.11. This approach can be extended with non-orthogonal coupled mode theory, as described in Section 5.3, which yields a more accurate estimate of the coupling constant, especially when the waveguides are close together. One advantage of using coupled mode theory is that after the waveguide mode has been calculated once, it is possible to calculate the coupling constant for many waveguide separations. Calculating the coupling constant $\mu(d)$ using the coupled mode theory as described in Sections 5.2 and 5.3 requires that the electromagnetic mode of one waveguide be known accurately at points within the other waveguide (which is located a distance d away.) Therefore, the accuracy of the coupling coefficient is limited by one's ability to compute the electromagnetic mode at points outside of the waveguide core. For the case of parallel slab waveguides, the exact functional form of the mode profile is known at all points outside of the waveguide core, and therefore the integral given in equation 5.11 can be calculated analytically for any waveguide separation. However, for more complicated geometries such as the square channel waveguide considered in this work (Fig. 5.11), the electromagnetic modes cannot be determined analytically. Numerical techniques such as the finite difference method or the finite elements method are typically employed to solve for the waveguide mode at a series of discrete points. The overlap integral in equation 5.11 can then be numerically approximated with a summation to obtain a value for μ. The requirement that the field be known accurately at points far outside the waveguide core places a high demand on the numerical modesolver. Recall that the mode profile typically decays exponentially outside of the waveguide core. Therefore, even if the waveguide mode is known to a high degree of accuracy inside the core, the relative accuracy at points far outside the core where the mode is much smaller might be considerably less. It is also important to point out that for very large waveguide separations the coupling coefficient becomes so small that the coupling between waveguides can be completely neglected.

For the channel waveguide depicted in Fig. 5.11, the waveguide mode at points far outside of the core can be approximated by replacing the core with an equivalent circular core region with the same cross sectional area. The advantage of this approach is that the circular waveguide is a very well studied structure which has analytical solutions for the waveguide modes [15]. The electromagnetic field in the region outside of the core is a Bessel function of the second kind. A more accurate functional expression for the electromagnetic field can be found by including higher order Bessel functions, and matching electromagnetic boundary conditions at discrete points along the core-clad interface [64]. The advantage of using the Bessel function approximation is that even though the modes are only approximate, they are expected to have the correct asymptotic form at points far outside of the waveguide core. Therefore, the approximate circular waveguide modes can be used to calculate the coupling coefficient μ for very large waveguide separations where a numerical modesolver cannot reliably predict the electromagnetic mode.

Figure 5.12 is a plot of the coupling coefficient $\mu(d)$ as a function of the waveguide separation d, for the channel waveguide depicted in Fig. 5.11. The coupling coefficients presented in Fig. 5.12 have been calculated using the methods described above: exact calculation of the symmetric and antisymmetric modes and non-orthogonal coupled mode theory using a numerical modesolver. In calculating the symmetric and antisymmetric modes, and the isolated waveguide modes, a finite difference modesolver program was used [27].

Careful examination of the data presented in Fig. 5.12 from the direct calculation of the symmetric and antisymmetric modes (symbol) reveals that there is some uncertainty in the calculated coupling constant μ at waveguide separations greater than about 18 microns. This uncertainty is evidenced

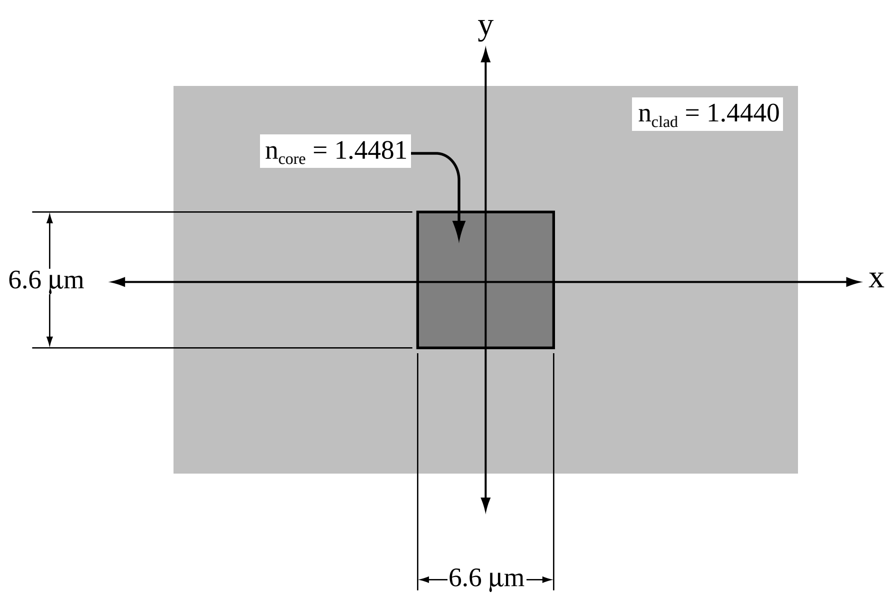


Figure 5.11 Channel waveguide. The square core is embedded within a cladding region with lower index of refraction.

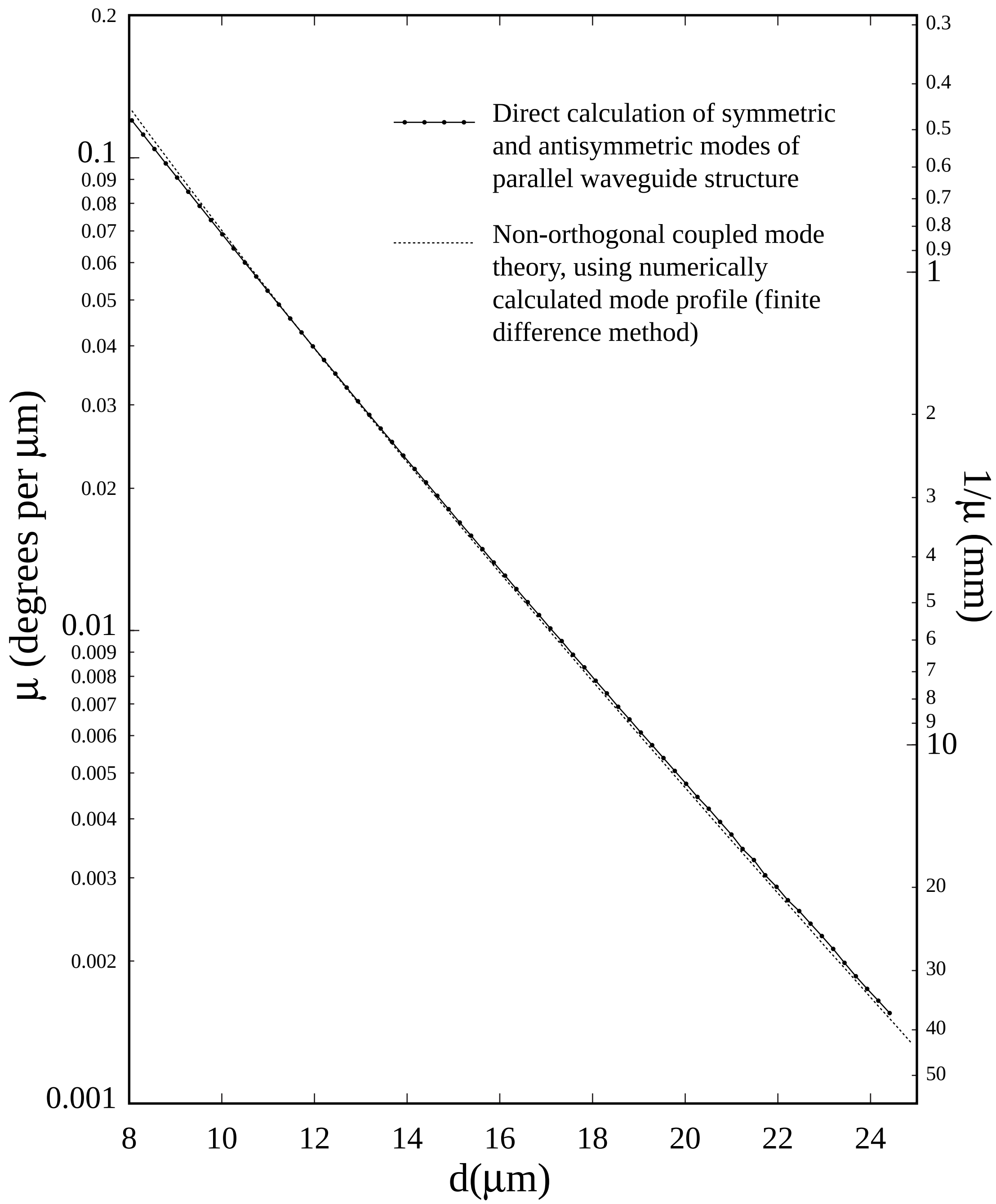


Figure 5.12 Calculated coupling constants of two parallel channel waveguides as a function of waveguide separation. The left scale expresses the coupling constant in units of degrees per micron, and the right scale measures the inverse coupling constant (often called the coupling length). These data have been calculated using two different methods: direct evaluation of the symmetric and antisymmetric propagation constants, and by non-orthogonal coupled mode theory, using a numerically determined mode profile.

by the deviation in the μ vs. d data from a smooth curve. As described in Section 5.2, this uncertainty results from the inability of the numerical modesolver to compute the symmetric and antisymmetric propagation constants to sufficient accuracy at large waveguide separations. However, the coupling constants calculated using non-orthogonal coupled mode techniques (dashed line) are in excellent agreement with the results obtained from direct calculation of the symmetric and antisymmetric modes. Because the coupled mode techniques do not involve computing the difference between two almost equal values, the resulting coupling constants are accurate to a greater waveguide separation.

Figure 5.13 presents the same data as in Fig. 5.12, plotted for a larger range of waveguide separations. Beyond a waveguide separation of about 40 microns, the coupling constant calculated using coupled mode theory with numerically calculated waveguide modes seems to approach a constant value. This erroneous and unphysical result arises because the modes calculated by the finite difference modesolver have not been computed to sufficient accuracy at large distances from the waveguide core. A careful examination of the numerically determined mode profile reveals that when the computation window for the modesolver is extended beyond about 40 microns, the calculated mode tends to level off at a constant (although very small) value rather than continue to decay exponentially. This levelling of the calculated mode can only be noticed when the mode amplitude is plotted on a logarithmic scale, and for many applications it is unimportant because the constant value which it approaches is several orders of magnitude smaller than the peak value of the mode profile. The unphysical levelling of the mode profile is an artifact of the algorithm used in the numerical modesolver. The range of validity of the modesolver could be extended by configuring the software to solve for the modes to a higher degree of precision, but only at the cost of increased computation time. It is computationally difficult to find a numerical solution for the mode profile which is accurate over several orders of magnitude, and therefore the modesolver cannot be expected to correctly predict the exponential decay of the mode profile at points far away from the waveguide core. The levelling-off of the numerically calculated mode profile causes the calculated values of μ to approach a constant value for waveguide separations larger than about 40 microns.

If the channel waveguide mode is approximated as that of an equivalent circular waveguide, the mode amplitude can be written in terms of first order Bessel functions. The Bessel function correctly predicts the exponential decay of the waveguide mode, even for points far outside the waveguide core where the mode amplitude is suppressed by several orders of magnitude. The coupling constants calculated using this approximation correctly exhibit the exponential decay of μ at large waveguide separations, as shown in Fig. 5.13. Moreover, the results from the circular harmonic approximation are in excellent agreement with the results independently calculated using other methods, even for relatively small waveguide separations.

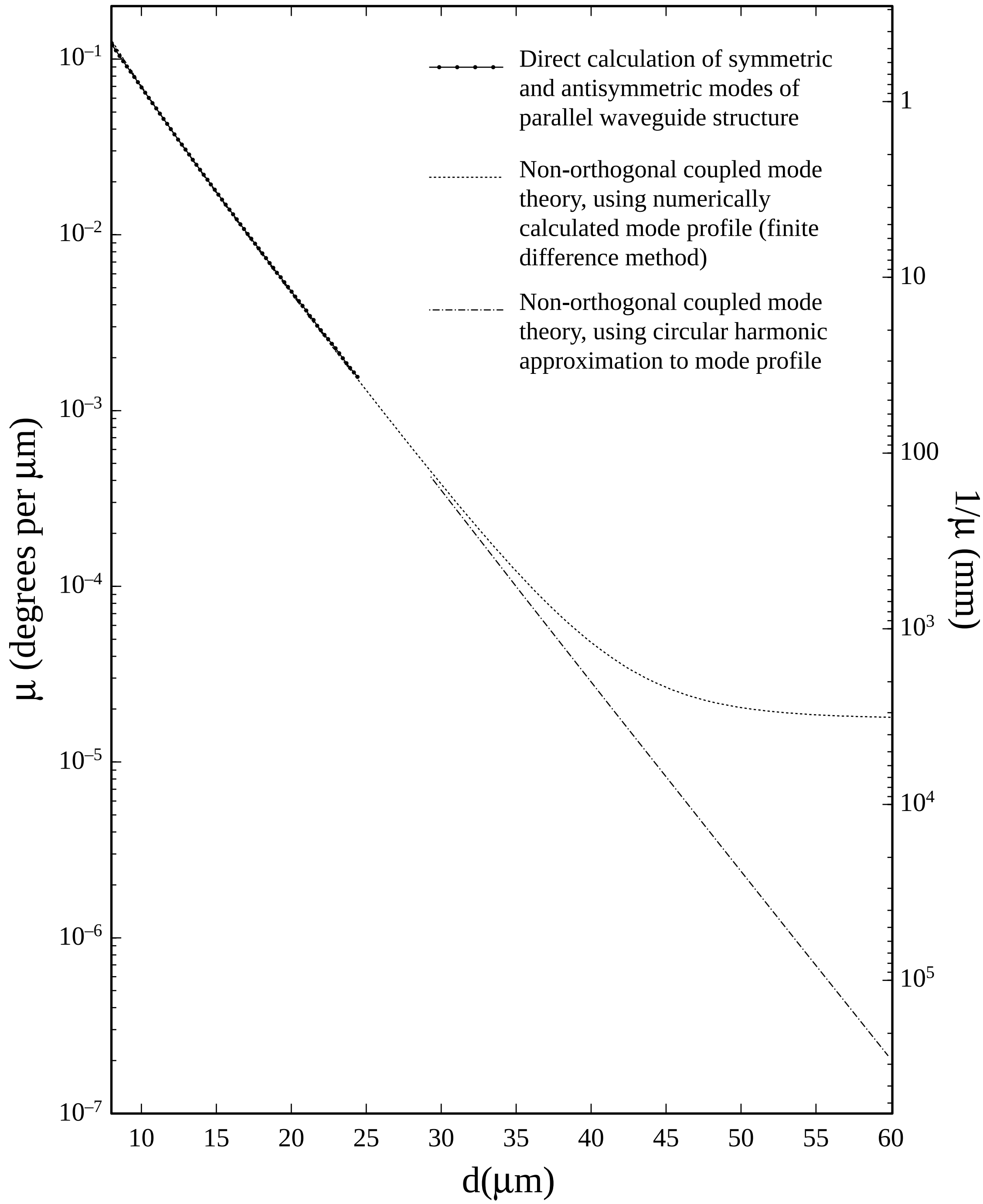


Figure 5.13 Calculated coupling constants of two parallel channel waveguides as a function of waveguide separation. As in Fig. 5.12, the left scale expresses the coupling constant in units of degrees per micron, and the right scale measures the inverse coupling constant (often called the coupling length). Three different methods were used to calculate the data presented in this figure: direct calculation of symmetric and antisymmetric modes, non-orthogonal coupled mode theory (NCMT) with a numerically calculated mode profile, and NCMT using a circular harmonic approximation for the mode profile.

5.6 Michelson Interferometer

In treating the coupling of modes (either codirectional coupling as presented earlier in this chapter, or contradirectional coupling as presented in Chapters 3 and 4), we have always expressed the solutions to the coupled mode equations in matrix form (equations 5.15 and 3.14.) The matrix which describes the evolution of the mode amplitudes is often called a transfer matrix. When the coupled mode solutions are expressed in matrix form, it is relatively easy to treat more complicated sequences structures by simply multiplying the appropriate transfer matrices for each portion of the structure [48]. For example, the transfer matrix of two concatenated codirectional couplers of lengths L_1 and L_2 can be computed by multiplying the transfer matrices for each coupler:

$$\begin{gathered}\mathbf{T}_{\text{total}} = \mathbf{T}_2\mathbf{T}_1 = e^{i\beta L_2}\begin{bmatrix} \cos(\mu L_2) & i\sin(\mu L_2) \\ i\sin(\mu L_2) & \cos(\mu L_2) \end{bmatrix} e^{i\beta L_1}\begin{bmatrix} \cos(\mu L_1) & i\sin(\mu L_1) \\ i\sin(\mu L_1) & \cos(\mu L_1) \end{bmatrix} \\ \mathbf{T}_{\text{total}} = e^{i\beta(L_1+L_2)}\begin{bmatrix} \cos\big(\mu(L_1+L_2)\big) & i\sin\big(\mu(L_1+L_2)\big) \\ i\sin\big(\mu(L_1+L_2)\big) & \cos\big(\mu(L_1+L_2)\big) \end{bmatrix}\end{gathered} \tag{5.49}$$

The transfer matrix approach can be applied to the case of the Michelson interferometer, as illustrated in Fig. 5.14. The interferometer can be treated as a concatenation of a codirectional coupler with a pair of identical Bragg gratings. Because there are two waveguides in the device, each of which can support a forward and backward mode, four mode amplitudes are needed to describe the system. The mode amplitudes can be collected as a vector of four quantities:

$$\mathbf{a}(z) = \begin{bmatrix} a_{1+}(z) \\ a_{2+}(z) \\ a_{1-}(z) \\ a_{1-}(z) \end{bmatrix} \tag{5.50}$$

where those quantities labelled + denote forward travelling modes, those quantities labelled – denote backward travelling modes, and the integer subscript (1 or 2) is used to number the two waveguides. Figure 5.14 illustrates the form of the transfer matrices for each section of the Michelson interferometer. Notice that the transfer matrix $\mathbf{T_1}$ associated with the coupler contains only terms which couple the modes travelling in the same direction, (i.e. the coupler does not produce any coupling between oppositely travelling modes.) The transfer matrix $\mathbf{T_2}$ associated with the Bragg gratings contains no terms which couple modes in waveguide 1 with those in waveguide 2, because throughout the grating region the waveguides are sufficiently separated to be completely decoupled.

The 4 x 4 transfer matrix $\mathbf{T_1}$ representing the directional coupler is defined by:

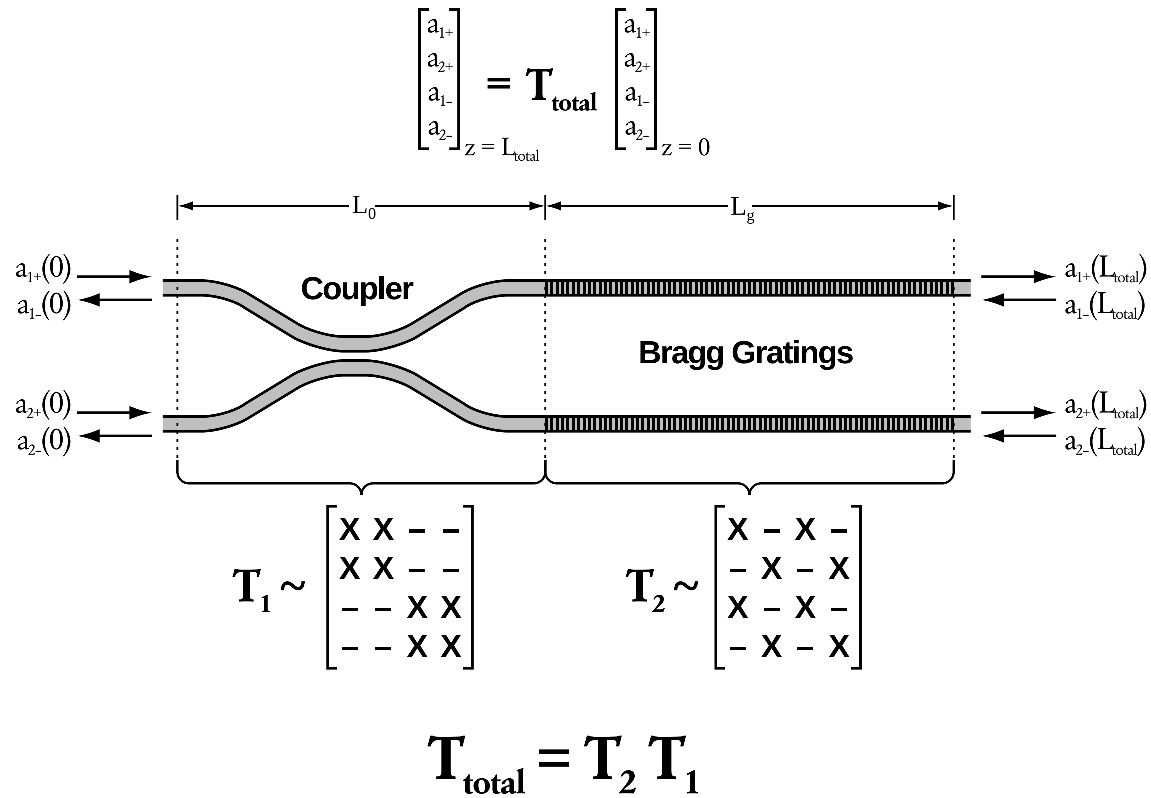


Figure 5.14 The integrated Michelson interferometer can be treated as a concatenation of a directional coupler with a pair of grating segments. The composite device can be analyzed by simply multiplying the transfer matrices for each of the concatenated segments. The form of the transfer matrices is illustrated in this figure.

$$\mathbf{a}(L_0) = \mathbf{T_1}\,\mathbf{a}(0) = \begin{bmatrix} e^{+i\beta L_0}\mathbf{I_2} & 0 \\ 0 & e^{-i\beta L_0}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} \mathbf{R} & 0 \\ 0 & \mathbf{R}^* \end{bmatrix} \mathbf{a}(0) \tag{5.51}$$

$$\mathbf{R} \equiv \begin{bmatrix} \cos(\mu L_{eff}) & i\sin(\mu L_{eff}) \\ i\sin(\mu L_{eff}) & \cos(\mu L_{eff}) \end{bmatrix} \tag{5.52}$$

L_{eff} is the effective coupling length of the directional coupler, L_0 is the total propagation length of the coupler, $\mathbf{I_2}$ is the 2 x 2 identity matrix, and $\mathbf{R}$ is the 2 x 2 transfer matrix associated with the coupling of forward travelling modes in a single waveguide. In equation 5.51, the full 4 x 4 transfer matrix for the coupler is expressed in terms of 2 x 2 sub-matrices. Notice that the forward travelling modes do not interact with the backward travelling modes, and the 2 x 2 transfer matrix describing the coupling of the backward travelling modes is simply the complex conjugate of the matrix describing the coupling of forward travelling modes.

The Bragg grating section can be likewise be represented by a transfer matrix. For the two identical Bragg gratings, the forward and backward modes of each waveguide are coupled but because the guides are far separated by design, there is no interaction between waveguide 1 and waveguide 2. The 4 x 4 transfer matrix for the two identical Bragg gratings follows directly from the 2 x 2 result given in equation 3.14:

$$\mathbf{a}(L_0 + L_g) = \begin{bmatrix} e^{\frac{i}{2}k_g L_g}\mathbf{I_2} & 0 \\ 0 & e^{-\frac{i}{2}k_g L_g}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} A & 0 & B & 0 \\ 0 & A & 0 & B \\ B^* & 0 & A^* & 0 \\ 0 & B^* & 0 & A^* \end{bmatrix} \mathbf{a}(L_0) \tag{5.53}$$

where the quantities A and B are defined by:

$$\begin{aligned} A &\equiv \cosh(\gamma L_g) + i\frac{\delta}{\gamma}\sinh(\gamma L_g) \\ B &\equiv \frac{\kappa}{\gamma}\sinh(\gamma L_g) \end{aligned} \tag{5.54}$$

In the above equation L_g is the length of the grating section, as depicted in Fig. 5.14, k_g is the k-vector of the grating, κ is the grating strength, δ is the deviation from the Bragg condition as described in equation 3.7, and γ is defined in equation 3.12. It is convenient to express the transfer matrix for the identical Bragg gratings in terms of 2 x 2 sub-matrices:

$$\mathbf{T_2} = \begin{bmatrix} e^{\frac{i}{2}k_g L_g}\mathbf{I_2} & 0 \\ 0 & e^{-\frac{i}{2}k_g L_g}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} A\mathbf{I_2} & B\mathbf{I_2} \\ B^*\mathbf{I_2} & A^*\mathbf{I_2} \end{bmatrix} \tag{5.55}$$

The 4 x 4 transfer matrix of the Michelson interferometer can be found by simply multiplying the transfer matrices of the coupler and the grating given in equations 5.51 and 5.55, as depicted in Fig. 5.14. The resulting transfer matrix of the complete structure is:

$$\mathbf{T_{total}} = \mathbf{T_2}\mathbf{T_1} = \begin{bmatrix} e^{\frac{i}{2}k_g L_g}\mathbf{I_2} & 0 \\ 0 & e^{-\frac{i}{2}k_g L_g}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} A\mathbf{I_2} & B\mathbf{I_2} \\ B^*\mathbf{I_2} & A^*\mathbf{I_2} \end{bmatrix} \begin{bmatrix} e^{+i\beta L_0}\mathbf{I_2} & 0 \\ 0 & e^{-i\beta L_0}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} \mathbf{R} & 0 \\ 0 & \mathbf{R} \end{bmatrix} \tag{5.56}$$

where each of the matrix elements given in equation 5.56 is understood to be a 2 x 2 sub-matrix. Equation 5.56 can be simplified to:

$$\mathbf{T_{total}} = \begin{bmatrix} e^{\frac{i}{2}k_g L_g}\mathbf{I_2} & 0 \\ 0 & e^{-\frac{i}{2}k_g L_g}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} A\mathbf{R} & B\mathbf{R}^* \\ B^*\mathbf{R} & A^*\mathbf{R}^* \end{bmatrix} \begin{bmatrix} e^{+i\beta L_0}\mathbf{I_2} & 0 \\ 0 & e^{-i\beta L_0}\mathbf{I_2} \end{bmatrix} \tag{5.57}$$

The transfer matrix for the Michelson interferometer relates the mode amplitudes at opposite ends of the device in the following way:

$$\mathbf{a}(L_{total}) = \begin{bmatrix} a_{1+} \\ a_{2+} \\ a_{1-} \\ a_{2-} \end{bmatrix}_{z=L_{total}} = \mathbf{T_{total}} \begin{bmatrix} a_{1+} \\ a_{2+} \\ a_{1-} \\ a_{2-} \end{bmatrix}_{z=0} \tag{5.58}$$

If all of the forward and backward mode amplitudes are known at the left end of the device (taken as z = 0), then multiplying the mode amplitude vector by the transfer matrix $\mathbf{T_{total}}$ will yield the mode amplitudes at the right end of the Michelson interferometer. As described in section 3.2, the boundary conditions are usually not completely specified at z = 0. Rather than specifying both the input and output signals at z = 0 and deriving the input and outputs at $z = L_{total}$, we want to specify the inputs at z = 0 and the inputs at $z = L_{total}$ and derive the outputs at z = 0 and $z = L_{total}$. Substituting these boundary conditions into equation 5.58 gives:

$$\begin{bmatrix} \mathbf{y_1} \\ \mathbf{x_1} \end{bmatrix} = \mathbf{T_{total}} \begin{bmatrix} \mathbf{x_0} \\ \mathbf{y_0} \end{bmatrix} \tag{5.59}$$

where $\mathbf{x_0}$, $\mathbf{y_0}$, $\mathbf{x_1}$, and $\mathbf{y_1}$ are 2-vectors representing inputs and outputs to the system. The **x**'s represent signals entering the device (inputs) whereas the **y**'s represent signals emerging from the device (outputs.) $\mathbf{x_0}$ and $\mathbf{y_0}$ represent the input and output vectors respectively at the left edge of the device:

$$\begin{aligned} \mathbf{x_0} &\equiv \begin{bmatrix} a_{1+} \\ a_{2+} \end{bmatrix}_{z=0} && \text{(input signals launched into waveguides from the left)} \\ \mathbf{y_0} &\equiv \begin{bmatrix} a_{1-} \\ a_{2-} \end{bmatrix}_{z=0} && \text{(emerging signals on left side of device)} \end{aligned} \tag{5.60}$$

Likewise, $\mathbf{x_1}$ and $\mathbf{y_1}$ represent respectively the input and output vectors on the right side of the device (at $z = L_{total}$):

$$\begin{aligned} \mathbf{x}_1 &\equiv \begin{bmatrix} a_{1-} \\ a_{2-} \end{bmatrix}_{z=L_{total}} && \text{(input signals launched into waveguides from the right)} \\ \mathbf{y}_1 &\equiv \begin{bmatrix} a_{1+} \\ a_{2+} \end{bmatrix}_{z=L_{total}} && \text{(emerging signals on right side of device)} \end{aligned} \tag{5.61}$$

Of particular interest is the case where $\mathbf{x_1} = 0$, meaning no light is launched into the waveguides from the right-hand side. Using the derived expression for the transfer matrix (equation 5.57) it is possible to solve for the output signal $\mathbf{y_0}$ in terms of the input signal $\mathbf{x_0}$:

$$\begin{bmatrix} \mathbf{y_1} \\ 0 \end{bmatrix} = \begin{bmatrix} e^{\frac{i}{2}k_g L_g}\mathbf{I_2} & 0 \\ 0 & e^{-\frac{i}{2}k_g L_g}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} A\mathbf{R} & B\mathbf{R}^* \\ B^*\mathbf{R} & A^*\mathbf{R}^* \end{bmatrix} \begin{bmatrix} e^{+i\beta L_0}\mathbf{I_2} & 0 \\ 0 & e^{-i\beta L_0}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} \mathbf{x_0} \\ \mathbf{y_0} \end{bmatrix} \tag{5.62}$$

which can be simplified to:

$$\begin{bmatrix} \mathbf{y_1} e^{-\frac{i}{2}k_g L_g} \\ 0 \end{bmatrix} = \begin{bmatrix} A\mathbf{R} & B\mathbf{R}^* \\ B^*\mathbf{R} & A^*\mathbf{R}^* \end{bmatrix} \begin{bmatrix} \mathbf{x_0} e^{+i\beta L_0} \\ \mathbf{y_0} e^{-i\beta L_0} \end{bmatrix} \tag{5.63}$$

The lower set of equations can then be used to arrive at a relationship between $\mathbf{x_0}$ and $\mathbf{y_0}$:

$$0 = B^*\mathbf{R}\mathbf{x_0} e^{+i\beta L_0} + A^*\mathbf{R}^*\mathbf{y_0} e^{-i\beta L_0} \tag{5.64}$$

$$\mathbf{y_0} = -e^{+i2\beta L_0} \left(\frac{B}{A}\right)^* \left(\mathbf{R}^*\right)^{-1} \mathbf{R}\,\mathbf{x_0} \tag{5.65}$$

$\mathbf{R}$ is the 2 x 2 matrix associated with the coupler, and the constants A and B are defined in equation 5.54. Substituting equations 5.52 and 5.54 into equation 5.65 gives:

$$\begin{bmatrix} a_{1-} \\ a_{2-} \end{bmatrix}_{z=0} = -\frac{\frac{\kappa^*}{\gamma}\sinh(\gamma L_g)}{\cosh(\gamma L_g) - i\frac{\delta}{\gamma}\sinh(\gamma L_g)} e^{+i2\beta L_0} \begin{bmatrix} \cos(2\mu L_{eff}) & i\sin(2\mu L_{eff}) \\ i\sin(2\mu L_{eff}) & \cos(2\mu L_{eff}) \end{bmatrix} \begin{bmatrix} a_{1+} \\ a_{2+} \end{bmatrix}_{z=0} \tag{5.66}$$

The first term in equation 5.66 is no more than the reflection spectral response of a Bragg grating of length L_g. This term is identical to the result derived for an isolated Bragg grating in equation 3.16. The remainder of equation 5.66 is equivalent to the transfer matrix for a directional coupler with an effective coupling length of $2L_{eff}$. Signals which enter the device from the left pass once through the directional coupler and are then reflected by the Bragg gratings. The reflected signals once again pass through the coupler. Thus, the result of equation 5.66 could almost be predicted without resorting to the mathematics presented earlier: the signal passes through the directional coupler twice (which accounts for the factor of 2 in the matrix terms of 5.66), and the filtering action of the Bragg grating is contained entirely in the first term. Any light which is not reflected by the Bragg grating simply emerges on the right side of the device.

As described in section 5.1, when the coupler is designed to split the light evenly between the upper and lower waveguides, the Michelson interferometer can be used to redirect the filtered signal to a different waveguide. Equation 5.66 describes this relationship mathematically. When the coupler is designed so that $\mu L_{eff} = \pi/4$, a single pass through the coupler will divide an incident signal 50% between the upper and lower waveguides, and two passes through the coupler will transfer the signal completely from one guide to the other. Substituting $\mu L_{eff} = \pi/4$ into equation 5.66 gives:

$$\begin{bmatrix} a_{1-} \\ a_{2-} \end{bmatrix}_{z=0} = -\frac{\frac{\kappa^*}{\gamma}\sinh(\gamma L_g)}{\cosh(\gamma L_g) - i\frac{\delta}{\gamma}\sinh(\gamma L_g)} e^{+i2\beta L_0} \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \begin{bmatrix} a_{1+} \\ a_{2+} \end{bmatrix}_{z=0} \tag{5.67}$$

Therefore, if light enters the device in the waveguide 1 the reflected filtered signal will emerge in waveguide 2, with no reflection in waveguide 1. Likewise, if light enters waveguide 2, the filtered signal will emerge in waveguide 1. In this way, the device spatially separates the input and the output signals without losing any of the reflected power from the gratings.

5.7 Summary

In order for a grating-based filter to be practical, it is necessary to somehow separate the reflected, filtered signal from the input signal. One way that this separation can be accomplished is by constructing an integrated Michelson interferometer, as depicted in Fig. 5.2. The interferometer consists of two identical gratings connected to a directional coupler which divides light between the two waveguides. Sections 5.2–5.5 describe how this directional coupler can be analyzed and designed to obtain the desired amount of power transfer between the waveguides.

The interaction between two parallel waveguides can be analyzed by treating the two waveguides as an aggregate structure which has a symmetric and an antisymmetric mode. The difference between the propagation constants of the symmetric and antisymmetric modes describe the rate at which power is transferred between the parallel waveguides. A more convenient method of modeling two parallel waveguides is to use coupled mode techniques. Coupled mode theory can be extended, as described in section 5.3, to account for the non-orthogonality of the waveguide modes. In a practical directional coupler the waveguides are not parallel, rather they are gradually brought together and apart over a curved path. Section 5.4 describes how coupled mode techniques can be used to model such a tapered structure.

Designing a tapered directional coupler requires that the coupling constant be known accurately for a range of waveguide separations. Section 5.5 describes three different methods of calculating the coupling constant μ, and discusses the range validity of these methods.

Chapters 3 and 4 present a complete description of the integrated Bragg gratings, and this chapter describes the physics of codirectional couplers. In order to model the behavior of the Michelson interferometer depicted in Fig. 5.2, we need to piece together the solutions for the directional coupler with those of the Bragg grating. This is most easily accomplished by expressing the solutions to the coupled mode equations in matrix form. Each segment of a device has a transfer matrix which relates the backward and forward mode amplitudes at the right edge of the segment to those at the left edge of the segment. The transfer matrix for a sequence of concatenated waveguide segments can be easily found by simply multiplying the matrices of the constituent segments. Section 5.6 demonstrates how the transfer matrix method is applied to the analysis of the Michelson interferometer.

CHAPTER 6: DESIGNING FOR MINIMUM LOSS

6.1 Loss Considerations

Chapters 3–5 describe the operation of integrated Bragg gratings and directional couplers, but they do not provide any concrete design specifications for practical devices. Chapter 4 describes the physics of dielectric waveguides, but it offers no explanation for how to select the index of refraction contrast and waveguide dimensions. Likewise, although Chapter 5 describes how to model a waveguide coupler, it gives no motivation for how to best select the parameters of the coupler such as the arc radius and bend angle (cf. figure 5.9). Based upon the material presented thus far, there are several possible choices for these waveguide parameters all of which would seem to yield equivalent devices. This chapter seeks to address the question of how to select the device parameters in order to obtain the optimal performance.

One of the primary figures of merit for any integrated optical device is the total insertion loss, defined as the total amount of attenuation suffered between the input fiber and the output fiber. Most optical device designs seek to minimize the total insertion loss wherever possible. The matched filters described in this thesis are meant to be used for filtering amplified spontaneous emission noise from an optical signal. For such a noise filter, minimizing insertion loss is even more important, because the filtered signal cannot be re-amplified without reintroducing amplified spontaneous emission noise.

There are several factors which contribute to the insertion loss of the device. One place where loss occurs is at the junction between the optical fiber and the waveguide chip. Additionally, there is some intrinsic attenuation associated with the passive material of which the waveguide is composed. If the waveguide walls are rough, granular, or in any way non-smooth, a portion of the bound light will be scattered into radiation modes, causing attenuation of the signal. Whenever the waveguide follows a curved path rather than a straight path, some of the light is radiated. Finally, the grating only reflects a fraction of the incident light; therefore the grating filter is inherently a lossy device. This chapter addresses each of these factors in order to seek a device design that yields the minimum possible waveguide loss while maintaining reasonable filter performance.

6.2 *Fiber Coupling Loss*

Fiber optic cables are the primary transmission medium for most optical communication systems. Therefore, in designing integrated optical components, it is important to consider how efficiently light is transferred from a fiber into the integrated waveguide and vice versa. One of the more common methods of launching light into an integrated waveguide is to cleave or polish the edges of the waveguide chip, allowing light to be end-fire coupled directly into the waveguide cross section on the edge of the chip. The fiber is often polished to a flat surface so that it can be abutted to the chip facet, a technique known as butt-coupling. If there is a gap of air between the fiber and the chip facet, reflections at the facet will occur because of the index of refraction discontinuity between the air and dielectric. These reflections are often reduced by polishing the fiber and chip facets at a slight angle [65], and also by inserting an index matching oil or epoxy between the fiber and the chip to reduce the index discontinuity [35, 65]. Even if these precautions are taken, the coupling between the fiber and the waveguide will not be efficient unless the modes of the two structures are similar. If there is a large mode mismatch between the optical fiber and the integrated waveguide, one of the structures must be altered to achieve good coupling. One solution is to taper the fiber to a point, which creates a lens that focuses the light emerging from the fiber [34]. Another approach is to taper the integrated waveguide so that the structure of the waveguide at the facet is better matched to the optical fiber [66, 67, 68]. Some authors have proposed using integrated periodic structures at the chip facet to expand the mode of the integrated waveguide in an effort to improve the coupling to an optical fiber [69].

The amount of loss incurred at an abrupt boundary between two dissimilar waveguides such as those depicted in figure 6.1 has been theoretically treated by many authors [15, 25, 70, 71]. In most cases, it is not possible to satisfy the electromagnetic boundary conditions at the interface between the two guides by using a simple solution consisting of forward and backward travelling modes in each region. For weakly guiding structures, it is approximately valid to calculate the coupling efficiency at the abrupt junction by evaluating the overlap integral of the transverse modes of the two waveguides. This result can be derived from expanding the mode of waveguide 1 in terms of the bound and radiation modes of waveguide 2:

$$\Psi(x,y) = \sum_i a_i \phi_i(x,y) + \int A(\beta)\phi(x,y;\beta)d\beta \tag{6.1}$$

where $\Psi(x,y)$ is the bound mode of waveguide 1, $\phi_i(x,y)$ are the discrete bound modes of waveguide 2, and $\phi(x,y;\beta)$ are the continuous radiation modes of waveguide 2 parameterized by their propagation constants β. In equation 6.1, the scalar mode approximation is used to represent the modes of both waveguides because for this work we are interested primarily in weakly guiding structures. Assuming that waveguides 1 and 2 have only one bound mode, the expansion coefficient a_1 of the bound mode in waveguide 2 is given by:

$$a_1 = \frac{\iint \Psi(x,y)\phi_1(x,y)dA}{\sqrt{\iint |\phi_1(x,y)|^2 dA}} \tag{6.2}$$

The expansion of equation 6.1 assumes that the mode of waveguide 1 can be treated as an initial condition for the wave propagation in waveguide 2, and it therefore neglects the effect of reflection back into waveguide 1. In this approximation, the efficiency of the power transfer between the bound modes of the two waveguides is governed by the expansion coefficient obtained when the mode of waveguide 1 is resolved in terms of the modes of waveguide 2. The transmission coefficient for the abrupt junction is simply the ratio of the power transferred to the first bound mode of waveguide 2 to the power contained in the bound mode of waveguide 1 [25]:

$$T \approx \frac{|a_1|^2}{\iint |\Psi(x,y)|^2 dA} = \frac{\left|\iint \Psi(x,y)\phi_1(x,y)dA\right|^2}{\iint |\phi_1(x,y)|^2 dA \iint |\Psi(x,y)|^2 dA} \tag{6.3}$$

The limitation of the mode overlap approach can be seen by considering the junction-coupling between the TE modes of two planar waveguides, as illustrated in figure 6.2. The two slab waveguides have the same geometrical size, but different indices of refraction. Recall from the analysis of Section 4.1.2, the transverse TE mode of a slab waveguide depends only upon the index contrast ($n_{core}^2 - n_{clad}^2$) between the core and cladding. Therefore, it is conceivable to have two slab waveguides which have identical mode profiles yet have different indices of refraction, provided the difference ($n_{core}^2 - n_{clad}^2$) is

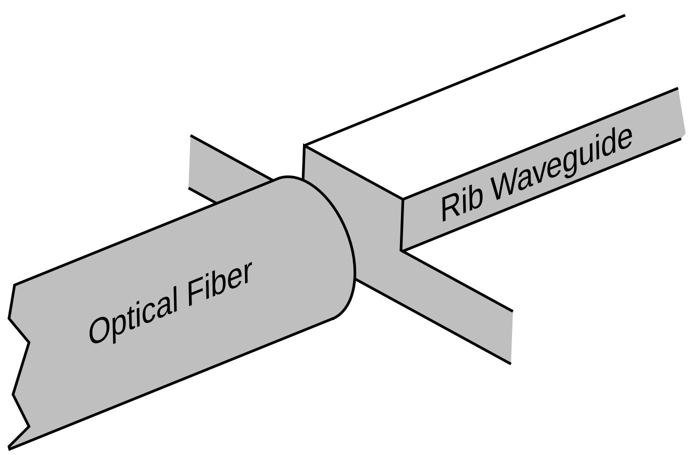


Figure 6.1 An example of an abrupt junction between dissimilar waveguides: an optical fiber coupled into an integrated rib waveguide. The amount of loss suffered at the abrupt junction depends upon how well the modes of the two structures overlap.

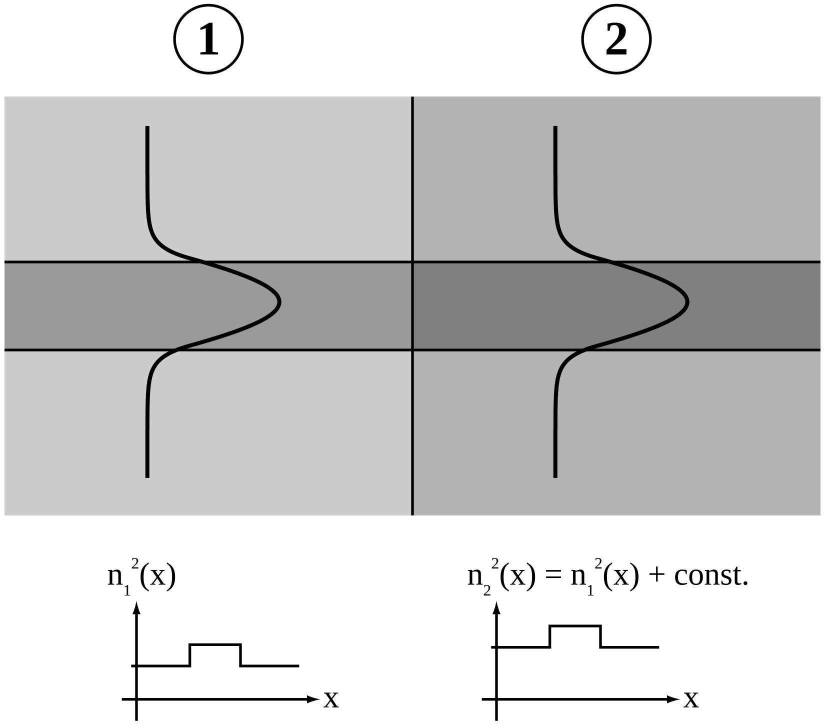


Figure 6.2 Junction between two slab waveguides. Both guides have the same slab width, but different index profiles. If the index contrast (n_{core}^2 - n_{clad}^2) is the same for the two waveguides, then the transverse modes will overlap perfectly, even if the guides are different, as illustrated in the above figure. Because the index of refraction of guide 2 is uniformly higher than in guide 1, perfect transmission cannot occur. This (albeit contrived) example illustrates that the mode overlap approach does not provide a complete description of junction loss when there is a large index difference between the two guides.

the same for both guides, as depicted in figure 6.2. This would be the case if the index profile in guide 2 were related to that in guide 1 by:

$$n_2^2(x) = n_1^2(x) + \text{constant} \tag{6.4}$$

From physical arguments, however, it is not possible for full transmission to be realized in such a structure because of the abrupt index discontinuity between the waveguides. (Consider that as the core-clad index contrast approaches zero, the reflection and transmission must be given by the Fresnel equations for plane wave incidence at a dielectric interface [24].) Because the index of refraction in waveguide 2 is uniformly higher than that of waveguide 1, some Fresnel reflection must occur at the boundary. Therefore, this somewhat contrived example illustrates one case where the modes of the two waveguides are perfectly matched, but full power transmission is not achieved. Some authors have proposed a modified method of calculating the coupling efficiency which contains a product of a Fresnel reflection term with a mode overlap term [72]. However, when the indices of refraction between the two waveguides materials are close, the mode overlap approximation provides an accurate estimate of the coupling efficiency between the waveguides [15]. Moreover, even if there is a substantial change in the bulk index of refraction between the waveguides, it is reasonable to assume that well-matched modes will couple more efficiently than poorly matched modes. Thus, although the mode overlap approximation does not completely describe the physics of abrupt waveguide junctions, it provides a practical starting point for waveguide design.

One of the advantages of using SiO_2-based materials for waveguide fabrication is that the index of refraction of glass materials is very closely matched to that used in optical fibers. Accordingly, it is possible to obtain very good matching between the mode of the optical fiber and that of a glass waveguide, if the waveguide is properly designed. We present here a simplified analytical model of the junction coupling loss between an optical fiber and a channel waveguide. Both waveguides are treated as weakly guiding structures (cf Section 4.1.3), so that the modes can be described by a single scalar function representing one of the transverse field components.

The mode of the optical fiber can be expressed analytically in terms of Bessel functions, but for simplicity, we approximate the mode of the fiber as a circular Gaussian mode of radius w:

$$\Psi(x,y) = \sqrt{\frac{2}{\pi w^2}}\exp\left(-\frac{x^2+y^2}{w^2}\right) \tag{6.5}$$

The Gaussian mode approximation for optical fibers is described completely in [15]. The scalar mode profile given in equation 6.5 has been normalized so that:

$$\iint \Psi(x,y)dA = 1 \tag{6.6}$$

For a channel waveguide, analytical solutions for the mode profile can be approximated by treating the structure as a superposition of two orthogonal slab waveguides, as illustrated in figure 6.3 [73, 74, 22]. This method ignores the fields in the shaded regions in the four corners outside of the core region. The advantage of this approximation is that for the index profile depicted in figure 6.3, the scalar wave equation is separable, that is, the solution may be written as a product of two orthogonal slab modes:

$$\phi(x,y) = \phi_x(x)\phi_y(y) \tag{6.7}$$

$$\phi_x(x) = \begin{cases} A_x\cos(k_x x) & |x| < \dfrac{d_x}{2} \\ B_x\exp(-\alpha_x|x|) & |x| > \dfrac{d_x}{2} \end{cases} \tag{6.8}$$

similar expression for $\phi_y(y)$

The slab modes are found by solution of the scalar wave equation for a TE mode of a planar slab waveguide, as demonstrated in Section 4.1.2. Note that because the channel waveguide considered in this work is a weakly guiding structure, the TE boundary conditions can be used for both the horizontal and vertical slab modes $\phi_x(x)$ and $\phi_y(y)$. The constants A_x and B_x are found by solving the eigenvalue

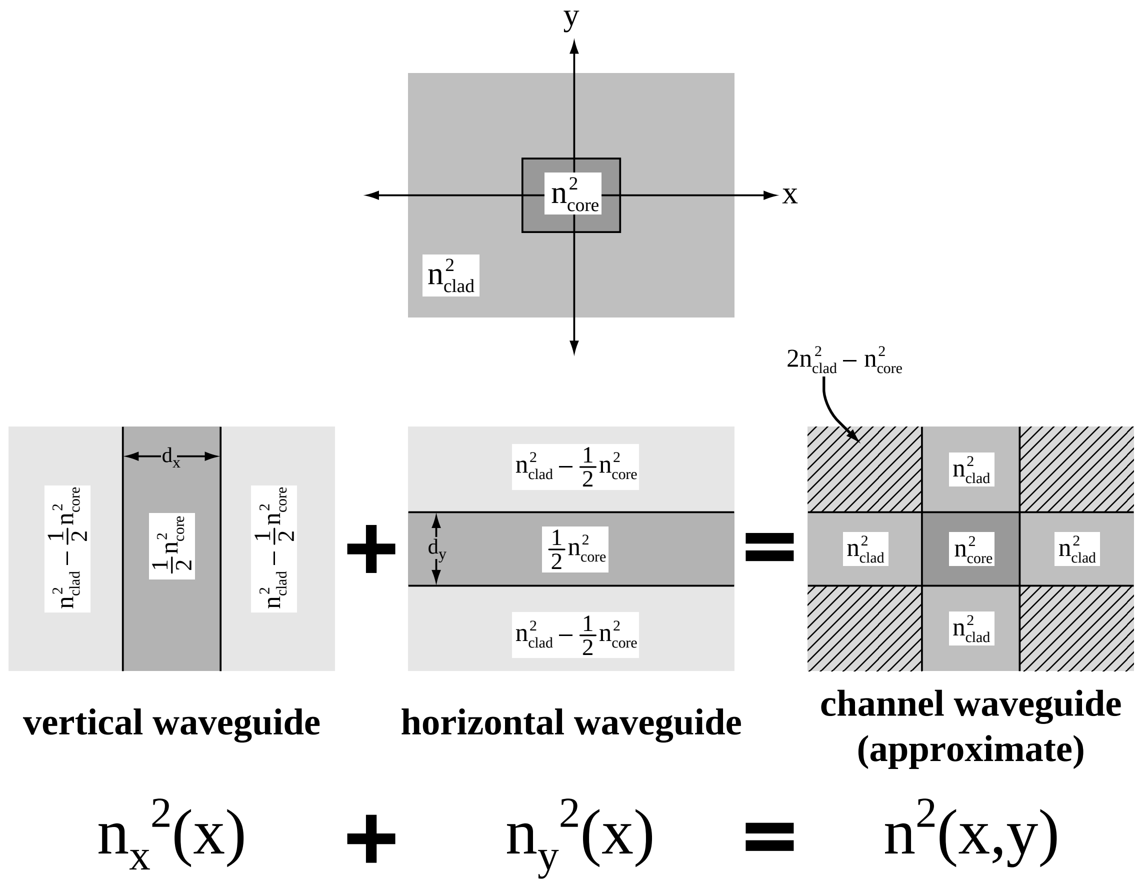


Figure 6.3 The electromagnetic field of the channel waveguide (depicted at the top of this figure) can be approximated as a superposition of a horizontal slab mode with a vertical slab mode. This approximation neglects the field in the hatched corner regions.

equation for the one dimensional slab mode, and requiring that the resultant mode be normalized to unity.

As with the approximate scalar mode of the channel waveguide, the approximately Gaussian mode of the optical fiber can be separated into a product of the x and y dependent functions:

$$\Psi(x,y) = \Psi_x(x)\Psi_y(x)$$
$$\Psi_x(x) = \left(\frac{2}{\pi w^2}\right)^{\frac{1}{4}} \exp\left(-\frac{x^2}{w^2}\right), \quad \Psi_y(x) = \left(\frac{2}{\pi w^2}\right)^{\frac{1}{4}} \exp\left(-\frac{y^2}{w^2}\right) \tag{6.9}$$

Because both of the scalar mode profiles are separable, the evaluation of the mode overlap (denoted η) between the fiber and the waveguide can be separately evaluated in the x and y directions:

$$
\begin{aligned}
\eta &= \iint \Psi(x,y)\phi(x,y)dA = \eta_x \eta_y \\
\eta_x &= \int \Psi_x(x)\phi_x(x)dx\,, \quad \eta_y = \int \Psi_y(y)\phi_y(y)dy
\end{aligned}
\tag{6.10}
$$

Therefore, the problem of calculating the coupling efficiency between a fiber and a channel waveguide is reduced to the evaluation of the one dimensional overlap integral of a Gaussian with a planar slab waveguide mode. We wish to express the result of this overlap integral in terms of the parameters of the slab mode (the slab width d, dielectric contrast $\Delta\varepsilon$, and optical frequency) and the parameters of the Gaussian fiber mode (w). The overlap integral between the slab mode and the Gaussian mode can be expressed as a function of two dimensionless parameters: V, the normalized waveguide frequency introduced in Chapter 4 in equation 4.17, and d/w, the ratio of the slab width to the Gaussian width.

Figure 6.4 is a contour plot of the overlap integral of a normalized Gaussian with a normalized slab waveguide mode. The x-axis describes the ratio d/w, and the y-axis is the normalized waveguide frequency parameter V for the slab waveguide. Note that when the value of V exceeds π, the slab waveguide will be multi-moded, as described in Section 4.1.2. Because we wish to design only single-moded waveguide structures we have plotted only the range $0 < V < \pi$ in figure 6.4.

The design problem can be stated as follows: given an optical frequency (k_0) and the Gaussian width w, what is the optimal value for the index contrast ($n_{core}^2 - n_{clad}^2$) and the slab width d. The two thick solid lines drawn in figure 6.4 represent lines of constant index contrast. The locus of points with the same index contrast ($n_{core}^2 - n_{clad}^2$) is a straight line passing through the origin, with slope:

$$
\text{slope} = k_0 w \sqrt{n_{core}^2 - n_{clad}^2}
\tag{6.11}
$$

The shaded region in figure 6.4 corresponds to points with less than 0.05 dB of estimated coupling loss. In order for the design operating point to lie within this low loss window, the index contrast must be selected so that it lies between the two solid lines depicted in the figure. Thus, good coupling can only be achieved for a narrow range of index contrasts. Once the index contrast is selected, the slab width d can then be chosen to achieve a suitably low coupling loss. The contour plot of figure 6.4 applies only for the one dimensional overlap of a Gaussian with a TE slab mode, but the same normalized plot can be employed to determine the amount of coupling loss associated with mode mismatch in the y-direction. Because the loss contours in figure 6.4 are expressed in decibels, the x and y components of the coupling loss can be simply added to obtain an estimate of the two dimensional overlap integral. Clearly, the optimal configuration would be a square channel waveguide, with equal mode mismatch in the x and y directions.

For the Corning SMF–28 optical fiber operating at 1.550 μm wavelength, the nominal Gaussian mode field diameter is 2w = 10.5 μm [75]. We have selected an index contrast of 0.3% and a channel size of 6.6 μm square. This design is represented by the marked spot in the contour plot of figure 6.4. The index contrast and channel size have been selected to place the operating point near the center of the low-loss window. The simple overlap analysis described above predicts that the junction coupling loss for this design will be under 0.1 dB (less than a 0.05 dB contribution from both the x and y overlap components.)

In the preceding analysis, we approximated the mode of the optical fiber as a Gaussian and the mode of the channel waveguide as a product of two orthogonal slab modes. These approximations proved useful because they allow the mode overlap integral to be separated into x- and y- components

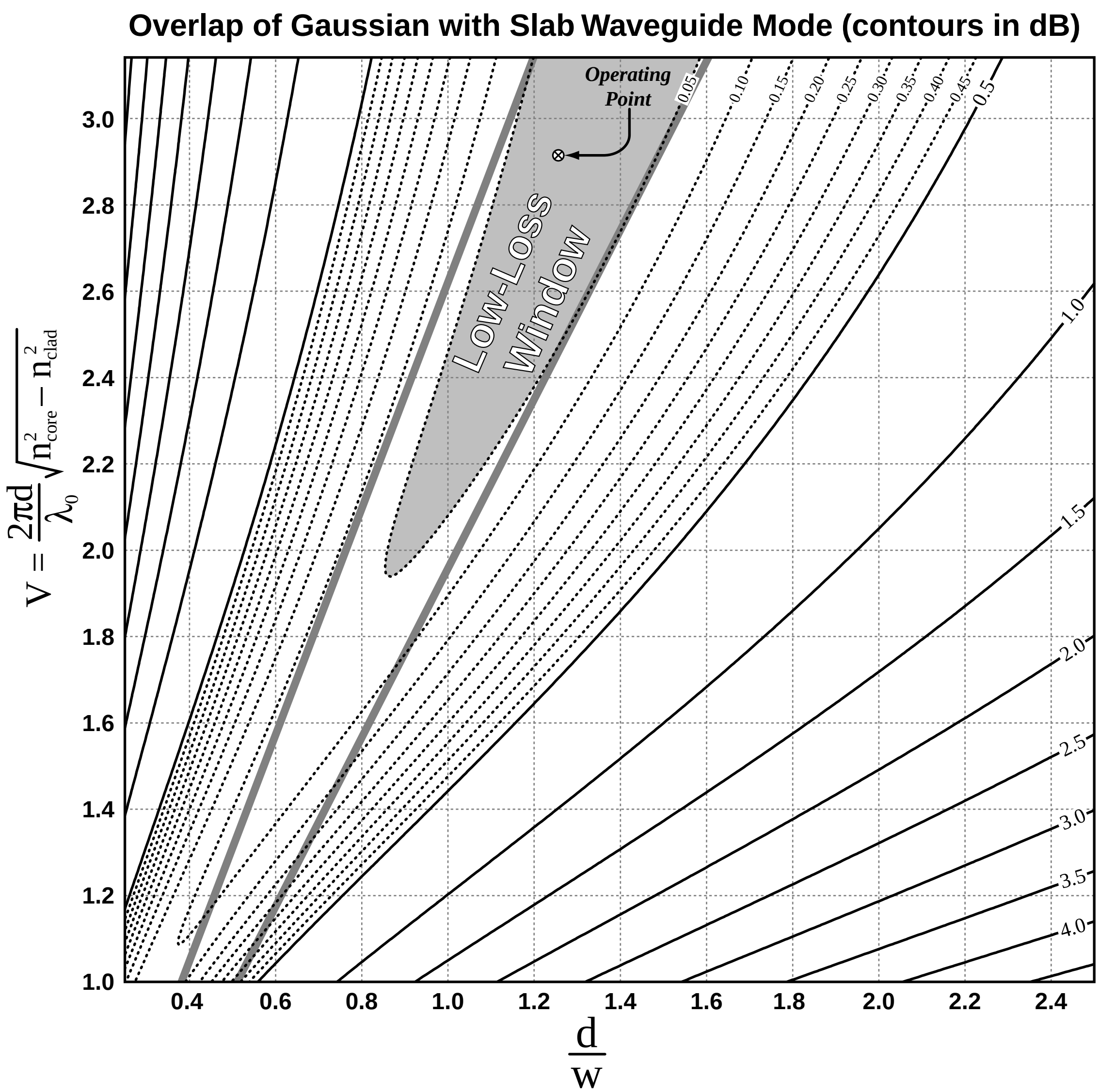


Figure 6.4 Contour plot of the junction coupling loss between a Gaussian mode and a 1-D slab waveguide mode. The coupling loss is found by computing the normalized overlap integral between the two modes. d represents the slab width, w is the gaussian field 1/e length, V is the normalized frequency parameter of the slab waveguide. The shaded region indicates the region of points for which the overlap is better than 0.05 dB. The straight solid lines are lines of constant index contrast.

which can be analytically evaluated. In order to obtain a more accurate measure of the mode overlap, it is necessary to use more accurate solutions for the modes of the fiber and the channel waveguide. The exact mode of the optical fiber can be expressed in terms of Bessel functions, but there is no such analytical expression for the modes of the rectangular channel waveguide. Therefore, in order to more accurately approximate the field of a channel waveguide, we have used a finite difference modesolver to solve for the mode at a series of discrete points in the waveguide cross section. Figure 6.5 presents a grayscale plot of the mode of the square channel waveguide next to the analytically calculated mode of the Corning SMF-28 optical fiber. The overlap integral between these two mode profiles can be approximated by a summation over the discrete mesh-points used by the modesolver. Using the more accurate modes of the fiber and the waveguide, we obtain a calculated junction coupling loss of 0.14 dB. This result is slightly higher than the junction loss predicted from the contour plot in figure 6.4. Nonetheless, the simple approximations described earlier provide a convenient and reasonably accurate starting point for waveguide design. (Note that junction losses of 0.10 dB and 0.14 dB correspond to normalized field overlaps of 0.9886 and 0.9840 respectively, and therefore the actual difference between the two losses is less than 1%.)

6.3 Bending Loss

The analysis of waveguides presented in Chapter 4 is founded upon the notion of a dielectric index profile which is a function of only the transverse coordinates x and y. However, most devices of interest (like the couplers described in Chapter 5) require waveguides which direct the light over a more complicated curved trajectory. If the curve is sufficiently gradual, the waveguide can be accurately treated as if it were straight waveguide, with phase fronts that are perpendicular to the propagation direction. If a waveguide is bent into a circular arc, it would seem that the phase fronts would be radial planes propagating in the tangential direction. However, this would mean that the phase velocity of the light (the speed at which the phase fronts propagate) is linearly proportional to the radial distance from the arc center. Therefore at some distance outside of the waveguide the phase velocity would exceed the local speed of light in the cladding material. Beyond this point the field can not remain in phase with the remainder of the mode, and it therefore breaks away from the bound mode and some power is radiated from the guide. In fact, a curved waveguide does not admit bound mode solutions like a straight waveguide. Whenever a waveguide is bent or curved, some of the light leaks out of the structure and is radiated away. This section describes a one-dimensional model which explains this phenomenon, and estimates the amount of attenuation or light leakage as a function of bend radius.

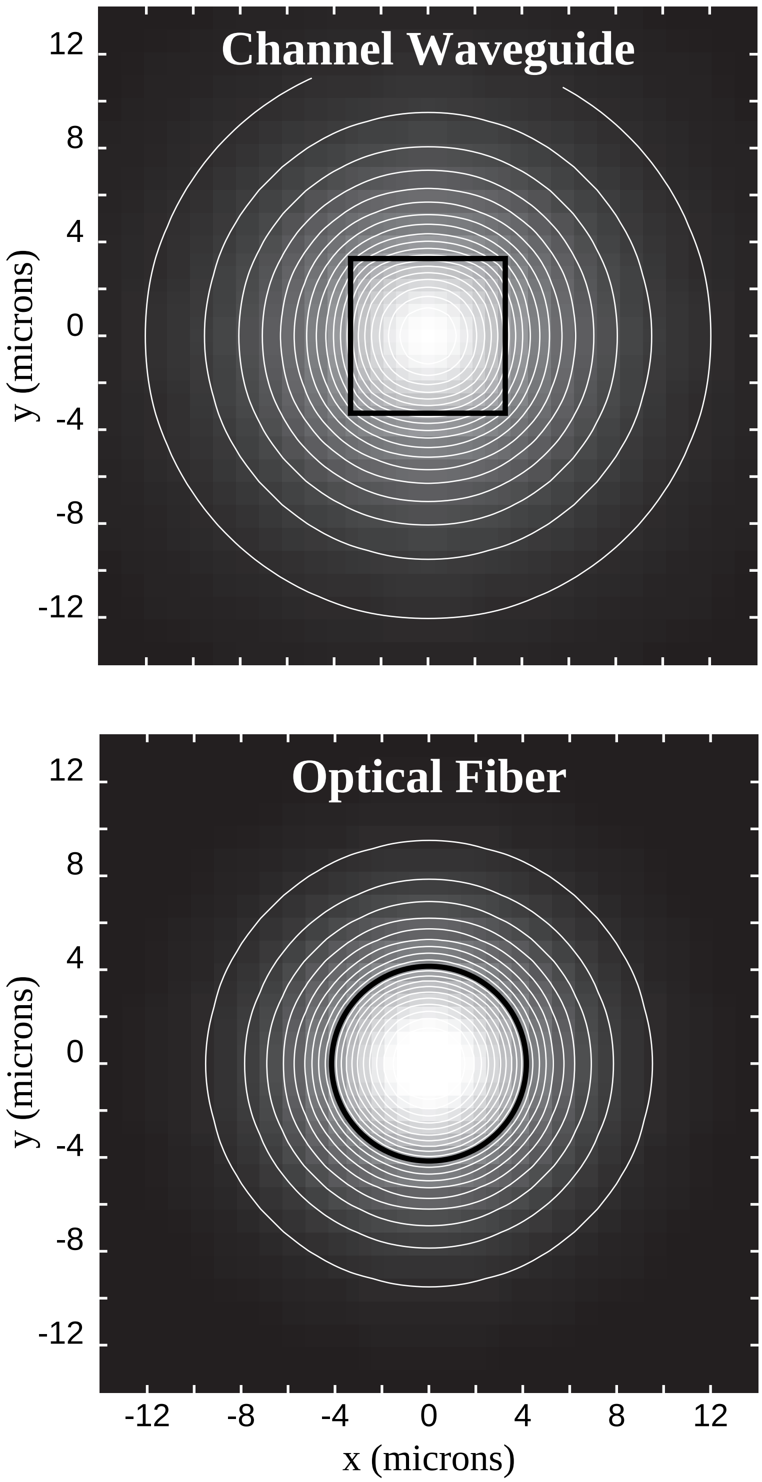


Figure 6.5 The modes of a channel waveguide (upper) and an optical fiber (lower). The grayscale level is proportional to the electromagnetic field of the mode. The mode of the channel waveguide was calculated numerically using a finite difference modesolver. The mode of the optical fiber was calculated analytically. The calculated coupling loss between these two modes is 0.14 dB.

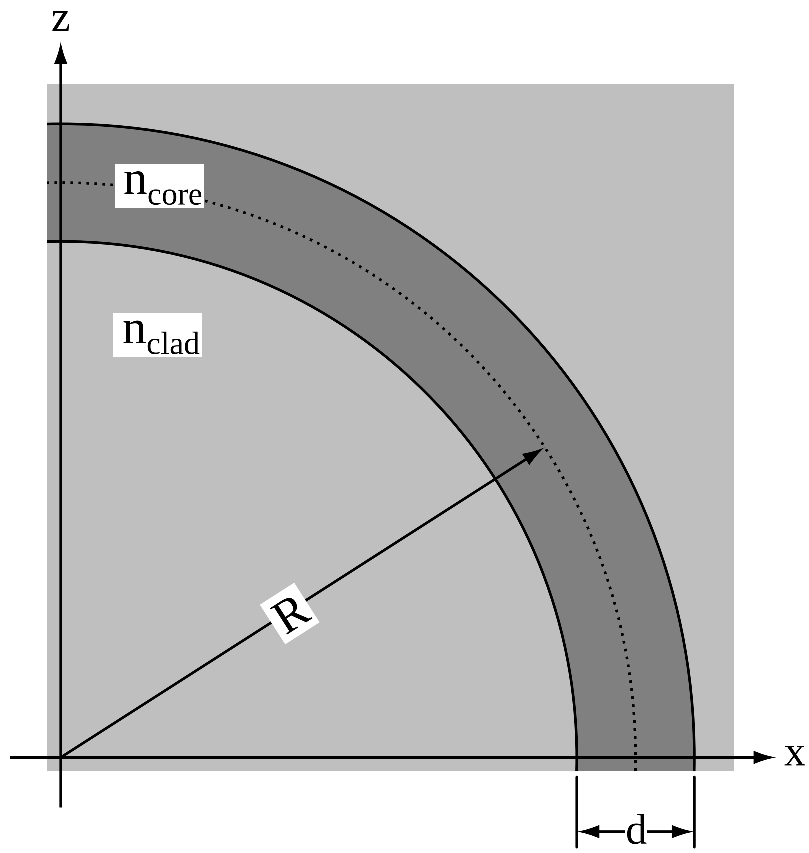


Figure 6.6 A bent slab waveguide. This waveguide consists of a uniform slab which is bent in the x-z plane, extending infinitely in the y direction (out of the page.)

6.3.1 Conformal Transformation

The simplest curved waveguide to analyze is the bent slab waveguide, as illustrated in figure 6.6. The index of refraction profile is a function of x and z only. The transverse electric field of the structure must satisfy the Helmholtz equation:

$$\left\{\nabla_{xz}^2 + k_0^2 n^2(x,z)\right\}\Psi(x,z) = 0 \tag{6.12}$$

Where $\Psi(x,z)$ represents the transverse electric field component E_y. $\Psi(x,z)$ and its derivative must be continuous at the boundaries between different media. The solution to this two dimensional problem can be simplified by applying a conformal transformation, which reduces the two dimensional problem to a one dimensional problem with an equivalent index profile [25, 72, 76, 77]. The conformal transformation approach seeks to find solutions in terms of a new pair of independent variables u and v, which are related to the cartesian variables x and z by the conformal mapping:

$$u + iv = R\ln\left(\frac{x + iz}{R}\right) \tag{6.13}$$

The conformal transformation described in the above equation can be thought of as a simple change of variables to a pair of variables that is more convenient for the cylindrically symmetric index profile.

$$\begin{aligned} r &= \sqrt{x^2 + z^2} = R\,\exp\left(\frac{u}{R}\right) \\ \phi &= \tan^{-1}\left(\frac{z}{x}\right) = \frac{v}{R} \end{aligned} \tag{6.14}$$

The variable u can be seen to represent variations in the radial direction, whereas v is related to variations in the angular direction. Therefore, after applying the conformal transformation of equation 6.13, the index of refraction is only a function of the variable u, and is independent of v. Equation 6.12 can be re-written in terms of the new variables u and v, making use of some of the properties of conformal transformations. First, the Laplacian operator is re-expressed in terms of the quantities u and v using the mapping of equation 6.13:

$$\nabla^2_{uv}\Psi\big(x(u,v), z(u,v)\big) = \exp\left(\frac{2u}{R}\right)\nabla^2_{xz}\Psi\big(x,z\big) \tag{6.15}$$

Substituting equation 6.15 into the equation 6.12 yields the following differential equation:

$$\left\{\nabla^2_{uv} + k_0^2 n^2(u)\exp\left(\frac{2u}{R}\right)\right\}\Psi(u,v) = 0 \tag{6.16}$$

Notice that the above equation is identical in form to equation 6.12, provided we replace the index of refraction by a modified index profile $\tilde{n}(u)$:

$$\tilde{n}(u) = n(u)\exp\left(\frac{u}{R}\right) \tag{6.17}$$

Therefore, the conformal transformation described in equation 6.13 transforms the bent slab waveguide into a straight waveguide as illustrated in figure 6.7. The problem is reduced from a two dimensional problem to an effective one dimensional problem, however the new index profile $\tilde{n}(u)$ is a continuously varying function of u rather than a piecewise constant function. Equation 6.16 can be solved by assuming a solution which is a travelling wave in the v direction:

$$\Psi(u,v) = \psi(u)e^{i\beta v} \tag{6.18}$$

Substituting equation 6.18 into equation 6.16 yields the following second order differential equation for the mode $\psi(u)$:

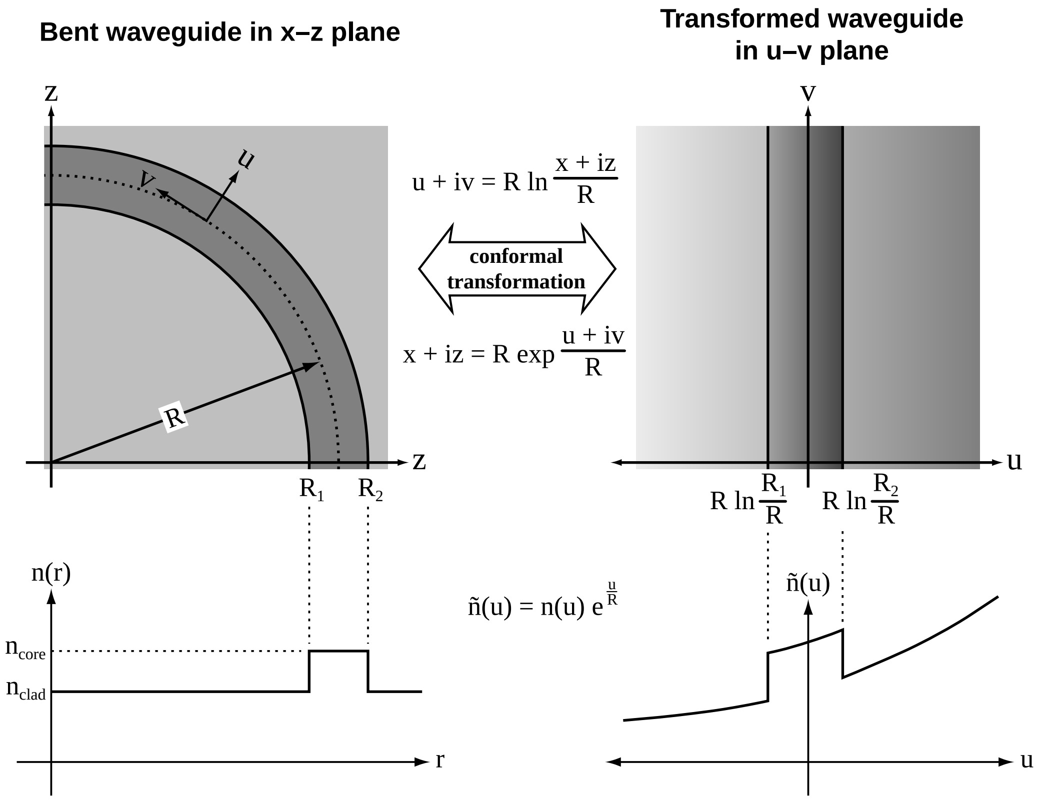


Figure 6.7 A conformal mapping transforms the curved waveguide in the x-z plane to a straight waveguide in the u-v plane, as illustrated. The transformed straight guide in the u-v plane is characterized by an index profile that is exponential.

$$\left\{\frac{d^2}{du^2} + k_0^2\tilde{n}^2(u) - \beta^2\right\}\psi(u) = 0 \tag{6.19}$$

The above equation is an eigenvalue equation, with eigenvalue β. When the solution to this equation is transformed back to the x-z plane, βR describes the angular propagation constant of the curved waveguide structure. Notice that the conformal transformation in equation 6.13 is parameterized by the constant R, which is conventionally chosen to be the radius of curvature of the waveguide (somewhere between R_1 and R_2). However, nothing in the analysis thus far has placed any constraints on how R is chosen. In fact, there is some arbitrariness in how R is selected; the conformal transformation is equally valid for any value of R [25]. However, when R is chosen to be the radius of curvature of the waveguide, the parameter v can be interpreted as the distance travelled along the axis of the curved waveguide. In many cases, the radius of curvature is much larger than the waveguide width, in which case the parameter u can be treated as the radial deviation from the radius R for points close to the waveguide core:

$$u = R\ln\left(\frac{r}{R}\right) = R\ln\left(1 + \frac{r - R}{R}\right) \approx r - R \ , \quad \text{for} \left|\frac{r - R}{R}\right| << 1 \tag{6.20}$$

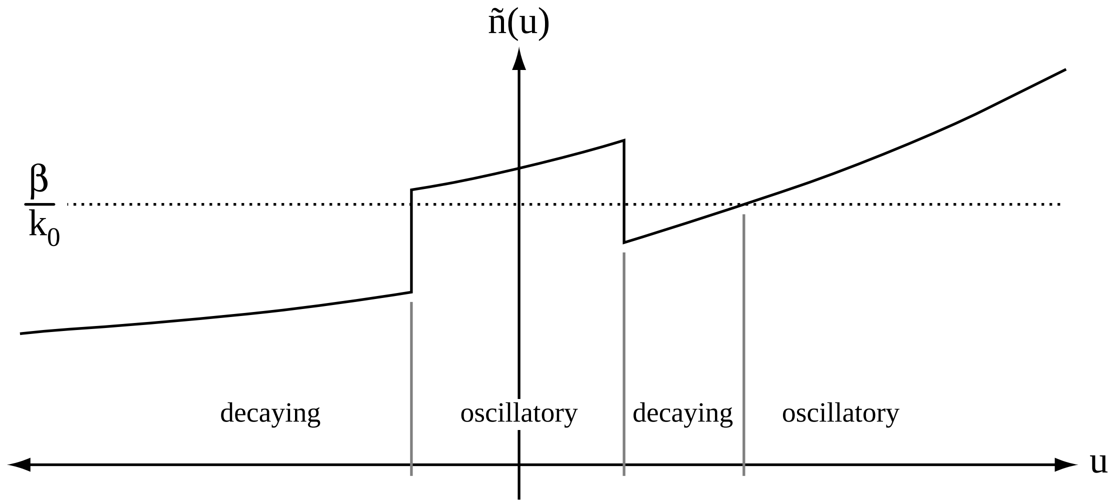


Figure 6.8 Because the index profile ñ(u) increases without bound as u increases, there are no bound modes for the bent waveguide. For any value of the propagation constant β, the mode solution is necessarily oscillatory as u approaches infinity.

Therefore, if we choose R to be midway between R_1 and R_2, we can treat u as the radial distance from the position R for all points in the vicinity of the waveguide core ($|r - R| << R$).

6.3.2 Leaky Modes an Resonances

Before describing the method used to solve equation 6.19, let us first examine the nature of the solutions. The index profile ñ(u) is monotonically increasing in the region to the right of the waveguide core, and exponentially decays to zero in the region to the left of the waveguide core. Because the index of refraction increases exponentially to the right of the core region, there are no bound modes. Recall from Chapter 4, that in order for a bound mode solution to exist, the waveguide mode must fall exponentially to zero in the regions outside of the waveguide core. Because ñ(u) increases without bound, the solutions to equation 6.19 will always be oscillatory in nature beyond a certain radial distance, as illustrated in figure 6.8. Therefore, the solutions to the equation 6.19 are not described by a series of discrete bound modes, but rather a continuum of radiation modes, characterized by oscillatory behavior in the far field. Nonetheless, the spectrum of the eigenmodes must somehow approach a discrete spectrum in the limit that the radius of curvature R becomes infinite.

The connection between the discrete modes of the unbent waveguide and the continuum radiation modes of the curved waveguide can best be understood by making an analogy with a Fabry-Perot cavity. Figure 6.9 illustrates a Fabry-Perot cavity comprised of two parallel reflective plates. If both of the mirrors are perfectly reflective, the cavity is characterized by a series of discrete modes where light is confined completely within the cavity. Only certain optical frequencies admit solutions which satisfy

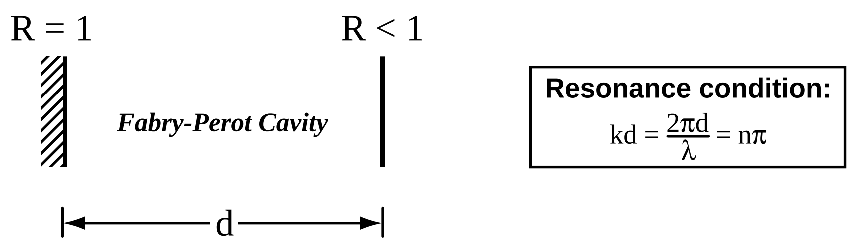


Figure 6.9 Fabry-Perot cavity. The behavior of a bent waveguide, with an index profile such as that depicted in figure 6.8 can be understood by making an analogy with a Fabry-Perot cavity. When the right mirror is not perfectly reflective, the cavity does not have discrete bound modes, but rather it is exhibits by a resonance behavior.

the boundary conditions (namely that the transverse electric field must be zero at the mirrors). However, if the right mirror is not perfectly reflective, the cavity is no longer characterized by a series of discrete longitudinal modes, but rather it has a continuous spectrum (i.e., any optical frequency is allowed within the structure.) However, even though any optical frequency admits a valid solution, certain optical frequencies are *resonant* with the cavity. The phenomenon of resonance is similar to that of a discrete bound mode, in that resonances occur when an integral number of half-wavelengths fit within the cavity. However, the resonant frequencies are characterized by a non-zero linewidth, which is related to the lifetime of photons within the cavity. Consider the case where a plane wave is incident on the Fabry-Perot cavity of figure 6.8 from the right-hand side. Because the left mirror is completely reflective, all of the power from the plane wave must ultimately be reflected by the structure, thereby creating a standing wave in the region to the right of the cavity. If we plot the ratio of the peak field intensity inside the cavity to the peak field intensity of the standing wave outside the cavity as a function of optical frequency, we find that only around the resonant frequencies is there an appreciable amount of light stored within the cavity. The resonance peaks of the Fabry-Perot cavity are characterized by a Lorentzian profile, whose width is related to the mirror reflectivity [14]. Figure 6.10 plots the resonance spectrum of the Fabry-Perot cavity for four different mirror reflectivities: 25%, 50%, 75%, and 95%. As illustrated in figure 6.10, the resonance linewidth decreases with increasing mirror reflectivity. Therefore in the limit that the right mirror becomes perfectly reflective, the resonance linewidth shrinks to zero and the spectrum of the cavity becomes discrete rather than continuous, as expected. Figure 6.11 plots the electromagnetic field of the Fabry-Perot cavity for points near resonance and points away from resonance, illustrating how only at points near the resonance is there an appreciable fraction of power stored within the cavity. Although the resonance is not, strictly speaking, a discrete mode of the Fabry-Perot cavity, it can be thought of as a leaky mode of the cavity. The linewidth of the resonance is related to the rate at which power leaks out of the cavity.

The bent waveguide structure depicted in figure 6.8 is analogous to the Fabry-Perot cavity discussed above. Although the structure has a continuous mode spectrum (i.e., any value of β admits a valid solution for $\psi(u)$), we expect that there will be one value of β which is *resonant* with the structure. More specifically, the resonance should have a Lorentzian profile, with a non-zero width which is related

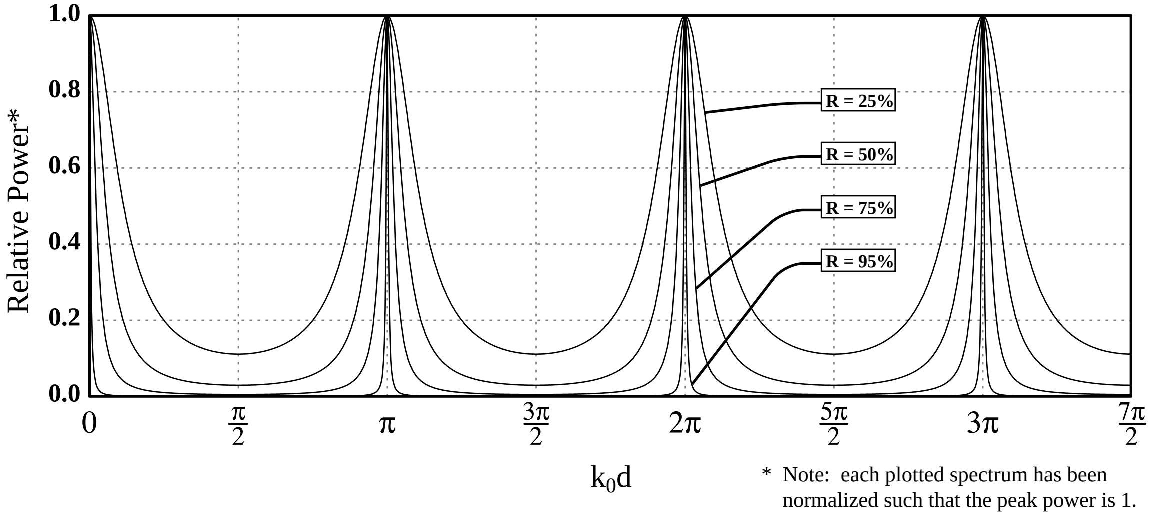


Figure 6.10 Resonance behavior of the Fabry-Perot cavity. When one of the mirrors is not perfectly reflective, the cavity no longer has discrete modes, but rather resonances with finite linewidths. The above figure plots the ratio of the peak field intensity inside of the cavity to that outside the cavity for four different mirror reflectivities: 25%, 50%, 75% and 95%. The width of the resonance peaks decreases as the mirror reflectivity approaches 1. In order to facilitate comparison of the four plots, each spectrum has been normalized to a peak value of 1.

to the radius of curvature of the structure [78]. As the radius R increases, the Lorentzian linewidth narrows so that in the limit that R becomes infinite, the structure is again characterized by a discrete bound mode solution. The Lorentzian linewidth of the resonance is directly related to the rate at which power leaks out of the structure. Therefore, by analyzing the spectrum of the bent waveguide index profile, we can determine the rate of power loss and therefore the attenuation coefficient of the curved waveguide.

6.3.3 Methods of Solving for Mode Spectrum

Because the index profile $\tilde{n}(u)$ is exponentially varying, the solutions to equation 6.19 are not as simple as for a straight slab waveguide where the index of refraction is constant within the three regions. However, approximate solutions can be found by replacing the continuously varying profile with a series of thin slabs, each of which has a uniform index of refraction, as illustrated in figure 6.12. The analysis of the slab waveguide presented in Section 4.1.2 can be extended to treat an arbitrary series of homogeneous stratified layers. The form of the solution is assumed to be exponentially decaying in the left-most region of the structure. By proceeding from left to right and applying boundary conditions at each interface, it is possible to calculate the mode $\psi(u)$ in each region of the structure, for any value of the propagation constant β. Matching boundary conditions does not yield a series of discrete eigenvalues βi. *Any* value of the propagation constant β yields a valid solution to the eigenvalue equation. However, by applying the boundary conditions, we can determine the ratio of the power confined within the core region to amplitude of the standing wave in the right-most region of the structure. If this ratio is

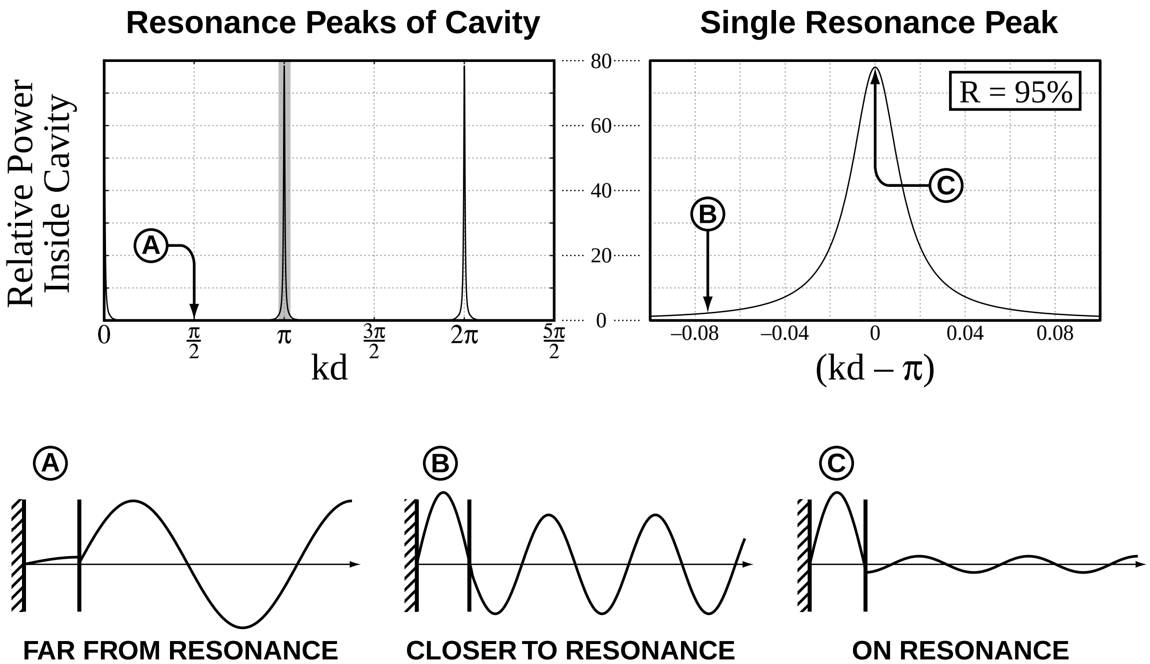


Figure 6.11 Electromagnetic field inside Fabry-Perot cavity at three different frequencies. The upper portion of the figure plots the spectrum of the cavity (the right figure is a magnified plot of the first resonance peak, illustrating the Lorentzian lineshape.) These spectra plot the ratio of the peak field intensity within the cavity to that outside the cavity as a function of frequency (kd). The lower portion of the figure plots the electromagnetic field of the cavity at the three frequencies indicated. (A) Far away from resonance there is very little power stored in the resonator compared to the field intensity outside the cavity. (B) Closer to resonance. (C) On resonance a relatively large amount of power is present inside the cavity.

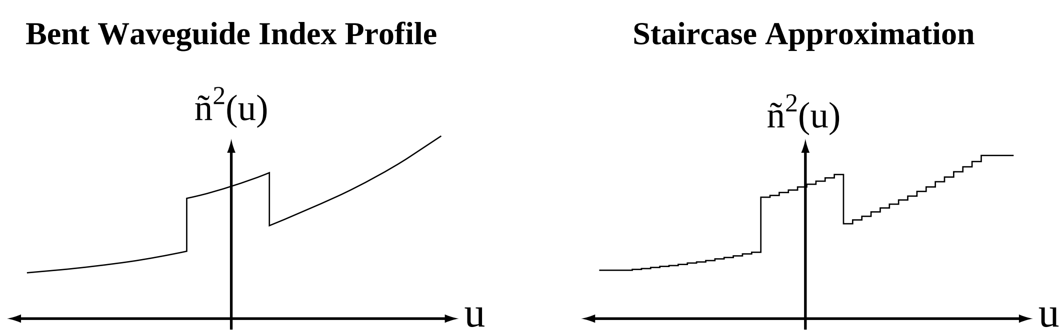


Figure 6.12 One way to treat the graded index profile is to approximate it with a series of homogenous slabs, as illustrated in the left figure.

plotted as a function of β, the resulting spectrum has a Lorentzian profile. Thus, the structure exhibits a resonance behavior as expected: certain values of β yield solutions for $\psi(u)$ in which a large portion of the power is confined within the waveguide core region. The Lorentzian linewidth then describes the attenuation rate of the waveguide. This approach was described by Thyagarajan et al. [79, 80].

When the exponential index profile is approximated by a series of thin slabs of uniform index, the form of the solutions in each region can be expressed in terms of sine's cosine's and exponential functions. The simple functional form of the solutions makes it relatively easy to match boundary conditions, however in order to obtain a good approximation of the index profile $\tilde{n}(u)$, it is necessary to use a series of many thin slabs covering the core region and extending well into the cladding regions. An improved method of analysis is to approximate the function $\tilde{n}^2(u)$ as a series of piecewise linear functions, as described in [81]. Often, it is sufficient to approximate the index profile $\tilde{n}^2(u)$ with three piecewise linear functions, as depicted in figure 6.13. The advantage of this approach is that boundary conditions must only be matched at two places, and the linear function provides a better approximation to the exponential index profile in the regions outside of the core, especially for gradual bends. The disadvantage is that the solutions to the Helmholtz equation in each region are no longer simple exponential or sinusoidal functions. Rather, the functional form of the solutions in each region is described by Airy functions. This method of calculating bending loss is described completely in [81], therefore in this chapter we present only a brief explanation of the method and explain some of the numerical techniques which were employed. Wherever possible, we have employed the notation of [81].

The refractive index profile in the i^{th} region is approximated by a linear function:

$$\tilde{n}^2(u) = N_i^2 + \alpha_i u \tag{6.21}$$

where N_i^2 is the value of $\tilde{n}^2(u)$ at the left edge of the i^{th} region, α_i is the slope of the index profile in the i^{th} region, and u represents the distance from the left edge. In region 1, N_i^2 represents the value of $\tilde{n}^2$ at the right edge of the region and u is the distance from this point. When the bend radius is large com-

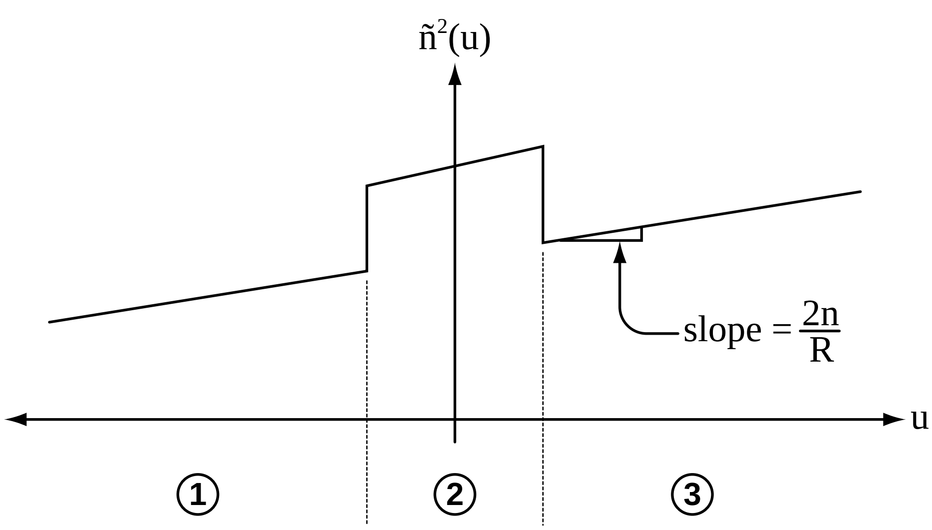


Figure 6.13 An improved treatment of bent waveguide index profile. $\tilde{n}^2(u)$ is approximated by a series of linear functions. Often it is sufficient to use only three piecewise linear functions. The slope of the index profile is inversely related to the radius of curvature of the bend.

pared to the waveguide dimensions, the index profile $\tilde{n}^2(u)$ can be expanded to first order in the quantity u/R:

$$\tilde{n}^2(u) = n^2(u)\exp\left(\frac{2u}{R}\right) \approx n^2(u)\left(1 + \frac{2u}{R} + \ldots\right) \tag{6.22}$$

Therefore, the slope of the index profile in the i^{th} region can be written as:

$$\alpha_i = \frac{2n_i}{R} \tag{6.23}$$

Substituting the linear index profile into equation 6.19 gives:

$$\left\{\frac{d^2}{du^2} + k_0^2\left(N_i^2 + \alpha_i u\right) - \beta^2\right\}\psi_i(u) = 0 \tag{6.24}$$

If we then apply a change of variables, equation 6.24 can be cast into the following form:

$$\left\{\frac{d^2}{dZ_i^2} - Z_i\right\}\psi_i(Z_i) = 0 \tag{6.25}$$

where the new dimensionless independent variable Z_i is linearly related to u in the following way:

$$Z_i \equiv \frac{\beta^2 - k_0^2 N_i^2 - \alpha_i u}{(k_0^2\alpha_i)^{\frac{2}{3}}} \tag{6.26}$$

Equation 6.25 has well known analytic solutions, which are the Airy functions Ai(Z) and Bi(Z) [82]:

$$\psi_i = C_i \mathrm{Ai}(Z_i) + D_i \mathrm{Bi}(Z_i) \tag{6.27}$$

In the left-most region, the coefficient D_1 must be zero in order for the function $\psi(u)$ to decay to zero as u approaches $-\infty$. C_1 may be set to 1, because the mode profile $\psi(u)$ can only be determined up to an arbitrary proportionality constant. Once C_1 and D_1 are known, the remaining coefficients C_2, D_2, C_3 and D_3 can be determined by requiring that $\psi(u)$ and its first derivative be continuous at the two boundaries. As with the staircase approximation described earlier, the process of matching boundary conditions can be reduced to a set of 2 x 2 matrix equations relating the coefficients in adjacent regions [80, 81]. An analysis of the mode spectrum can again be carried out in the same way described above. The coefficients C_2 and D_2 provide a measure of how much power is confined in the core region of the waveguide, whereas the coefficients C_3 and D_3 describe the amplitude of the standing wave pattern in the right-most region of the structure. If the ratio:

$$\frac{C_2^2 + D_2^2}{C_3^2 + D_3^2} \tag{6.28}$$

is calculated as a function of the propagation constant β, the resulting spectrum has a Lorentzian lineshape whose width describes the amount attenuation suffered along the waveguide arc segment.

$$\frac{C_2^2 + D_2^2}{C_3^2 + D_3^2} \cong \frac{\text{const.}}{(\beta - \beta_r)^2 + \frac{\Gamma^2}{4}} \tag{6.29}$$

Γ is the full-width at half maximum (FWHM) of the Lorentzian*. When the curved waveguide is initially excited with the bound mode of the equivalent straight waveguide, the curved waveguide will carry a wavepacket which is centered about the resonant value of β. If, after propagating a distance v the wavepacket is then resolved back into the bound mode of the straight waveguide, the power that remains in the straight waveguide is:

$$P(v) = P(0)\exp\left(-\Gamma v\right) = P(0)\exp\left(-\Gamma R\phi\right) \tag{6.30}$$

Therefore, the Lorentzian FWHM describes the rate at which power leaves the waveguide as it propagates through the structure.

To summarize, a bent waveguide structure does not have a series of discrete modes, but rather it is characterized by a continuous spectrum, with one or more resonances corresponding to the bound modes of the equivalent straightened waveguide. When light from a straight waveguide is launched into the bent waveguide, the initial mode excitation of the bent waveguide forms a Lorentzian wave packet. The Lorentzian wave packet propagates and spreads in such a way that when the wavepacket is projected back onto the bound mode of the straight waveguide after a distance v, the power remaining in the bound mode is attenuated exponentially. The exponential attenuation factor Γ is simply the Lorentzian FWHM, as described in equation 6.30.

6.3.4 Junction Losses

The preceding analysis describes how the attenuation of bent slab waveguide can be approximated by analyzing the resonances of the bent waveguide structure. Another place where losses can occur is at the junction between a straight waveguide and the bent waveguide. As described in Section 6.3.1, the mode of the bent waveguide can be derived by using a dielectric profile which includes an exponential factor as in equation 6.17. Accordingly, the resonant modes of a bent waveguide cannot be expected to match perfectly with the bound modes of the equivalent straight waveguide. This causes a mode mismatch, which can be another source of loss, as described in Section 6.2. One of the more noticeable differences between the leaky modes of the bent waveguide and the bound modes of the

* A direct calculation of the fractional power contained in the waveguide core is difficult to evaluate analytically when the mode is expressed in terms of Airy functions. However, near resonance the power in the core is proportional to the ratio of the squared amplitudes as given in equation 6.28 [78]. The proportionality factor is not important because the bending loss is related only to the width of the resonance peak.

equivalent straight waveguide is that the bent waveguide modes are displaced towards the outer boundary of the guide [72, 83, 84]. It was originally suggested by Neumann that one method to reduce these junction losses is to add an abrupt displacement to the waveguide at the junction between straight and curved segments in order to account for the outward mode displacement in the curved waveguide [85]. Some authors have reported reduced bending loss by addition of such an offset [86].

The method of calculating bending losses described in Section 6.3.3 can also be used to optimize the junction coupling loss. Once the resonance spectrum of the bent waveguide has been solved, it is straightforward to evaluate the mode $\psi(u)$ at the center of the resonance peak. Actually, it is slightly misleading to refer to the function $\psi(u)$ as a mode because it is not a normalizable bound mode, rather it is a radiation mode of the bent waveguide evaluated at the resonant value of β. Some authors refer to this resonant function $\psi(u)$ as a quasi-mode or a leaky mode [15, 78]. The quasi-mode closely resembles the bound mode of the straight waveguide, with the exceptions that (1) it is displaced towards the outer boundary of the guide, and (2) it exhibits oscillatory behavior beyond a certain radius instead of decaying monotonically. The junction coupling loss can be minimized by finding the junction offset value which gives the best overlap between the bound mode of the straight waveguide and the quasi-mode of the bent waveguide.

6.3.5 Calculated Bending Loss Results

One of the limitations of the bend loss analysis presented in this chapter is that it applies only to slab waveguides, or more generally one dimensional index profiles. More complicated structures can be treated by using the effective index method, which is described completely in [25, 48]. The effective index method seeks to reduce a two dimensional refractive index profile to an effective one dimensional index profile. Figure 6.14 illustrates how the effective index method would be applied to a rib waveguide. We first treat the variations in the vertical direction, by separating the waveguide into three vertical stacks. Each of these stacks can be treated as a slab waveguide and the analysis of Chapter 4 can be applied to determine the propagation constants of the bound modes of each vertical stack. Having found the propagation constants in each of the three vertical stacks, we define an effective index for each region as the ratio β/k_0 for each stack. Using these three effective stack indices, we then construct an equivalent horizontal stack. The same analysis is then applied in the horizontal direction, using the effective indices derived from each of the vertical stacks. Thus, the effective index method seeks to reduce a two-dimensional dielectric index profile to an equivalent one-dimensional profile by defining an effective index of refraction for each of the horizontal regions of the structure. For a channel waveguide, such as the one analyzed in this work, the outer horizontal regions are homogeneous, and therefore there is no vertical dielectric stack in these regions. The effective index of these regions is taken as the cladding index n_{clad}.

After reducing the channel waveguide to an equivalent slab waveguide structure, the bending loss analysis can be applied to the equivalent slab waveguide in order to approximate the degree of bending loss, using the method of analysis described in Section 6.3.3. Figure 6.15 illustrates the effective one-dimensional index profile for the 6.6 μm square channel waveguide considered in this thesis. Superposed on this plot is a plot showing the index profile ñ(u) of the transformed bent waveguide for a radius of 35 mm. Note that because the radius of curvature is orders of magnitude larger than the mode size, the function ñ(u) is linear over the region of interest.

To calculate the bending loss, the index of refraction profile $\tilde{n}^2(u)$ is approximated by a series of piecewise linear functions. For a bent waveguide, it is often sufficient to use only three linear regions, but the numerical algorithm which was implemented (using MATLAB) has the capability to handle an

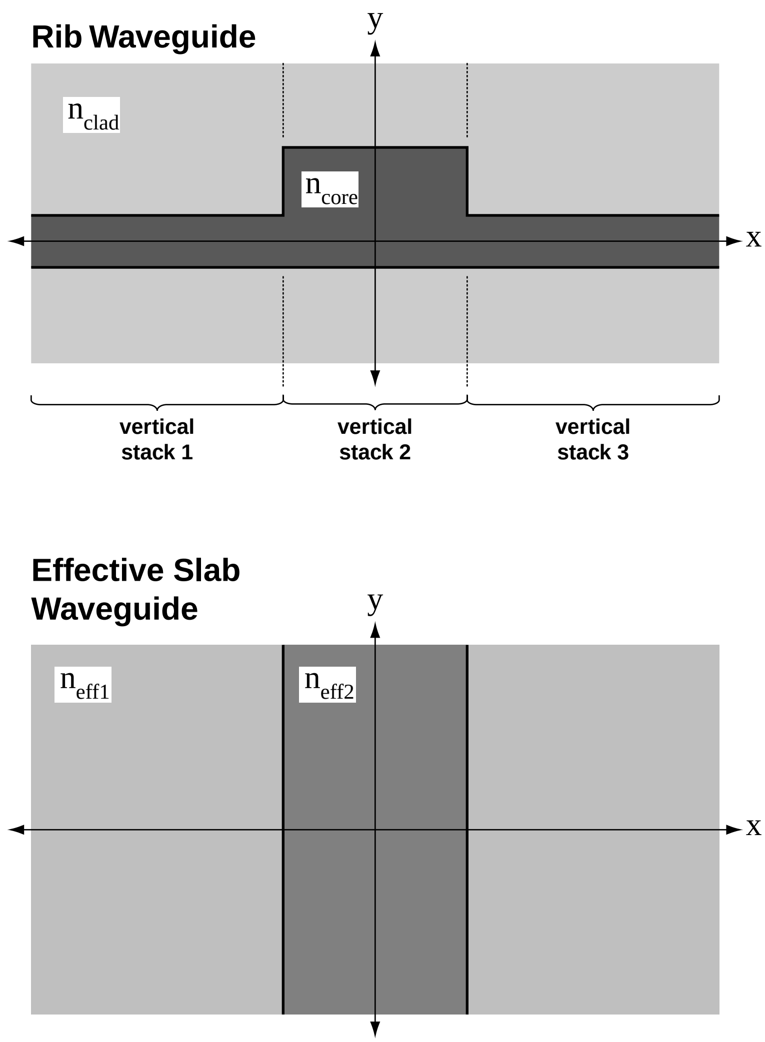


Figure 6.14 The effective index method reduces a two-dimensional refractive index profile into an effective one dimensional index profile. Each of the three horizontal regions is treated as an effective vertical dielectric stack. The effective indices of each vertical stack are used to construct an effective slab waveguide in the x direction.

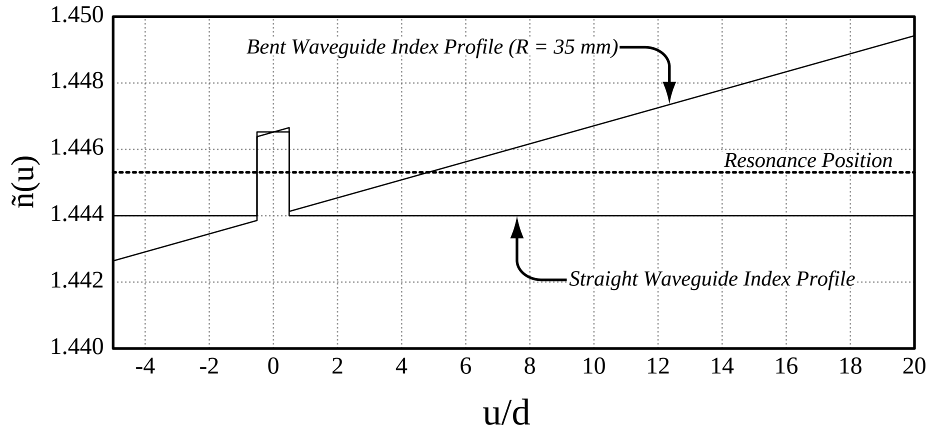


Figure 6.15 The effective index method is applied to a 6.6 μm square channel waveguide, reducing it to an effective one dimensional slab waveguide. When this guide is bent with a radius of 35 mm, the index profile is transformed into the sloped profile indicated.

arbitrary number of regions. The Airy functions Ai and Bi and their derivatives were calculated in terms of the Bessel functions of fractional order, as described in [82]. By matching boundary conditions at the edge of each linear region, the program calculates the mode profile $\psi(u)$ for several different values of the propagation constant β. For each value of β, the ratio of the confined power to the standing wave amplitude in the right-most region is calculated. A non-linear least squares analysis is used to fit the resulting data to a Lorentzian shape, and the FWHM is then associated with the bending loss. Figure 6.16 illustrates the Lorentzian shape of the resonance curve, calculated using the method described above, along with the best fit Lorentzian.

It is illustrative to examine the nature of the solutions $\psi(u)$ at points near resonance and away from resonance. Figure 6.17 presents a sequence of plots of the solutions $\psi(u)$ at four different frequencies. For comparison, the mode of the unbent waveguide is also plotted. At points far away from resonance (plot A), the mode amplitude within the core region is small compared to the amplitude of the oscillations outside the core. As the propagation constant β approaches the resonance peak, the function $\psi(u)$ becomes more confined within the waveguide core. At resonance, the leaky mode $\psi(u)$ can be seen to closely resemble the bound mode of the straight waveguide, the only difference being the small oscillations outside the core, and a slight displacement of the mode. The solutions plotted in figure 6.17 correspond to the structure depicted in figure 6.15, with a 35 mm. radius of curvature.

Figure 6.18 plots the calculated bending loss for the glass channel waveguide design which we will be using for constructing the matched filters. The amount of bending loss decreases exponentially with increasing radius of curvature. For the radii of curvature used in these filters, the calculated mode displacement is quite small, and therefore we have chosen not to include a junction offset (as described in Section 6.3.4) in the designs.

6.4 *Optimum Directional Coupler Design*

Because the degree of bending loss decreases exponentially with radius of curvature, it might seem that the best strategy for device design would be to use the largest possible radius of curvature. However, there are other sources of loss which factor into the total device loss. For example, even in a straight waveguide there is some attenuation, perhaps due to intrinsic loss in the materials, or roughness in the walls of the channel waveguide. For glass waveguides, these losses are expected to be approximately 0.1 dB/cm [41]. To minimize the total intrinsic waveguide loss, devices should be designed to be as short as possible. There is a trade-off between bending loss and intrinsic waveguide loss in that devices which use a large radius of curvature or a small bend angle tend to be relatively long.

One of the constraints when designing a directional coupler (such as those described in Chapter 5) is that the waveguides must be sufficiently separated at the edge of the chip to allow two adjacent optical fibers to be butt-coupled into the waveguides. Although fiber cores are only a few microns in size, the diameter of the fiber cladding is typically 125–150 µm. Most fiber mounting stages which hold pairs of optical fibers are designed to hold the fibers in etched V-grooves separated by 250 µm center to center. Therefore, one of the design constraints for a directional coupler device is that the waveguides

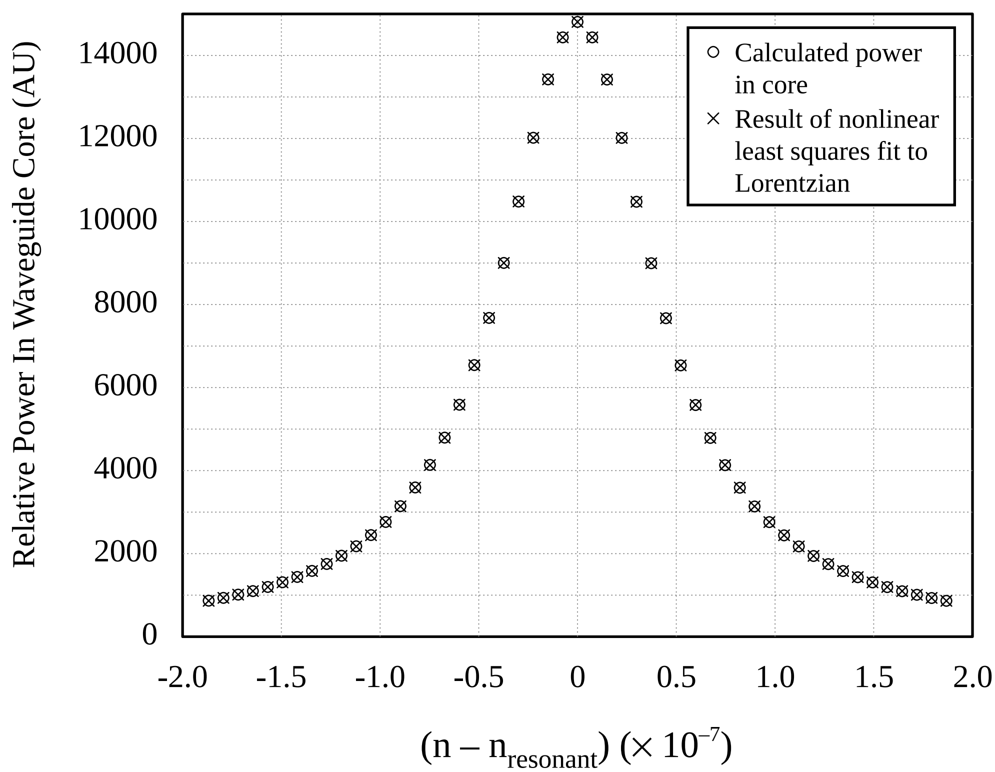


Figure 6.16 Resonance spectrum for bent waveguide. This spectrum was calculated by using piecewise linear approximation to the index profile $\tilde{n}^2(u)$. Superposed on the plot is the best fit Lorentzian. The radius of curvature for this plot was 35 mm.

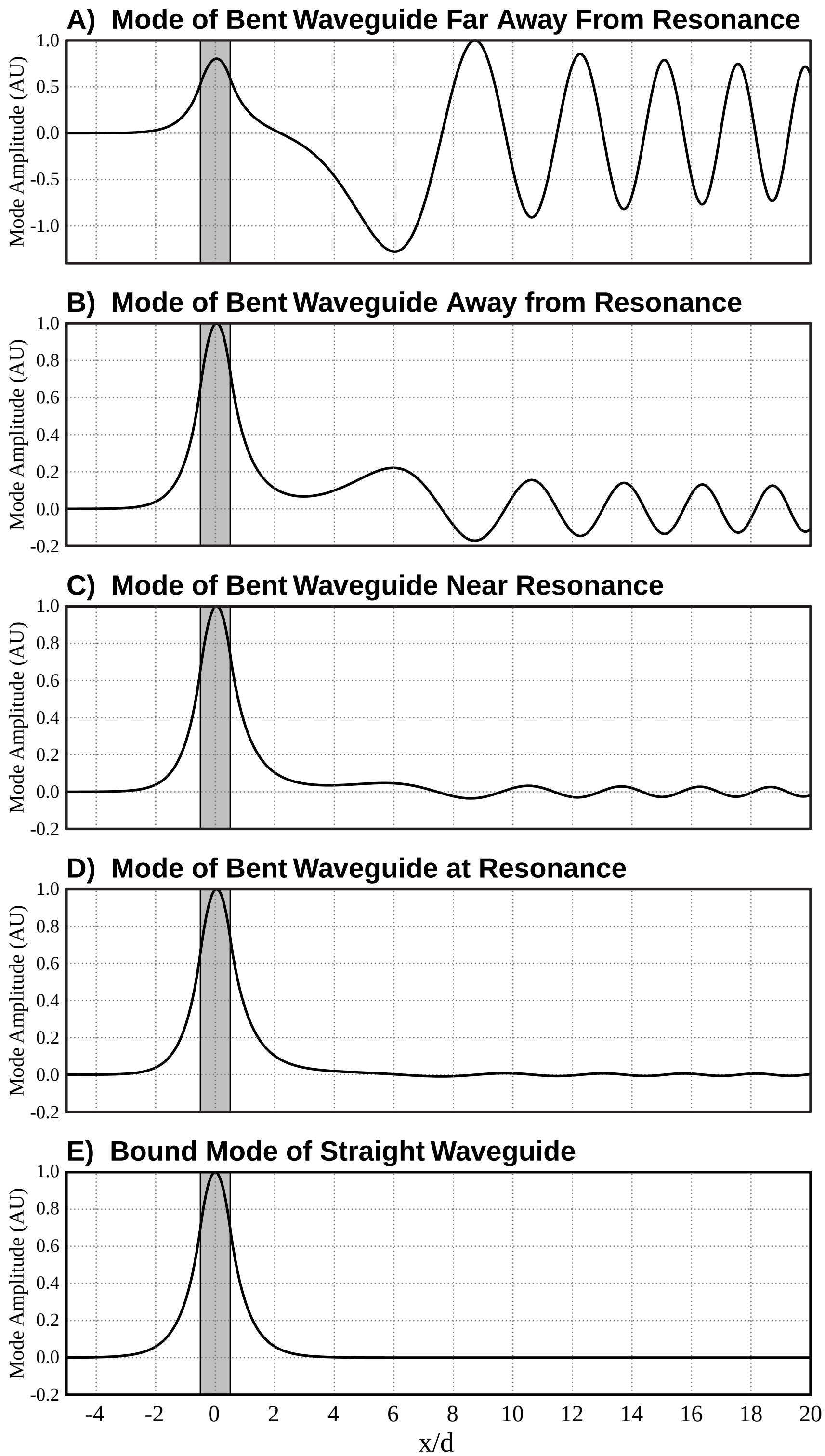


Figure 6.17 Solutions to the Helmholtz equation for the bent channel waveguide at resonance (plot D) and away from resonance (plots A–C). For comparison, the bound mode of the straight waveguide is presented in plot D.

must reach a separation of 250 μm at the edges of the chip in order to be compatible with commercially available paired fiber mounts. In designing the directional coupler, it is necessary to select the radius of curvature and bend angle in such a way that the total loss of the device (including bending loss and material loss) is minimized.

6.4.1 Optimized S-bend

Figure 6.19 depicts an S-shaped waveguide bend, described by a bend radius R, bend angle θ, and a total vertical waveguide separation of Δd. Such an S-shaped bend is a fundamental building block of waveguide couplers. For the moment, we shall take R and Δd to be fixed parameters, and consider the problem of how to best choose the bend angle θ so that the total loss is minimized. There is a trade-off between selecting a small bend angle (for which the bending loss is small, but the device must be made longer to achieve a total vertical displacement of Δd) and selecting a large bend angle (for which the bending loss is proportionately larger but the total path length of the device is shorter.) We shall assume that the intrinsic waveguide loss is described by an attenuation length z_m, and that for the specified radius of curvature the bending loss is characterized by an attenuation length z_a^*. The total path length of the S-bend is given by:

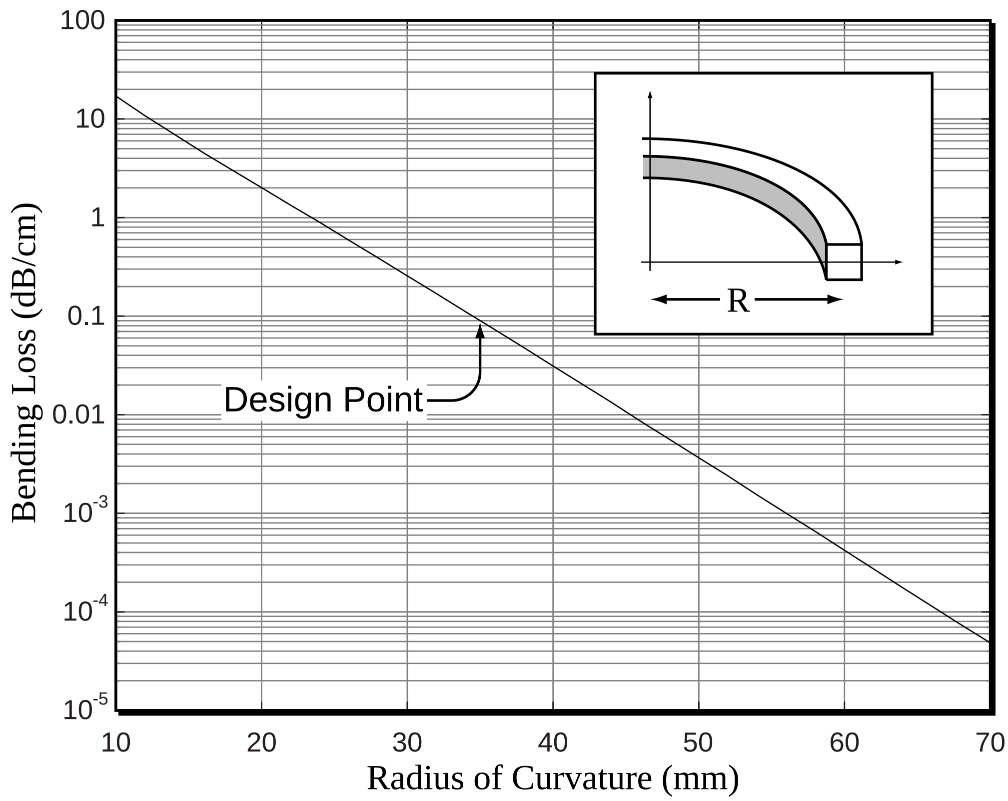


Figure 6.18 Bending loss for channel waveguide as a function of radius of curvature. These results were obtained using the method described in Section 6.3.3.

$$s_{total} = 2R\theta + \left(\Delta d - 2R(1-\cos\theta)\right)\csc\theta \tag{6.31}$$

and the total length of the bent sections only is given by:

$$s_{bend} = 2R\theta \tag{6.32}$$

In order to optimize the structure, we must minimize the quantity:

$$\frac{s_{total}}{z_m} + \frac{s_{bend}}{z_a} \tag{6.33}$$

with respect to the angle θ. Substituting equations 6.31 and 6.32 into equation 6.33, differentiating with respect to θ, and setting the result equal to zero yields (after some algebraic manipulation):

$$2R\left(1+\frac{z_a}{z_m}\right)\cos^2\theta - (2R-\Delta d)\cos\theta - 2R\frac{z_a}{z_m} = 0 \tag{6.34}$$

The above equation can be seen to be a quadratic equation for the quantity $\cos(\theta)$. The solution to this equation is:

$$\cos\theta = \frac{1}{2\left(1+\frac{z_a}{z_m}\right)}\left\{1-\frac{\Delta d}{2R}+\left[\left(1+2\frac{z_a}{z_m}\right)^2-\frac{\Delta d}{R}+\frac{\Delta d^2}{4R^2}\right]^{\frac{1}{2}}\right\} \tag{6.35}$$

where one of the solutions to the quadratic equation has been discarded because it yields a root larger than 1. When the separation Δd is much smaller than the radius of curvature, it is valid to approximate

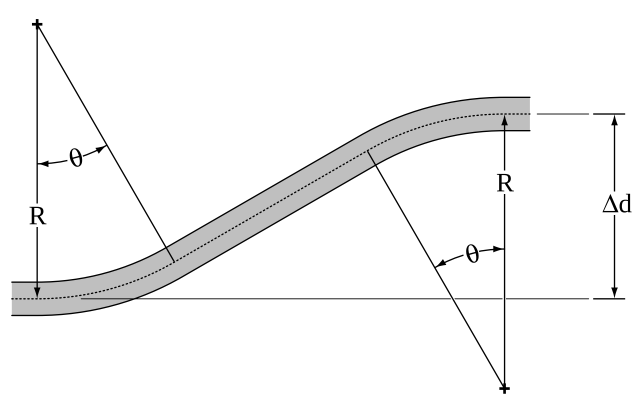


Figure 6.19 S-Bend waveguide segment. Such a bend segment is a building block for directional couplers and related devices. The bend is characterized by a total vertical displacement of Δd, a radius of curvature R and a bend angle θ.

* By attenuation length, we mean the path length over which the optical power decreases by a factor of 1/e. Thus, for an attenuation length z_0, the power decreases exponentially as $P(z) = P(0)\exp(-z/z_0)$.

the solution in equation 6.35 by Taylor expanding to lowest order in terms of $\Delta d/R$. This gives the following simplified result:

$$\cos\theta \approx 1 - \frac{\theta^2}{2} = 1 - \frac{\Delta d}{2R}\frac{1}{\left(1 + 2\frac{z_a}{z_m}\right)} + O\left(\frac{\Delta d}{R}\right)^2 \tag{6.36}$$

The optimal bend angle is then given approximately by:

$$\theta_{opt} \approx \sqrt{\frac{\Delta d}{R}\left(1 + 2\frac{z_a}{z_m}\right)^{-1}} \tag{6.37}$$

Therefore, given a radius of curvature R and vertical displacement Δd, there is an optimal value of the bend angle θ_{opt} which yields the smallest possible bending loss for the structure, accounting for both intrinsic waveguide loss and bending loss.

6.4.2 Full Optimization of Directional Coupler

In order to achieve a specified vertical displacement using an S-shaped bend with a given radius of curvature, the bend angle may be chosen optimally to yield the minimum attenuation, as described in the preceding section. However, this analysis does not describe how the radius of curvature R should be selected in the directional coupler. Recall from Chapter 5 that in a waveguide coupler the total integrated coupling is related to the radius of curvature of the waveguides as they approach one another in the coupling region. Thus, the problem of optimizing the complete directional coupler structure is complicated by the additional constraint that the dimensions of the device must yield the correct amount of integrated coupling. However, the relationships described in Section 6.4.1 for calculating the optimal bend angle, along with the simple approximations used in Chapter 5 for estimating the total integrated coupling make it possible to quickly and easily calculate device structures and attenuation figures for a range of radii and waveguide separations.

Figure 6.20 illustrates the geometry of the directional coupler which we seek to analyze. All of the waveguide bends are constrained to have the same radius of curvature. d_1 is the waveguide separation at the input plane of the coupler and d_2 is the waveguide separation at the output plane. The waveguides reach a minimum center-to-center separation of d_0, and remain at that separation for a distance L. The bend angle on the input side of the coupler is denoted θ_1 and the bend angle on the output side of the coupler is denoted θ_2. The waveguide separations at the input and output planes are generally predetermined by the separation of the input fibers or by the requirement that the coupling between the two guides be negligible. The only free design parameters are the radius of curvature R and the minimum waveguide separation d_0. Optimization of the directional coupler is accomplished with the following procedure.

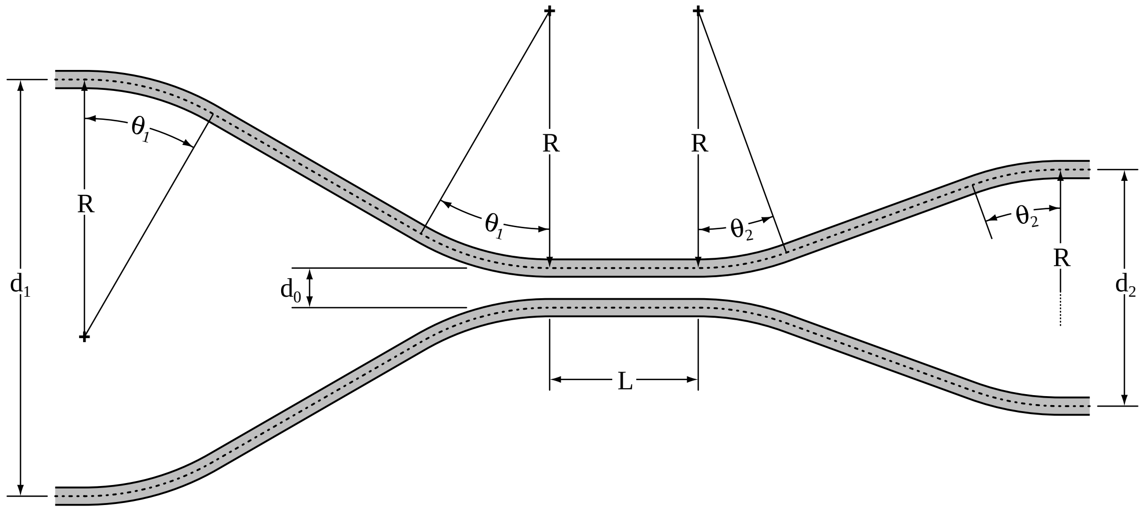


Figure 6.20 Geometry of directional coupler.

First, we characterize the waveguide coupling as described in Chapter 5: the coupling coefficient μ is calculated as a function of waveguide separation, and the resulting data is fitted to an exponential curve. Second, we calculate the bending loss of the waveguide, using the model described in Section 6.3.3, to determine the bending loss as a function of radius for the waveguide under consideration. Third, we select a range of radii (R) and minimum separations (d_0) which we will consider, forming a two dimensional matrix of values. For each value of R and d_0, we calculate all of the properties of the system. Based upon the minimum separation and the radius of curvature, we select the length of the parallel waveguide segment L using equation 5.47 so that the total integrated coupling of the device is $\pi/4$. We then optimally select the angles θ_1 and θ_2 as described in Section 6.4.1. Fourth, having determined all of the geometrical properties of the coupler, it is possible to calculate the total device length and total device attenuation. We repeat the procedure for all of the (R, d_0) pairs, and plot the results to determine which values of R and d_0 give the optimal design. Note that all of the calculations involved in this optimization involve simple algebraic relationships, making it mathematically simple and numerically efficient to consider a large number of data points.

Figure 6.21 plots the results of these calculations for the channel waveguide under consideration. The waveguide separation at the input plane of the device is chosen to be 250 μm and the separation at the output plane is chosen to be 75 μm. The intrinsic waveguide loss used in the calculations was 0.1 dB/cm, a value which is based upon results reported in the literature for waveguides with similar size and composition [41]. The data presented in figure 6.18 were used to describe the bending loss as a function of radius. The data presented in Chapter 5 (plotted in figures 5.12 and 5.13) describe the waveguide coupling as a function of waveguide separation. Notice in figure 6.21 that when the waveguides are placed closer together, the coupling between the guides is stronger and therefore the length L is shorter, which tends to reduce the total device length and attenuation. Interestingly, there is an optimum value for the

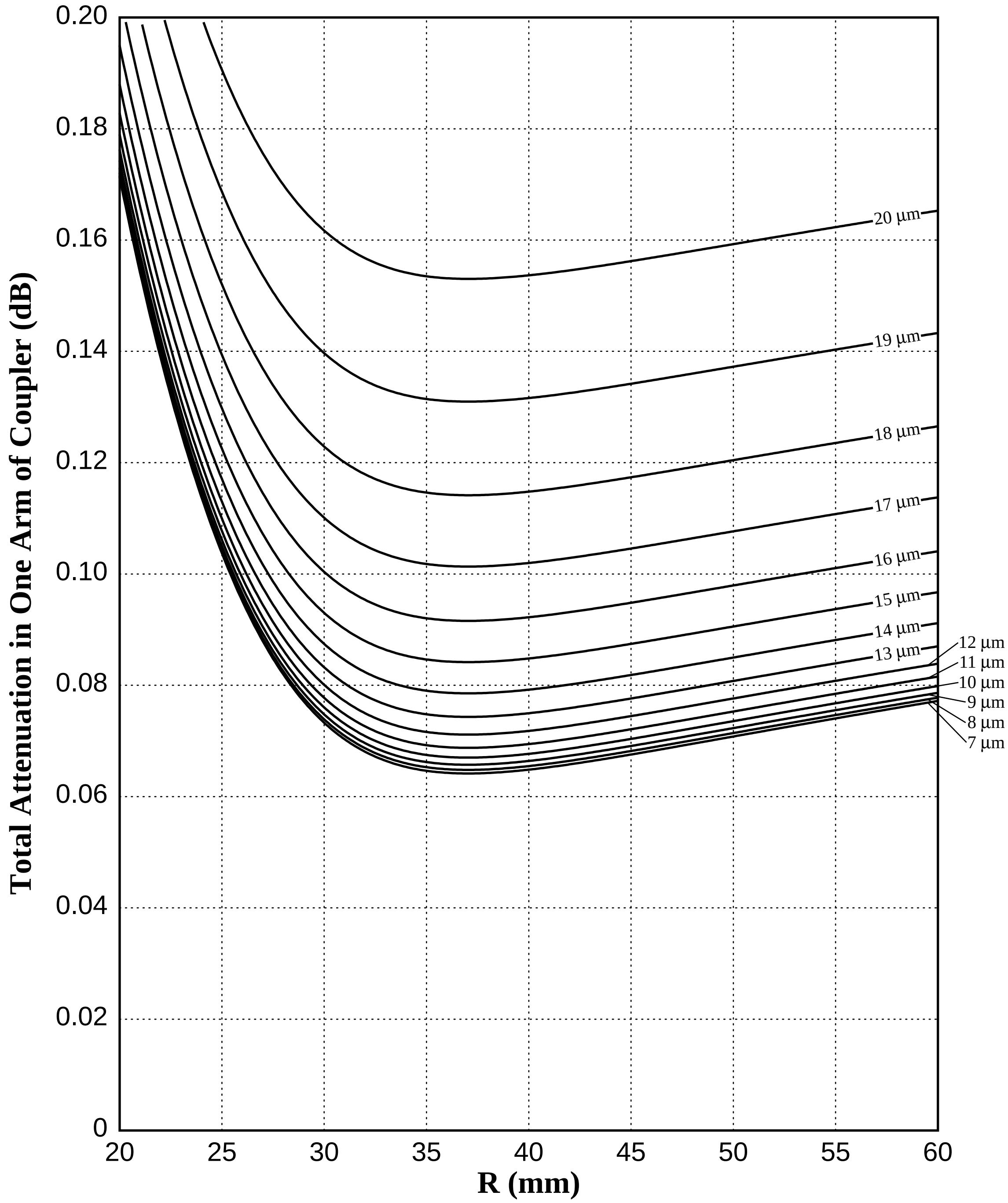


Figure 6.21 Total attenuation in a single arm of a codirectional coupler of the type depicted in figure 6.20. Each curve in this plot corresponds to a different value of d_0, the minimum waveguide separation. (d_0 ranges from 7 μm up to 20 μm.) For all values of d_0, the minimum total attenuation occurs for a radius of curvature of about 37 mm.

radius of curvature which gives the lowest total waveguide attenuation; that radius is around 37 mm. Figure 6.22 plots the total device length for a range of radii and minimum separations. The minimum device length occurs for a radius of curvature of about 30 mm.

Minimizing the total loss of the device is a reasonable goal for designing a directional coupler. However, there are other considerations which are involved in designing an "optimal" device, such as total material cost (which is related to the chip area consumed) and robustness of device performance. An optimal design must weigh all of these considerations in selecting the design parameters. Based upon the analysis described in this chapter, we have selected a 35 mm radius of curvature and a minimum waveguide separation of 10.5 μm. The justification for this choice of waveguide separation is given in Chapter 7, which treats the issues of device tolerances. This set of parameters yields a coupler which is about 6 mm long, with a total predicted attenuation of under 0.07 dB.

Several approximations were made in optimizing the directional coupler: the effective index method was used in analyzing the bending loss, the codirectional coupling expression given in equation 5.47 makes several assumptions about the nature of the waveguide coupling, and moreover we have approximated the intrinsic waveguide loss as 0.1 dB/cm, a value which is at present only a reasonable guess based upon values reported in the literature. In spite of these limitations, the optimization procedure described above provides a rational basis for choosing reasonable device parameters, which is certainly an improvement over arbitrarily selecting all of the device parameters. Even if the insertion loss figures predicted by this analysis prove to be somewhat inaccurate, it is important when designing such devices to at least make an order of magnitude estimate of the device loss, if only to insure that the loss is not prohibitively high.

6.5 Summary

This chapter has dealt with the problem of designing a device having a low total insertion loss. One place where loss occurs is at the junction between two dissimilar waveguides. In Section 6.2, we analyzed the problem of junction coupling between an integrated channel waveguide and an optical fiber. By appropriately choosing the waveguide cross section and index contrast, we can achieve efficient coupling between the fiber and the waveguide.

Another source of loss is bending loss, which is present in any structure that has non-straight waveguides, such as the directional couplers described in Chapter 5. Section 6.3 describes a one-dimensional model for calculating the radiation loss from a bent slab waveguide. The effective index method was employed to reduce the channel waveguide to an effective one dimensional waveguide. Then, using the one-dimensional model, the bending loss was calculated for the glass channel waveguide for a range of bend radii.

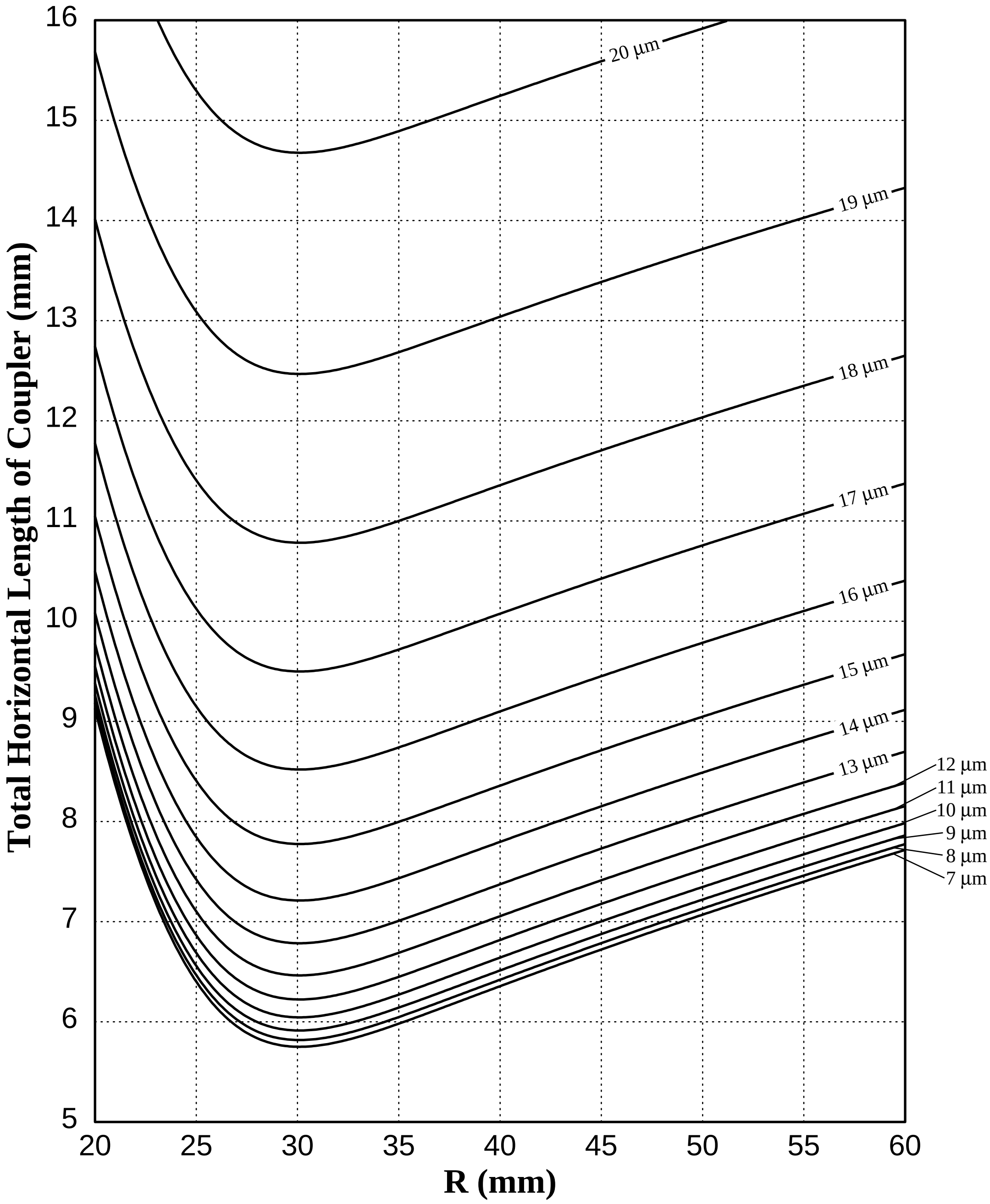


Figure 6.22 The total horizontal length of coupler, plotted for several values of R and d_0. Each curve in the above plot corresponds to a different value of d_0, ranging from 7 μm up to 20 μm.

Section 6.4 combines the bending loss analysis of this chapter with the directional coupler analysis given in Chapter 5 to arrive at some practical design specifications for low-loss couplers. The radius of curvature and bend angles were selected to minimize the amount of loss incurred in the directional coupler.

There is one final source of “loss” that was not discussed in this chapter, but was introduced in Chapter 3. A weak integrated Bragg grating reflects only a portion of the incident light, while the remainder of the light is transmitted through the grating. As described in Chapter 3, the spectral response of the weak Bragg grating becomes more ideal as the product κL is reduced, but this improvement comes at the expense of signal attenuation. In principle, we would therefore like to make the Bragg grating as weak as possible without otherwise hindering the device performance by a weak reflection signal. The loss optimization techniques described in this chapter therefore enable one to improve the filter performance. If one can tolerate only a certain amount of total insertion loss, then minimizing the loss in all other portions of the device allows one to build a weaker Bragg grating, which is predicted to have an improved spectral response.

CHAPTER 7: DEVICE TOLERANCES AND FABRICATION RESULTS

This chapter is concerned with some of the more practical aspects of device fabrication and device tolerances. The work presented earlier in this thesis seeks to analyze prescribed waveguide and grating geometries. If it were possible to accurately fabricate any desired waveguide geometry without any errors or deviations from the design, then perhaps this chapter would be unnecessary. However, almost nothing can be constructed to arbitrary accuracy. It is therefore necessary to consider how variations, deviations and other imperfections of the fabrication process will effect the performance of the device, and to seek a device design that is, wherever possible, insensitive to such deviations. In this chapter, we describe some of the device tolerances, including an analysis of alignment tolerances and waveguide coupler tolerances. We describe a fabrication sequence that addresses these issues, and we present the preliminary results of our initial efforts to fabricate integrated grating structures.

7.1 Generating Gratings with Interferometric Lithography

One of the questions that has not been addressed thus far is how a 1 cm long coherent grating can actually be constructed. By coherent, we mean that if the position of one grating tooth is known accurately, the positions of all other teeth in the grating can be accurately predicted because of the deterministic nature of the grating. As described in Chapter 3, the operation of the Bragg grating is based upon constructive interference between light reflected from different teeth of the integrated grating. In order for this constructive interference to occur, the grating must be coherent over its entire length.

One method of constructing a grating pattern is to use an electron-beam-lithography system, in which a scanning electron beam defines a periodic pattern on a suitable mask. The periodic pattern can then later be transferred from the mask to a substrate using x-ray lithography. The disadvantage of this approach is that the electron-beam system cannot produce a pattern that is coherent over a large area.

The patterns produced by an electron-beam system are subject to distortions caused by charging of the substrate, electromagnetic and acoustic interference, lens aberration, and stage motion. Some of these distortions are systematic and can therefore be corrected (provided they can be measured and characterized.) However, distortions associated with the motion of the stage introduce random errors. Stage motion is necessary in conventional electron-beam-lithography because a single beam cannot be deflected to cover an entire substrate; instead the beam is scanned over small fields and the substrate and stage are stepped to cover the full substrate. Such an approach is subject to field stitching errors at the boundaries between adjacent fields. Field stitching errors result when the patterns from adjacent fields do not match up perfectly at the field boundaries. If a grating pattern spans several electron-beam fields, stitching errors (also known as butting errors) destroy the coherence of a grating pattern. For semiconductor integrated circuit applications, a field stitching error of 30 nm might be acceptable, provided wires are still contiguous across field boundaries. However, for an integrated grating in glass, a displacement of 30 nm corresponds to 3% of the wavelength of light used in optical communications, or equivalently a phase discontinuity of 10°. Such a phase error, when accumulated over many adjacent e-beam fields comprising the grating would destroy the coherence of the structure.

The integrated grating structure must therefore be as coherent as the light with which it is to interact. The technology that we use to build coherent grating structures is interferometric lithography. In conventional interferometric lithography, a standing wave in space is formed by interfering coherent light from an argon ion laser. The standing wave pattern is then used to expose a pattern in a photoresist, either on a substrate or on a mask. Figure 7.1 illustrates schematically the process of interferometric lithography. Two plane waves interfere forming a standing wave in the region where they overlap. The period of the standing wave can be adjusted by changing the angle of incidence of the incident waves. The approach which we use is to first pattern a grating onto an x-ray mask which can then be used to repeatedly expose gratings onto optical substrates. The x-ray mask is a 1 μm-thick SiN_x membrane, supported on a silicon mesa which is bonded to a Pyrex ring [88]. The grating absorber patterns are formed by electroplating gold in the exposed regions of the mask. The process of constructing gratings on x-ray masks is described completely in [89].

Figure 7.2 provides a more detailed schematic of the interferometric lithography system [90]. The spatial filters depicted in Fig. 7.2 are pinhole apertures which can be thought of as point sources of coherent illumination. Therefore, the light that interferes on the substrate is actually not plane waves but rather large radius spherical waves. However, because the area of interest on the substrate is small compared with radial distance from the apertures, the phase-fronts are almost planar.

7.2 Alignment and Coupler Tolerances

One of the assumptions we made when analyzing the Michelson interferometer device in Chapter 5 was that the upper and lower arms of the interferometer are perfectly matched in length. That is,

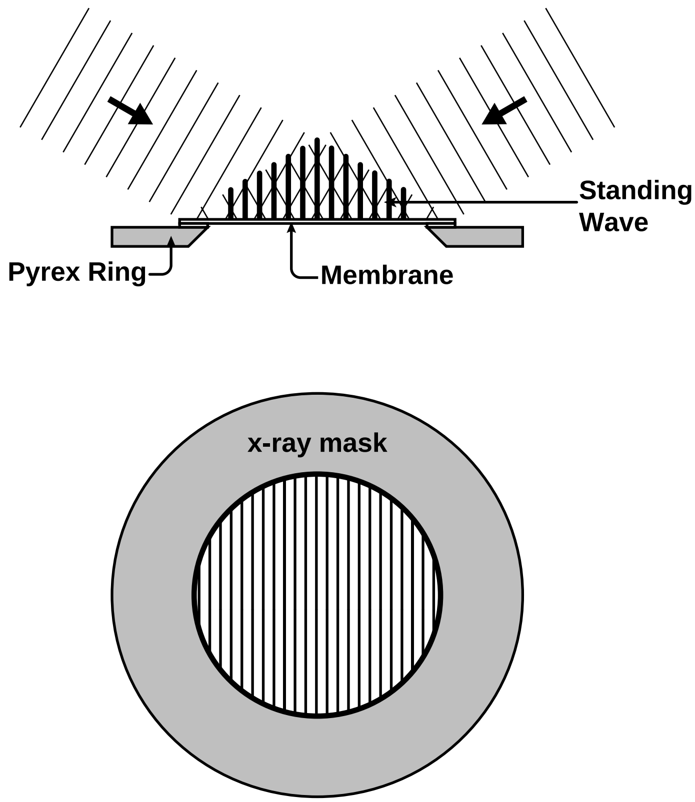


Figure 7.1 Interferometric lithography is used to expose a grating pattern onto an x-ray mask. Plane waves from a coherent source interfere forming a standing wave which exposes a resist on the x-ray mask. The period of the interference pattern can be adjusted by changing the angle of incidence of the plane waves.

the distance between the coupler and the start of the grating is exactly the same in the upper arm and the lower arm. We further assumed that the coupler provided exactly 50% coupling. We now consider how the device performance is effected if the coupling is not exactly 50% or the path lengths in the upper and lower arms are not the same. The analysis of Section 5.6 can be generalized to treat the case where the upper waveguide is longer than the lower waveguide by an amount ΔL. Figure 7.3 illustrates how this configuration can be analyzed using the matrix approach described in Chapter 5. We have inserted a third transfer matrix to represent the unequal waveguide lengths. It has the form:

$$\mathbf{T_D} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^* \end{bmatrix}, \text{ where } \mathbf{D} \equiv \begin{bmatrix} e^{i\beta\Delta L} & 0 \\ 0 & 1 \end{bmatrix} \tag{7.1}$$

Where β is the propagation constant of the waveguide. The transfer matrices $\mathbf{T_1}$ and $\mathbf{T_2}$ describing the coupler and grating respectively are identical to those presented in Section 5.6. The equations for $\mathbf{T_1}$ and $\mathbf{T_2}$ are summarized here for reference:

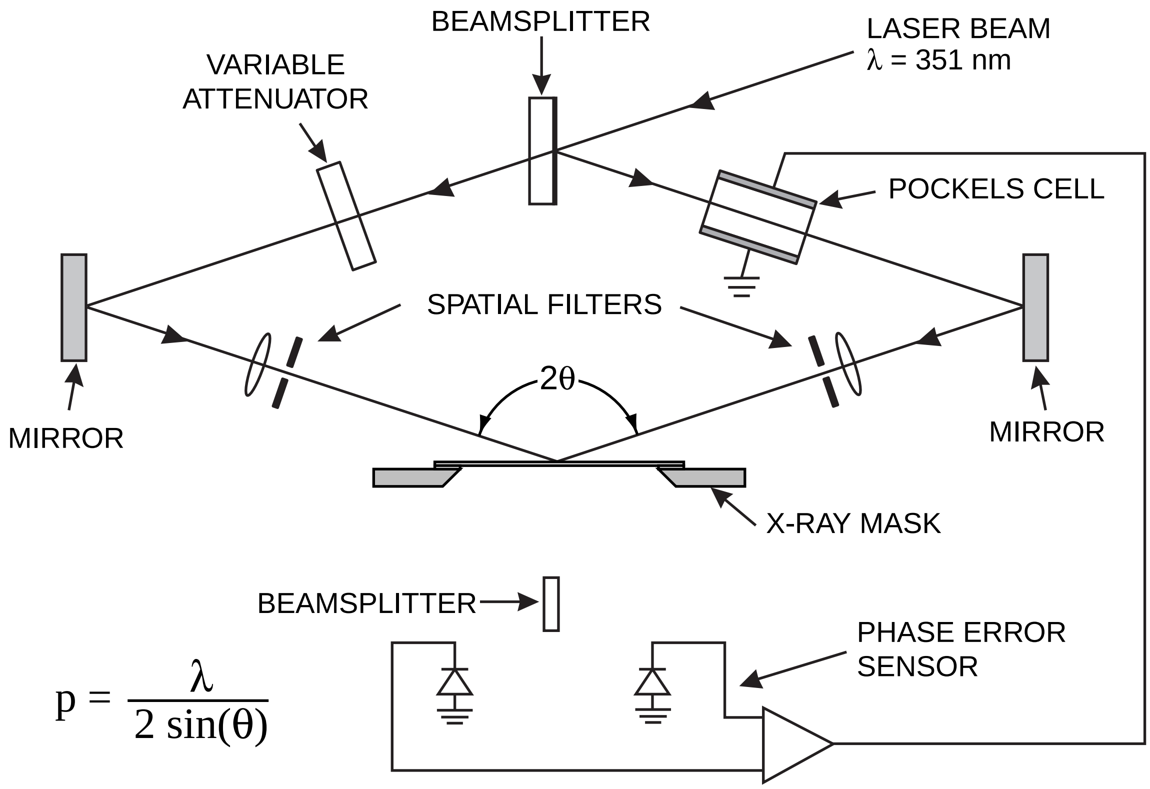


Figure 7.2 Schematic of interferometric lithography system. The spatial filters are pinhole apertures which expand the beam into spherical wavefronts which interfere at the surface. The phase error sensor and feedback loop is used to stabilize the interference pattern by actively adjusting the phase in one arm of the interferometer.

$$\mathbf{T_1} = \begin{bmatrix} e^{+i\beta L_0}\mathbf{R} & \mathbf{0} \\ \mathbf{0} & e^{-i\beta L_0}\mathbf{R}^* \end{bmatrix} \tag{7.2}$$

$$\mathbf{T_2} = \begin{bmatrix} e^{+\frac{i}{2}k_g L_g}\mathbf{I_2} & \mathbf{0} \\ \mathbf{0} & e^{-\frac{i}{2}k_g L_g}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} A\mathbf{I_2} & B\mathbf{I_2} \\ B^*\mathbf{I_2} & A^*\mathbf{I_2} \end{bmatrix} \tag{7.3}$$

Where each of the matrix elements in equations 7.2 and 7.3 are 2 x 2 block matrices. $\mathbf{I_2}$ is the 2 x 2 identity matrix, L_0 is the total path length of the coupler, L_g is the length of the grating segment, and k_g is the grating k-vector. The factors A and B describe the spectral response of the grating, given by:

$$\begin{aligned} A &\equiv \cosh(\gamma L_g) + i\frac{\delta}{\gamma}\sinh(\gamma L_g) \\ B &\equiv \frac{\kappa}{\gamma}\sinh(\gamma L_g) \end{aligned} \tag{7.4}$$

And the 2 x 2 matrix $\mathbf{R}$ in equation 7.2 is given by:

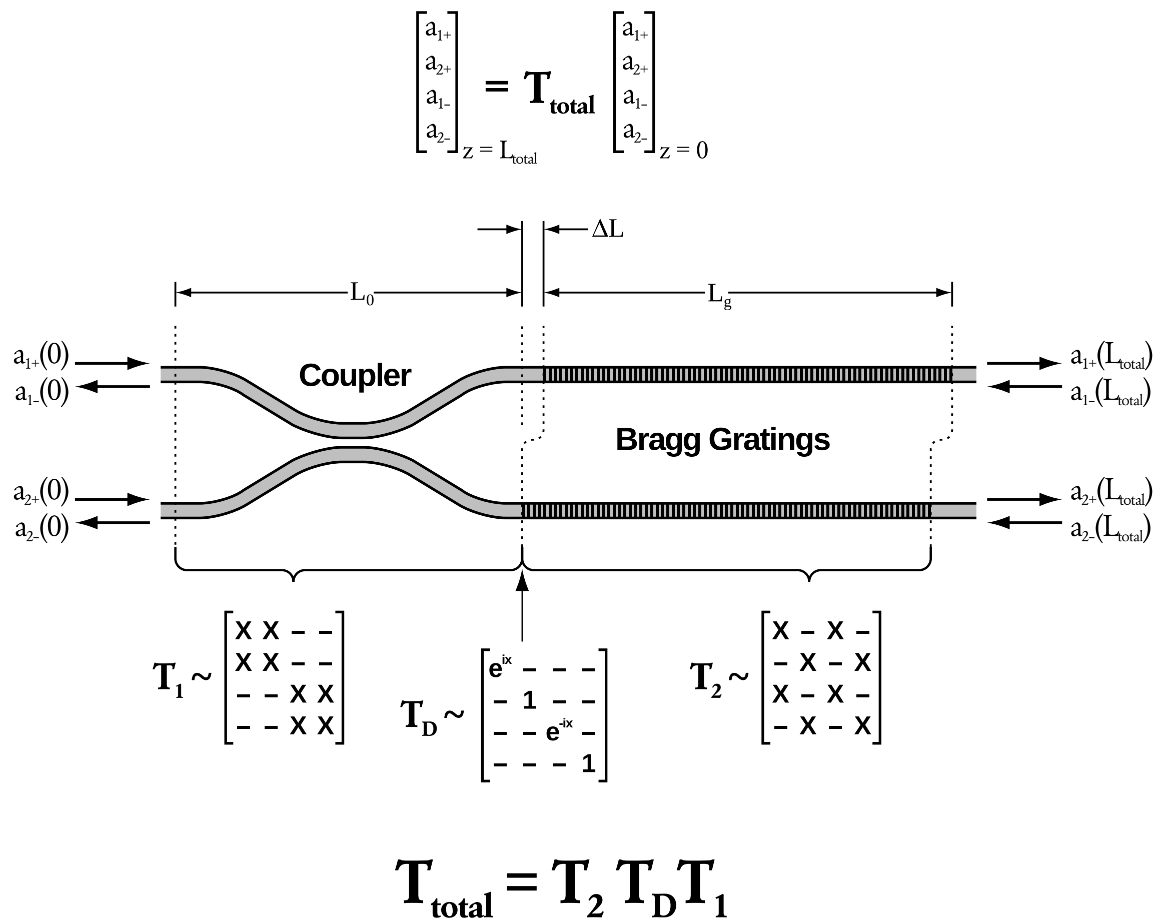


Figure 7.3 A Michelson interferometer device in which the upper and lower arms are not balanced. The imbalance of the interferometer arms can be treated by inserting a third diagonal transfer matrix between the coupler and the gratings, as depicted above.

$$\mathbf{R} \equiv \begin{bmatrix} \cos\phi & i\sin\phi \\ i\sin\phi & \cos\phi \end{bmatrix}, \text{ where } \phi = \int \mu(z)dz \tag{7.5}$$

The response of the entire device is then found by multiplying the three matrices in reverse order, as demonstrated earlier in Section 5.6.

$$\mathbf{T_{Total}} = \mathbf{T_2 T_D T_1} = \begin{bmatrix} e^{+\frac{i}{2}k_g L_g}\mathbf{I_2} & \mathbf{0} \\ \mathbf{0} & e^{-\frac{i}{2}k_g L_g}\mathbf{I_2} \end{bmatrix} \begin{bmatrix} A\mathbf{DR} & B\mathbf{D}^*\mathbf{R}^* \\ B^*\mathbf{DR} & A^*\mathbf{D}^*\mathbf{R}^* \end{bmatrix} \begin{bmatrix} e^{+i\beta L_0}\mathbf{I_2} & \mathbf{0} \\ \mathbf{0} & e^{-i\beta L_0}\mathbf{I_2} \end{bmatrix} \tag{7.6}$$

Equation 7.6 can be seen to be identical to equation 5.57, provided we make the following substitution:

$$\mathbf{R} \to \mathbf{DR} \tag{7.7}$$

Therefore, we can dispense with a redundant analysis of the generalized problem, and make use of the previously derived solution given in equation 5.65:

$$\mathbf{y_0} = -\,e^{+i2\beta L_0}\left(\frac{B}{A}\right)^*\left(\mathbf{D}^*\,\mathbf{R}^*\right)^{-1}\mathbf{D}\,\mathbf{R}\,\mathbf{x_0} = -\,e^{+i2\beta L_0}\left(\frac{B}{A}\right)^*\mathbf{R}\mathbf{D}^2\mathbf{R}\,\mathbf{x_0} \tag{7.8}$$

where the vector $\mathbf{x_0}$ is a 2-vector representing the forward travelling inputs in guides 1 and 2 at the left edge of the device and $\mathbf{y_0}$ represents the oppositely travelling outputs in guides 1 and 2. Substituting the matrices **R** and **D** into equation 7.8 yields:

$$\begin{bmatrix} a_{1-}(0) \\ a_{2-}(0) \end{bmatrix} = -\left(\frac{B}{A}\right)^* e^{+i(2\beta L+\alpha)}\begin{bmatrix} i\sin\alpha + \cos\alpha\cos(2\phi) & i\cos\alpha\sin(2\phi) \\ i\cos\alpha\sin(2\phi) & i\sin\alpha + \cos\alpha\cos(2\phi) \end{bmatrix}\begin{bmatrix} a_{1+}(0) \\ a_{2+}(0) \end{bmatrix} \tag{7.9}$$

where the angle α is defined by:

$$\alpha \equiv \beta\Delta L \tag{7.10}$$

We now apply the initial condition that light is launched into the upper waveguide, while no light is launched into the lower waveguide. The relative power (at the Bragg frequency) reflected back into the upper waveguide is therefore:

$$\left|\frac{a_{1-}(0)}{a_{1+}(0)}\right|^2 = \tanh^2(\kappa L_g)\left[\sin^2(\alpha) + \cos^2(\alpha)\cos^2(2\phi)\right] \tag{7.11}$$

Equation 7.11 reveals that there is no way for the device to completely extinguish the reflected signal in the input port unless both $\sin(\beta\Delta L) = 0$ *and* $\cos(2\phi) = 0$. The amount of reflection in the input port of the device depends upon how closely matched the two arms of the interferometer are, as well as how close to $\pi/4$ the total integrated coupling is. If the difference in path length between the arms is precisely an integral number of half wavelengths, then $\alpha = n\pi$, and the device behaves as if the arms were matched. This result should come as no surprise however, because a path difference of one half wavelength is equivalent to a path difference of one full grating period. Therefore a path difference of one half wavelength can be visualized as offsetting one of the gratings by exactly one grating period relative to the other such that the first tooth of the upper grating lines up with the second tooth of the lower grating, etc. Because the reflection from a single tooth of the grating is assumed small, we expect that this result should yield the same response as if the path lengths were matched between the arms. Therefore, the requirement that the arms of the interferometer be "matched" can be better restated as the condition that the grating lines must be perfectly perpendicular to the direction of the waveguides. If the grating is properly aligned in angle to the waveguide, the matching condition will automatically be satisfied. (This assumes that the lower waveguide is a mirror reflection of the upper waveguide.)

Figure 7.4 is a contour plot of lines of constant extinction in the input arm of the Michelson interferometer, as a function of the two phase parameters α and 2ϕ, which describe the length difference in the interferometer arms and the total integrated coupling, respectively. For the purposes of this plot, we have neglected the $\tanh^2(\kappa L)$ term in equation 7.11. In practice, the non-unity reflection response of the Bragg grating, combined with any other device attenuation would shift these contours figures outward a few dB. In order for the reflection in the input arm of the Michelson interferometer to be reduced by a factor of at least 30 dB, the arms of the interferometer must be matched to within 20 nm, and the total integrated coupling must be $45° \pm 0.9°$. Reflections in excess of 30 dB in the input arm could be large enough to cause lasing in the optical amplifier. Therefore one goal of our design is to achieve a device with at least 30 dB of extinction in reflection in the input arm of the device. If the reflections in the input arm are too high, an isolator could be used to eliminate the reflection. Nonethe-

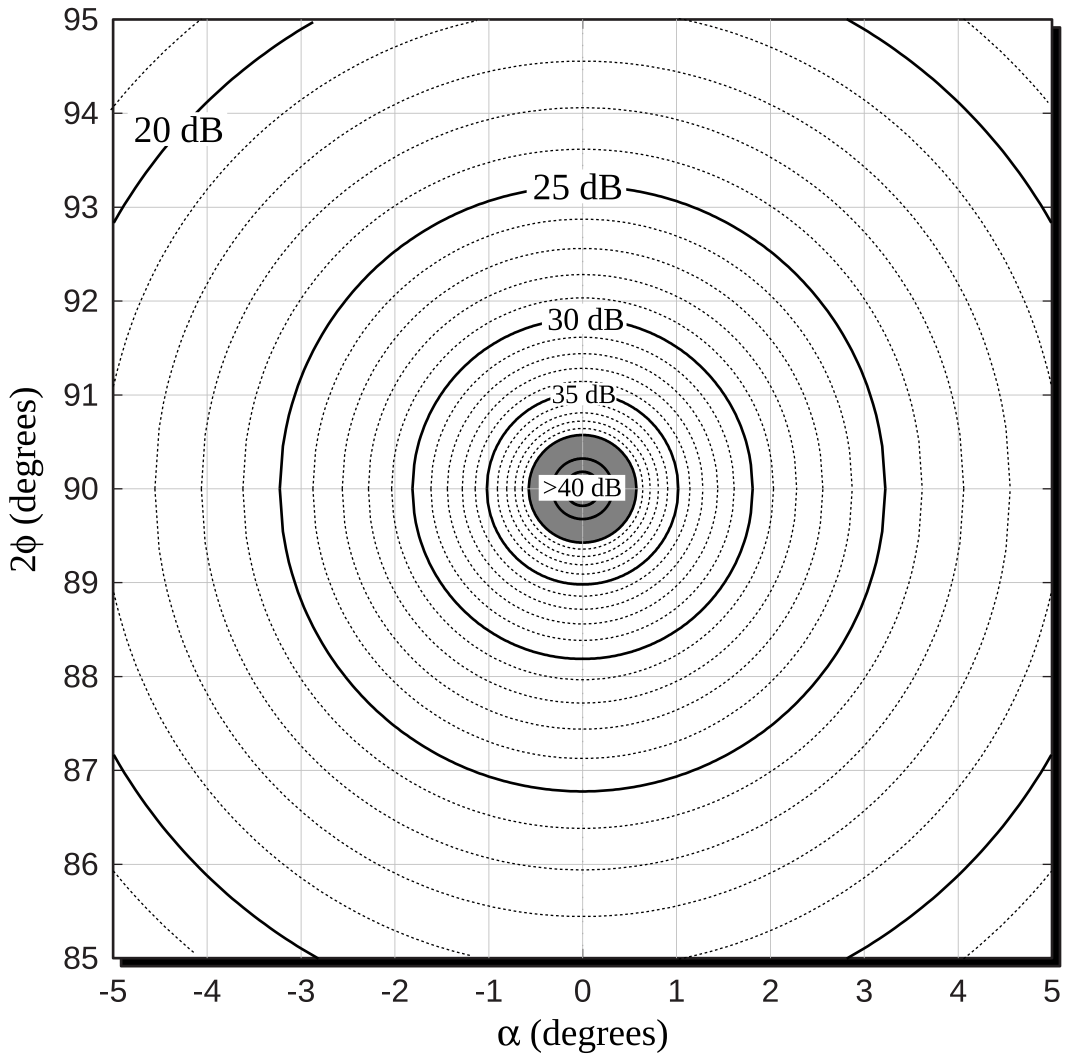


Figure 7.4 A contour plot of the relative return in the input arm of the interferometer, plotted as a function of the two angular parameters α and 2ϕ. α represents the length difference between the upper and lower arms. ϕ represents the total integrated coupling of the coupler.

less, angular alignment is important for insuring that the reflected signal in the output arm of the device is maximized.

Other approaches have been used to balance the arms of the interferometer. Often, a heating element is placed over one arm to allow temperature tuning [91]. Temperature tuning is possible because there is often a slight temperature dependence of the material index of refraction. Another approach is to use a photorefractive material in constructing the waveguides, and tuning one arm of the interferometer by exposing a section of the arm to UV radiation [53, 54]. In photorefractive materials, UV irradiation causes a permanent change in the material index of refraction. Germanium doped SiO_2 (which we are using as the core material for these devices) is photorefractive [92, 93], and this effect could allow us to trim one arm of the interferometer if sufficient angular alignment is not achieved. However, because these devices are planar-based integrated optical devices, it is possible to make use of special alignment technologies when patterning the gratings. Section 7.4 describes our proposed fabrication process, including an alignment scheme that should allow us to balance the arms of the interferometer without using external mechanisms such as temperature tuning or photorefractive trimming.

7.3 Coupler Geometry

Chapter 6 provides an analysis of the codirectional waveguide coupler, with the goal of minimizing the total insertion loss of the device. However, as described in the above section, another important criterion for the directional coupler is that it provide exactly 50% coupling. If the waveguide geometry deviates from the nominal design, because of imperfections in the fabrication process, we would like the device to still provide the correct amount of coupling. If possible, we would like the coupler performance to be relatively insensitive to variations in the waveguide size.

To investigate how a change in waveguide size would effect the total integrated coupling of the device, we have recalculated the coupling constant $\mu(d)$ for 9 different waveguide geometries. The nine cases considered comprise a 3 x 3 matrix of waveguide widths and heights, above and below the nominal design of 6.6 μm square. Figure 7.5 illustrates graphically the range of waveguide sizes considered. These nine cases cover values of d_x and d_y which are within 0.5 μm of the nominal design values. For each of the nine cases, the coupling constant was calculated as a function of waveguide center-to-center separation, using non-orthogonal coupled mode theory and a numerical modesolver, as described in Section 5.5. For each of the nine cases, a least-squares exponential fit was performed on the coupling data. Having characterized the coupling separately for each case, it is possible to apply the analysis described in Chapter 5 to determine how much the total integrated coupling changes as a result of deviations in waveguide size from the central value. The effective coupling length approximation given in equation 5.48 was used to account for coupling within the curved regions where the guides are not parallel. The length of the parallel segment is selected so that for the central case (6.6 μm square waveguide), the total integrated coupling is exactly 45°. Then using the same value for the length L, the total integrated coupling is calculated for the remaining 8 cases. The analysis can readily be carried out for a range of radii R and minimum waveguide separations d_0.

Figure 7.6 plots the deviation in the total integrated coupling (measured in degrees deviation from the nominal value of 90°) as a function of d_0, the minimum waveguide center-to-center separation, for the limiting case where R is taken to be 0. The special case where R = 0 is equivalent to neglecting any extra coupling arising from the curved waveguide segments. As illustrated in Fig. 7.6, when the minimum waveguide separation is large, the device performance tends to be more sensitive to deviations in the waveguide geometry. Interestingly, at a minimum separation of approximately 14 µm, the curves with the same value of d_y but different values of d_x seem to intersect. Likewise, at a minimum separation of about 9.5 µm, the three curves with the same value of d_x but different values of d_y intersect. Therefore, when we neglect the effect of coupling in the nonparallel regions, the coupling constant is almost stationary to variations in d_x at a waveguide separation of 14 µm, and µ(d) is stationary to variations in d_y at a waveguide separation of 9.5 µm.

The data plotted in Fig. 7.6 assumes that all of the coupling arises from one parallel waveguide segment, and it therefore neglects the coupling that results from the regions where the waveguides bend towards and away from each other. When we account for this extra coupling, the results are modified slightly. Figure 7.7 plots the same information as in Fig. 7.6, this time assuming that the waveguides

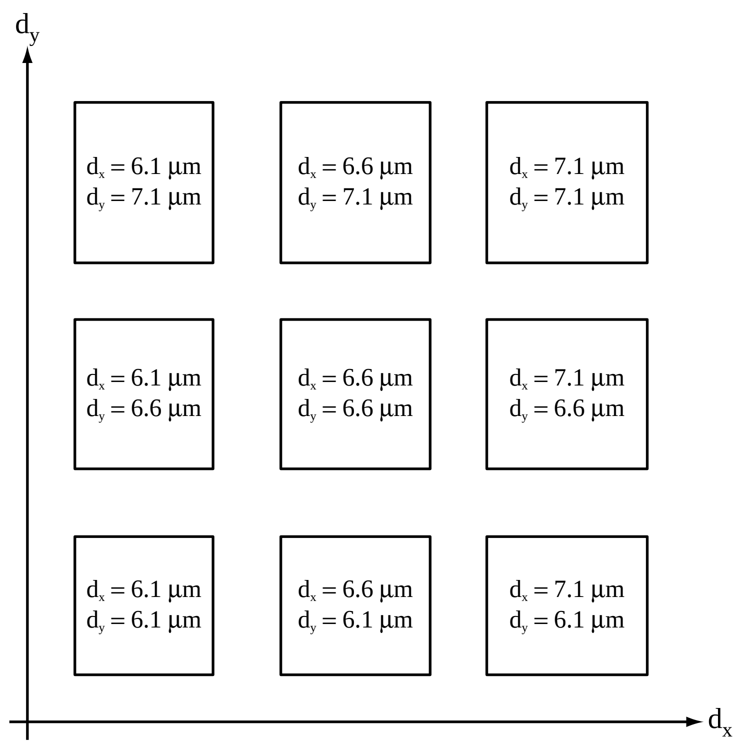


Figure 7.5 Range of waveguide sizes analyzed. In order to study how the waveguide coupling depends upon the geometry of the waveguide, we calculate the coupling for nine different cases, forming a 3 x 3 matrix of waveguide dimensions. The figure above illustrates the range of core sizes considered.

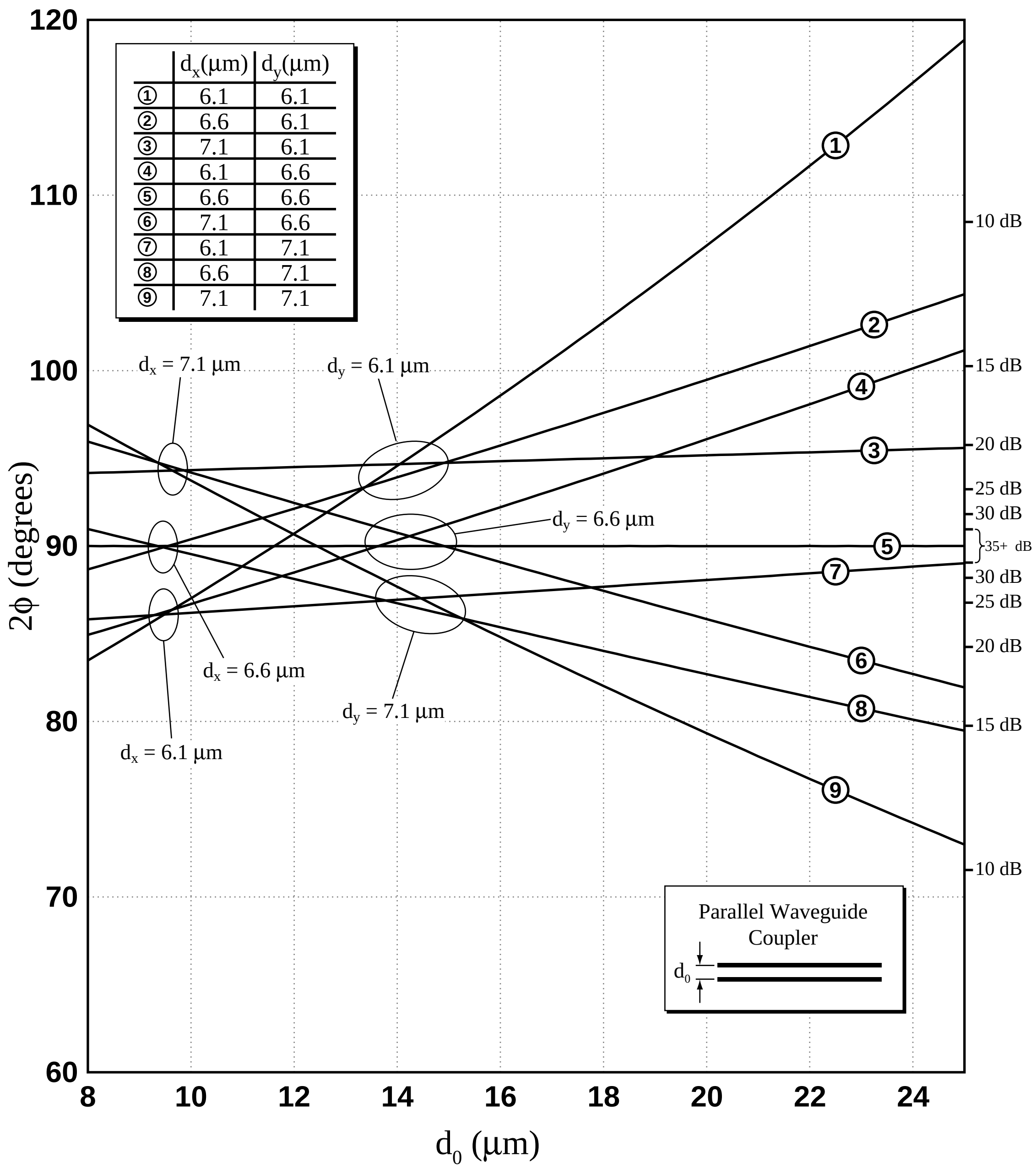


Figure 7.6 Total integrated coupling of parallel waveguide coupler, calculated for 9 different waveguide sizes. The left axis plots the 2ϕ, (twice the total coupling) for the coupler, and the x-axis plots the waveguide center-to-center separation. The right axis labels the amount of relative reflection in the input arm, assuming that the arms are matched in length. Note that when the waveguides are separated by 9.5 microns, the structure is insensitive to variations in d_y, and when the waveguide separation is 14 microns, the structure is insensitive to variations in d_x.

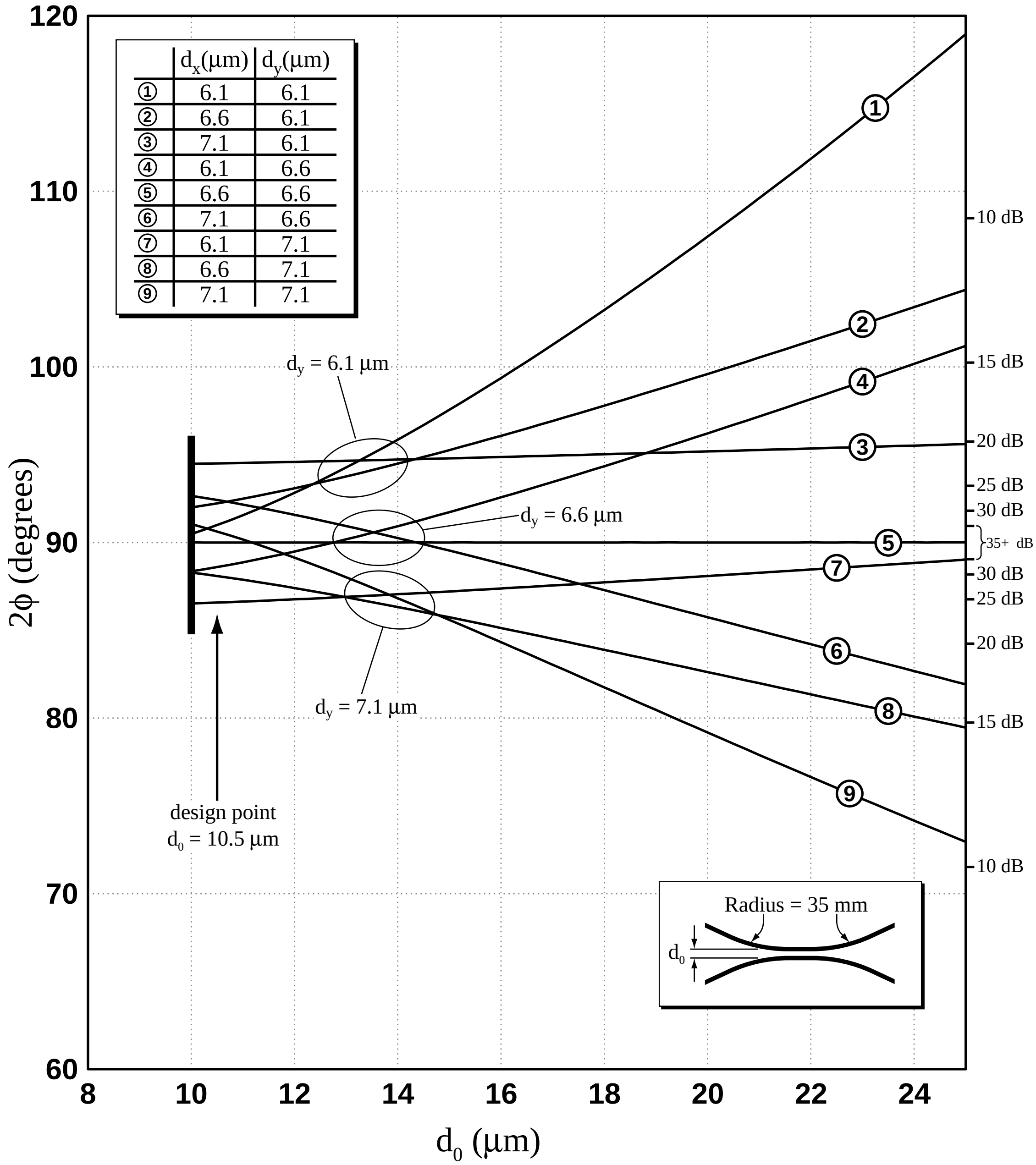


Figure 7.7 Total integrated coupling for nine different waveguide geometries, as a function of the minimum center-to-center separation. This plot is similar to the data plotted in Fig. 7.6, with the exception that the coupling in the curved regions is also included. When the radius of curvature is chosen to be 35 mm, it is not possible to bring the waveguides closer than 10 μm separation.

curve towards and away from each other with a radius of curvature of 35 mm. One important difference between Fig. 7.6 and Fig. 7.7 is that when we account for the coupling in the curved regions, the guides cannot be brought closer than about 10 μm before the total coupling from the curved segments alone exceeds 45°. The coupler still seems to be somewhat insensitive to variations in d_x at a separation around 14 μm, but the stationary point is not as well defined as in Fig. 7.6. Furthermore, because the waveguides cannot be brought closer than 10 μm, it is no longer possible to achieve a design for which the coupling is stationary under variations in d_y. Based upon this data, we have selected a minimum waveguide separation of 10.5 microns, as indicated in Fig. 7.7. For even the most extreme cases considered in these calculations, this design should give a variation in the total integrated coupling of at most a few degrees. To further ensure that at least some of the devices fabricated have the correct amount of coupling, we have included a range of coupling lengths on the photomask, which are calculated to yield a total integrated coupling which ranges from 40° and 50°.

7.4 Fabrication Process

7.4.1 Overview

Generally, the fabrication sequence for building integrated channel waveguides begins with a substrate material that includes a thick lower cladding covered by a planar slab of core material. In our case, the cladding is 20 μm of undoped SiO_2, and the core is a uniform slab of 6.6 μm-thick Ge-doped SiO_2. The amount of Ge doping is selected to achieve an index of refraction difference between the core and cladding of 0.3%. The underlying substrate is a four inch Si wafer. Both optical layers are deposited by a process of flame hydrolysis deposition, supplied by an outside vendor. The planar substrate is processed using the appropriate lithographic tools (described later in this section) to remove the core material everywhere but where the waveguides and gratings are to be. Finally, an upper cladding layer of 20 μm-thick SiO_2 is deposited over the top of the devices. This general sequence is outlined schematically in Figure 7.8. The remainder of this chapter describes the processing steps required to pattern the channel waveguide and gratings.

One of the challenges in constructing these devices is that the very-fine-period shallow Bragg gratings must be patterned and etched onto the top of the relatively tall channel waveguides. Figure 7.9 is a roughly-to-scale diagram illustrating the relative dimensions of the grating and the waveguide core. Sections 7.4.2 – 7.4.3 describe a unique fabrication sequence in which the gratings are patterned *first* onto the substrate, but the grating pattern is not etched until *after* the waveguides have been formed. This "pattern first, etch later" process is well suited for the problem of making fine period gratings on top of relatively tall features, because the nanolithography steps can be performed on flat surfaces.

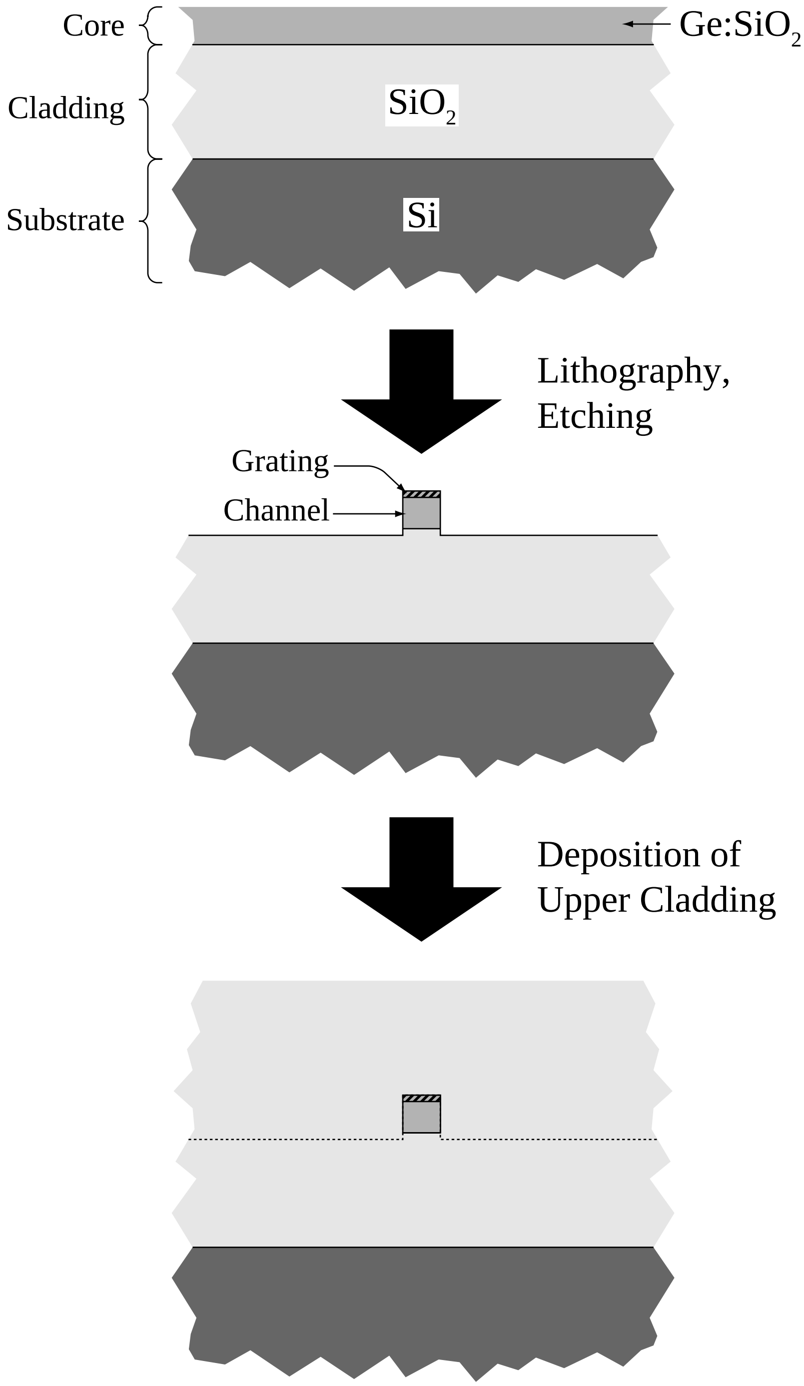


Figure 7.8 Overview of fabrication sequence. A planar slab waveguide is processed and etched to form a rectangular channel waveguide with a grating. A second, thick layer of glass cladding material is then deposited over the structure.

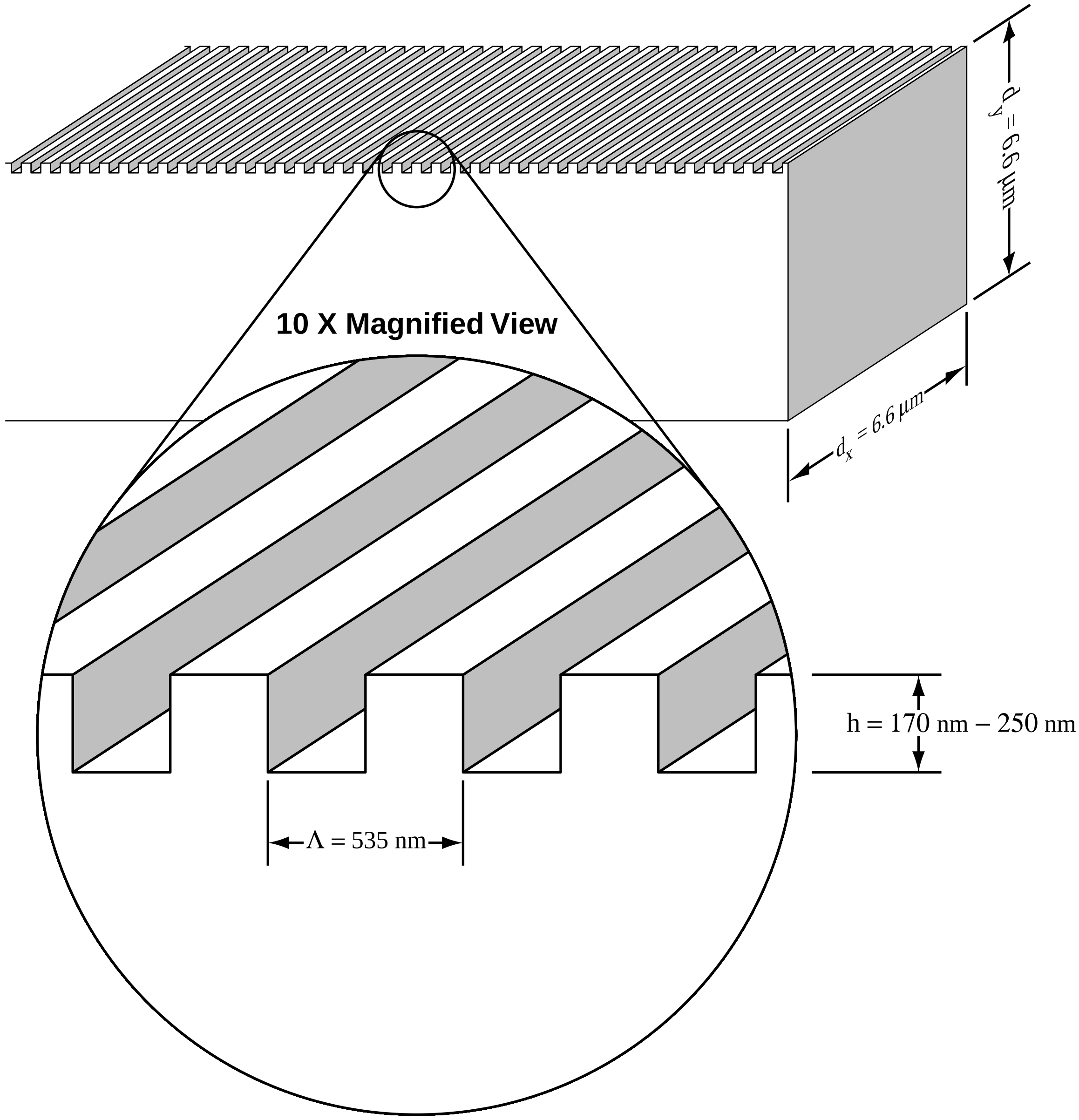


Figure 7.9 Roughly to scale diagram illustrating the relative dimensions of the grating and waveguide core.

7.4.2 Patterning the Grating

Initially, the grating regions are defined on the substrate using deep-UV contact photolithography in poly-methylmethacrylate (PMMA.) The grating regions are rectangular blocks on a conventional quartz photomask which define the regions where the gratings are to be patterned. These rectangular regions are placed such that they overlap with where the waveguides will later be located. However, the grating regions are designed to be 15.6 μm wide (9 μm wider than the waveguides), thus their lateral placement tolerance is ±4.5 μm. After developing the grating window pattern, PMMA remains on the substrate only in the regions where the gratings are to be placed. The remaining PMMA on the substrate then undergoes an x-ray exposure, through a fine-period grating mask, fabricated using the interferometric lithography scheme described in Section 7.1. After developing a second time, the PMMA remaining on the substrate forms a grating pattern which exists only within the rectangular windows defined by the first exposure. Thus, we use a two-step exposure process wherein the first (optical) exposure pre-patterns the resist into large rectangular grating windows and the second (x-ray) exposure defines a grating pattern in the remaining resist. The optical exposure must be developed (at least partially) before making the x-ray exposure because in order to align the x-ray exposure with the optical exposure some pattern features or alignment marks must be visible through the x-ray mask. After this two step expose and develop process, a 30 nm layer of chrome is evaporated and a liftoff is performed. The thin chrome layer is to be later used as a hard etch mask when reactive-ion etching (RIE) the fine-period grating onto the top of the waveguides. Figure 7.10 depicts schematically how the double exposure technique forms a series of grating segments in chrome on the substrate.

PMMA was selected as a suitable resist for this process because it is sensitive to soft x-ray radiation, and it is compatible with the double-exposure method described above. However, it should be pointed out that the process described above could be extended to allow for the use of chemically amplified resists with higher sensitivity. Unlike PMMA, a chemically amplified resist cannot be exposed and developed twice, therefore a sequence of two liftoffs would have to be performed. One liftoff would be done after the optical exposure, after which the sample would be re-coated and exposed with x-ray lithography. Following this step, a second liftoff would define a grating pattern on top of the pre-existing chrome window pattern.

7.4.3 Defining the Waveguides

Figure 7.11 illustrates the sequence of steps used to define the waveguides on top of the chrome grating pattern. After the gratings are patterned in chrome on top of the substrate, the waveguide patterns are defined on top of the grating patterns using optical lithography. First, a 500 nm thick uniform tungsten layer is deposited on top of the chrome grating pattern. The tungsten is to be used as a hard mask layer for etching the deep channel waveguides. Photoresist is then spun over the tungsten layer and the waveguides patterns are exposed in the photoresist, as shown in Fig. 7.11B. The photoresist then acts as a mask for reactive-ion etching the tungsten layer, as shown in Fig. 7.11C. A wet etchant could also be used for removing the tungsten, but because of the thickness of the tungsten layer this process is prone to undercutting. Where the tungsten has been removed, the underlying chrome pattern is exposed. This chrome layer can then be selectively removed with a chromium wet etchant or by RIE. The photoresist is then stripped off, leaving 500 nm-thick tungsten waveguide pattern covering a thin chrome grating pattern, as illustrated in Fig. 7.11D.

The deep channel waveguides are etched with RIE in CHF_3, using the relatively thick tungsten layer as a hard mask (Fig. 7.11E). The etch duration is selected to achieve an etch depth of at least 6.6 μm, which should etch all the way through the core region. After etching the channel waveguides, the tungsten is stripped using a wet chemical etch, which selectively removes the tungsten without effecting the underlying chrome grating pattern (Fig. 7.11F). Following this step, the grating pattern is etched

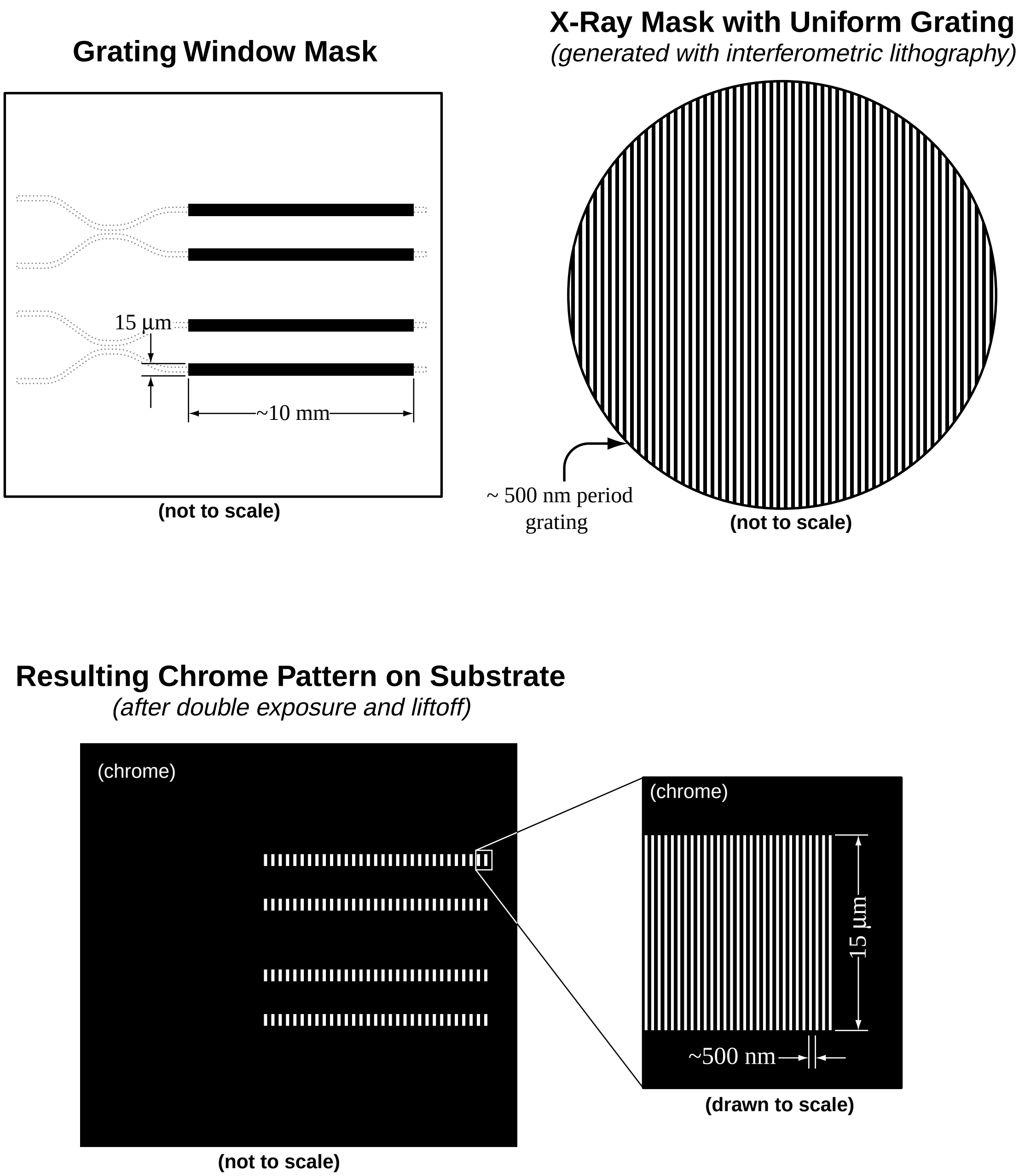
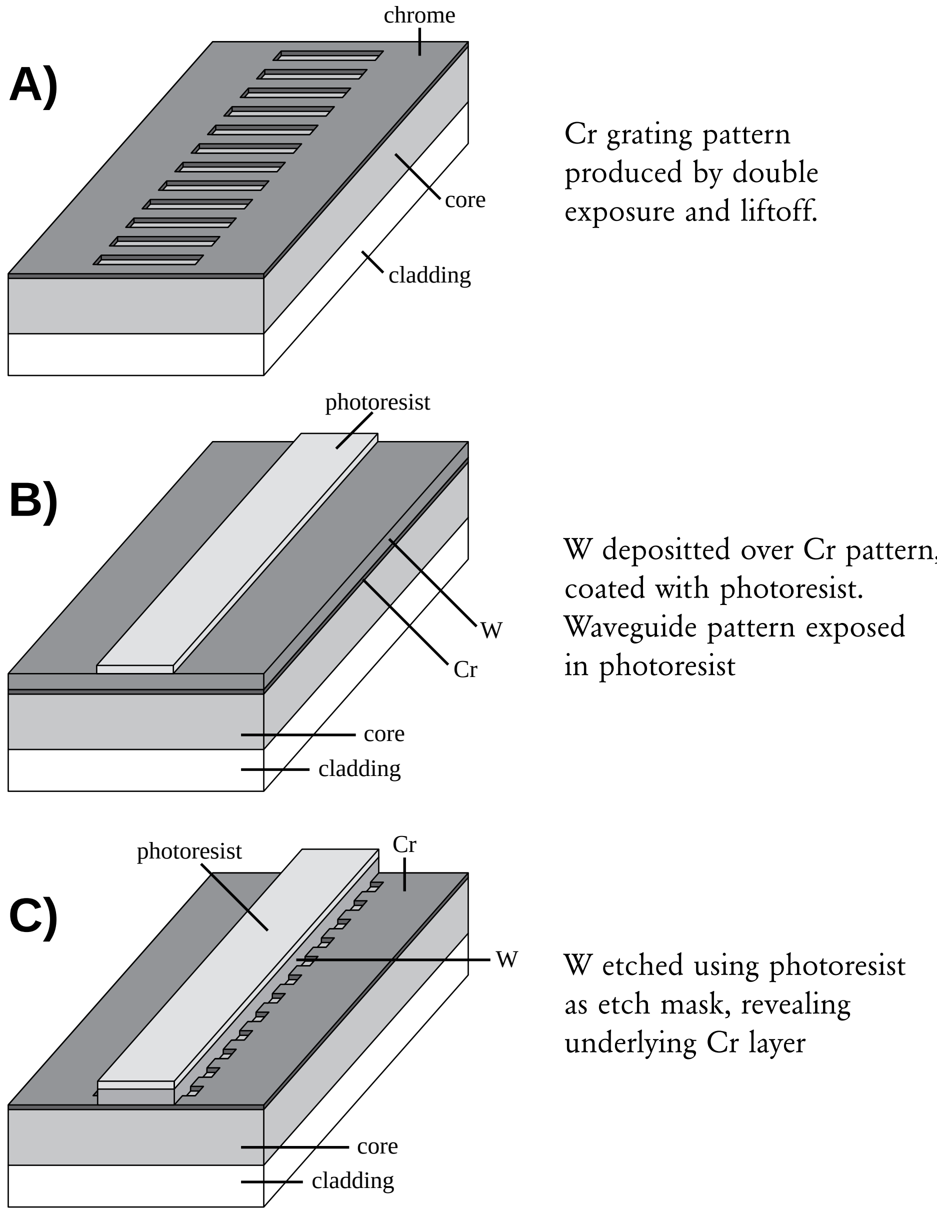


Figure 7.10 A double-exposure is used to define grating regions on the substrate. The first exposure is an optical exposure which defines the rectangular regions where the grating is to reside. The second exposure uses a uniform-grating x-ray mask to pattern the gratings within the rectangular regions.

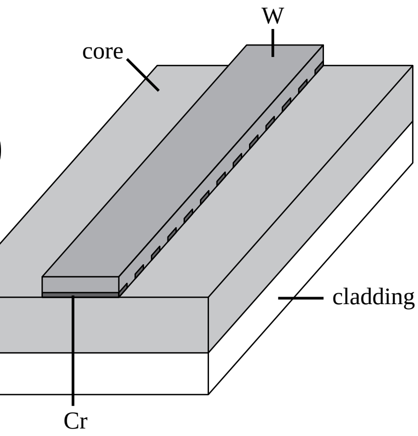


(Figure continued on following page...)

Figure 7.11 Processing sequence for defining waveguides. A second hard mask is patterned on top of the chrome grating pattern. The deep waveguides are etched first using the upper hard mask, after which the second hard mask is stripped revealing the underlying chrome grating pattern. Finally, the gratings are etched using the chrome mask.

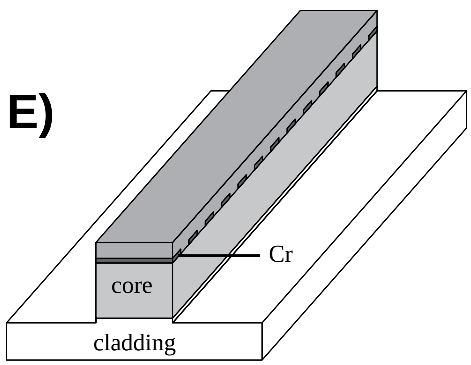
(...continued from previous page)

D)



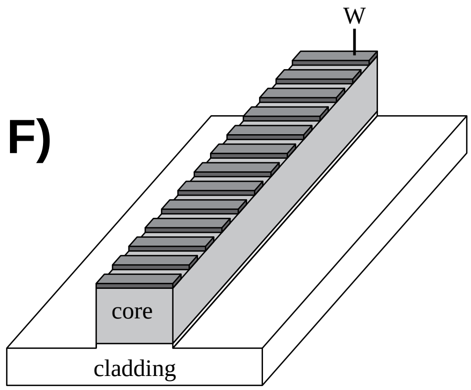
Cr layer chemically etched using W and photoresist as etch mask.

E)



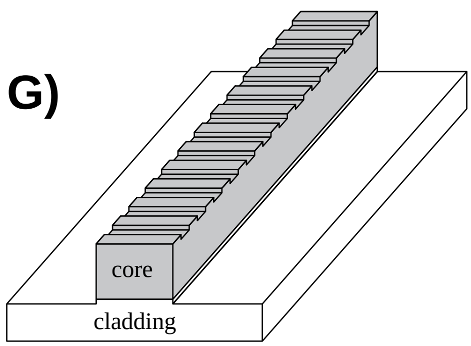
Channel waveguide etched with RIE using W pattern as hard mask.

F)



W selectively removed with wet chemical etch, without effecting underlying Cr pattern

G)



Cr used as hard mask for grating etch (RIE). Cr removed after etch.

onto the top of the core, this time using the thin chrome layer as an etch mask. Again, reactive ion etching is used with CHF_3. Finally the chrome hard mask is stripped leaving a 6.6 μm high channel waveguide with a shallow grating etched into the top surface, as depicted in Fig. 7.11G.

Tungsten and chrome have been selected as suitable metals because these materials can be independently and selectively removed. That is, chromium etchant does not attack the tungsten pattern, and likewise tungsten etchant does not attack the chromium pattern. Other pairs of masking materials might also be suitable for this process. For example, a different metal could be substituted for tungsten, provided the etching chemistry for that metal is harmless to the chrome. It is desirable for the waveguide etch mask (tungsten, in our case) to have a small grain structure so that the edges of the features are relatively smooth. A rough or granular etch mask would create vertical striations in the sidewalls of the waveguide, which could lead to scattering and additional attenuation. Researchers at NTT employ an amorphous silicon hard mask for the waveguide etch, which does not have a granular structure [37]. Further research is required to determine whether amorphous silicon is compatible with a secondary buried etch mask such as chrome.

One advantages of the "pattern first etch later" processing sequence is that all of the lithography steps are performed on surfaces with little or no topography. Both lithography steps are done first before any etching has been done; then the etching steps are performed in the reverse order. Another advantage is that the grating regions can be made wider than the waveguide structure, giving an extra degree of alignment tolerance. The chrome wet etch step ensures that the grating is confined to the top of the channel waveguide.

7.4.4 Alignment

The fabrication process described in the preceding sections involves three masks. The purposes of the three masks is illustrated schematically in Fig. 7.12. The first mask is an optical mask which defines the regions where the gratings are to be patterned. The second mask is an x-ray mask which contains a uniform grating with a period of approximately 500 nm. The third mask is an optical mask which defines the waveguide patterns. When building the devices, it is important that all three lithography steps be aligned to one another. The alignment of the grating windows with the waveguides allows a certain degree of misalignment, because the grating windows are intentionally made wider than the waveguides. The angular alignment of the grating with the waveguides is of critical importance because the arms of the interferometer will be mismatched if the grating is misaligned, as described in Section 7.2. Even if the grating windows overlay perfectly with the waveguide segments, a misalignment of the grating lines amounts to an effective path length difference between the upper and lower arms of the interferometer. Furthermore, if the angular misalignment of the gratings is too severe, some the light reflected by the grating perturbation could be scattered into radiation modes rather than reflected into the backward travelling bound mode. For these reasons, an alignment scheme is required which ensures that the grating lines are precisely perpendicular to the direction of the waveguides.

It is relatively straightforward to include alignment marks on the two optical masks in order to allow for alignment of the grating windows with the waveguides. However, because the grating x-ray mask is fabricated by interferometric lithography, alignment marks cannot be placed on the grating mask without further processing. Moreover, the 500 nm period grating pattern cannot be seen with conventional alignment optics, making it impossible to align the grating pattern with a substrate. To address this problem, we propose to use a modified grating mask, in which alignment marks are added to the mask with electron-beam-lithography, using the process depicted in Figure 7.13. Interferometric lithography is used to pattern a grating on an x-ray mask. However, before the grating pattern is developed, a second optical exposure is performed, exposing two large areas at the top and bottom of the mask pattern. After developing and electroplating the x-ray mask, the areas which were exposed optically will be plated. The x-ray mask is then replicated (or daughtered) onto a second x-ray mask, a process that reverses the polarity of the pattern. The daughter mask then has two transparent regions at the top and bottom of the patterns. It is in these regions that the alignment marks are to be written. The daughter mask is coated with PMMA and loaded into an electron-beam-lithography system. The e-beam system can then be used to examine the grating pattern and write a pair of alignment marks at the top and bottom of the mask which are precisely aligned with the grating on the mask. The PMMA is then developed and electroplated. After the alignment marks are written, the grating mask can be aligned with the other two lithography steps. It should be pointed out that in order for the electron-beam system to write alignment marks which are lined up with the grating, it is necessary to scan the electron beam over

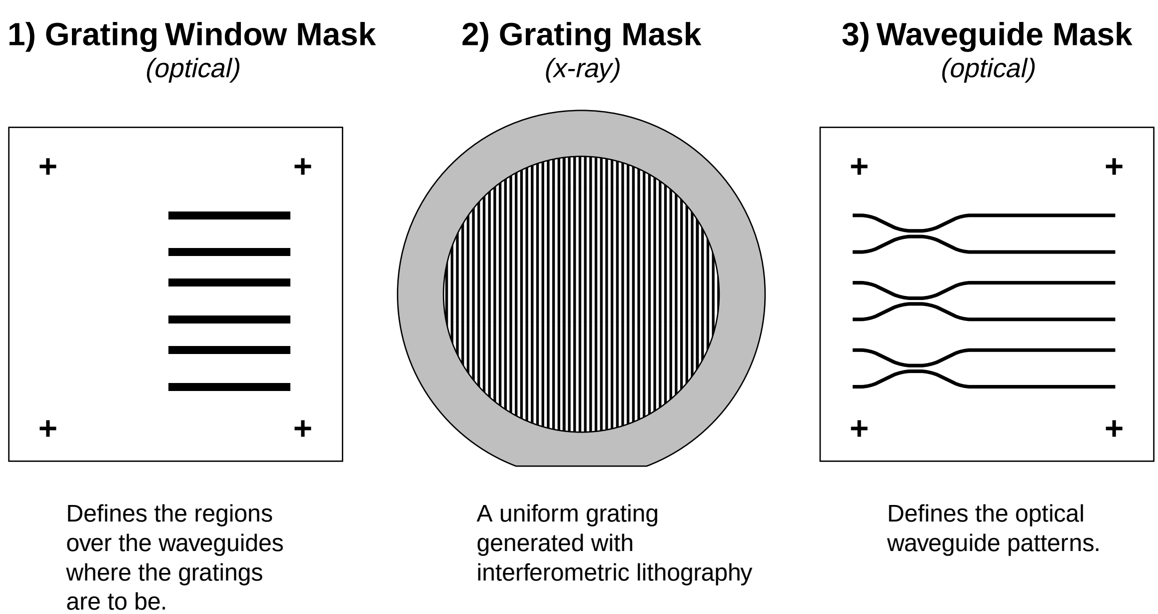


Figure 7.12 Overview of the three masks used to fabricate matched filter devices.

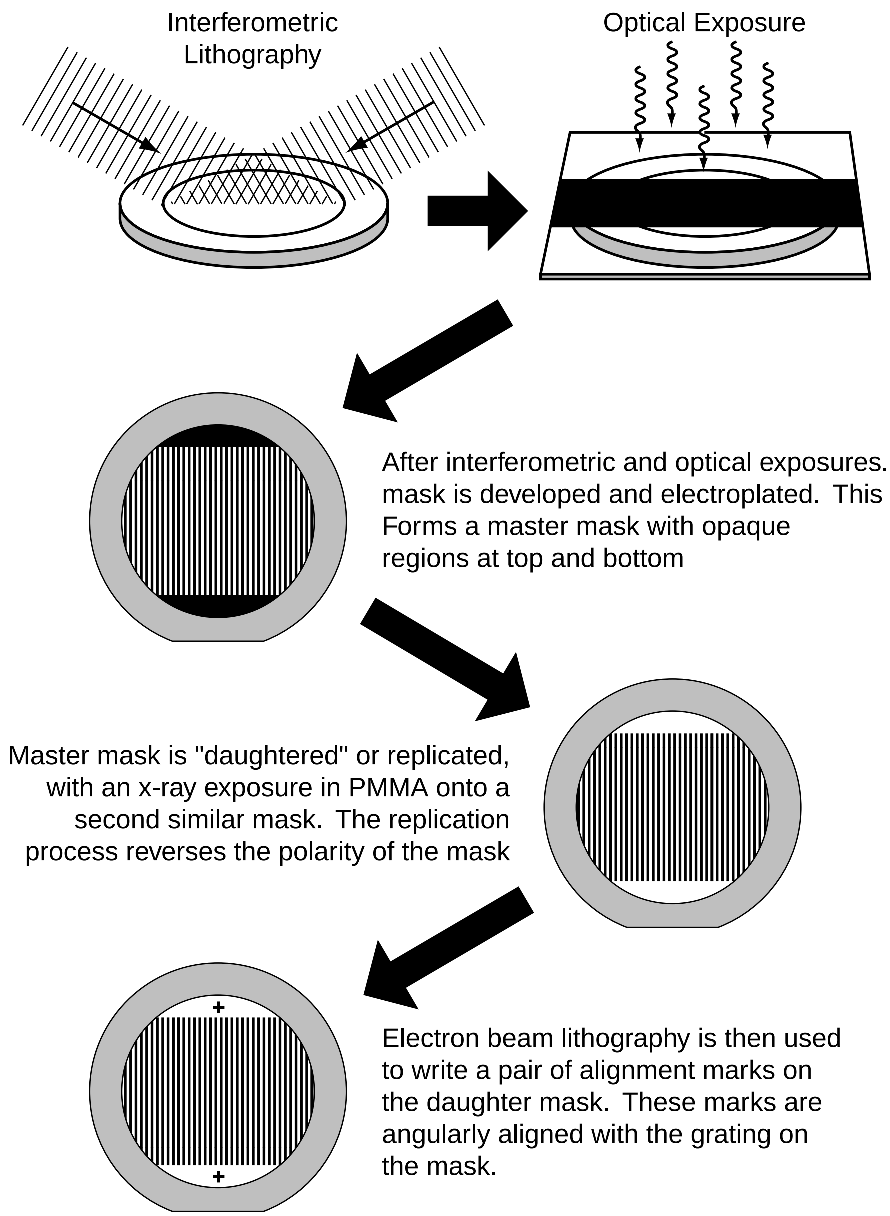


Figure 7.13 Process for adding alignment marks to a grating mask using electron-beam-lithography.

a portion of the gratings, which will cause the PMMA in these regions to be exposed. This is not a problem if the region which is exposed by the e-beam does not overlap with the grating windows defined on the optical mask.

The procedure described above places a pair of conventional alignment marks on the x-ray mask which are angularly aligned to the grating. In principle, it would be possible with the electron-beam system to follow a single grating line from the top of the mask to the bottom, and then write a pair of alignment marks which are aligned with the line of the grating. Later, when the waveguide mask is exposed onto the substrate, it can be aligned with these reference marks, thereby ensuring that the grating lines are perpendicular to the waveguide direction.

7.5 Preliminary Fabrication Results

7.5.1 Optical Mask Design

After deciding on a suitable waveguide coupler geometry and grating length, the next step is to construct an optical mask set which contains the necessary patterns. Two optical masks are needed to construct the matched filter devices. One mask defines the waveguide patterns and the second mask defines the regions where the grating sections are to be located. Both masks require alignment marks, as described in Section 7.4.4.

Chapter 2 describes how the spectral width of the weak grating filter is proportional to the length of the integrated Bragg grating. In order for the filter to be matched, the down-and back propagation time for light in the grating must be equal to one bit duration. In order to span a range of communications bit rates, the lengths of the grating windows on the optical mask were selected to yield down-and-back propagation times of 85, 90, 95, 100, and 105 ps. This corresponds to grating lengths ranging from 8.8 mm up to 10.9 mm.

The various types of waveguide devices included on the mask are summarized in Figure 7.14. The simplest device is a straight waveguide and grating. This device allows the grating filter to be used as a reflection filter, perhaps with a circulator or beamsplitter. A slight variant of this device is the straight waveguide where the right edge of the device is tapered to a point. By tapering the right edge to a point, the light which emerges from the right edge of the grating is gradually radiated rather than directed back to the opposite edge of the chip. Some of the devices have a slight s-shaped bend at the left edge of the chip. This is designed to provide a vertical offset between the input waveguide and the transmission waveguide. These devices will hopefully allow for easier alignment of optical fiber on the right and left edges of the chip: if the fiber is misaligned at the left edge, the uncoupled light from the misaligned fiber is not pointed directly at the output fiber at the right edge.

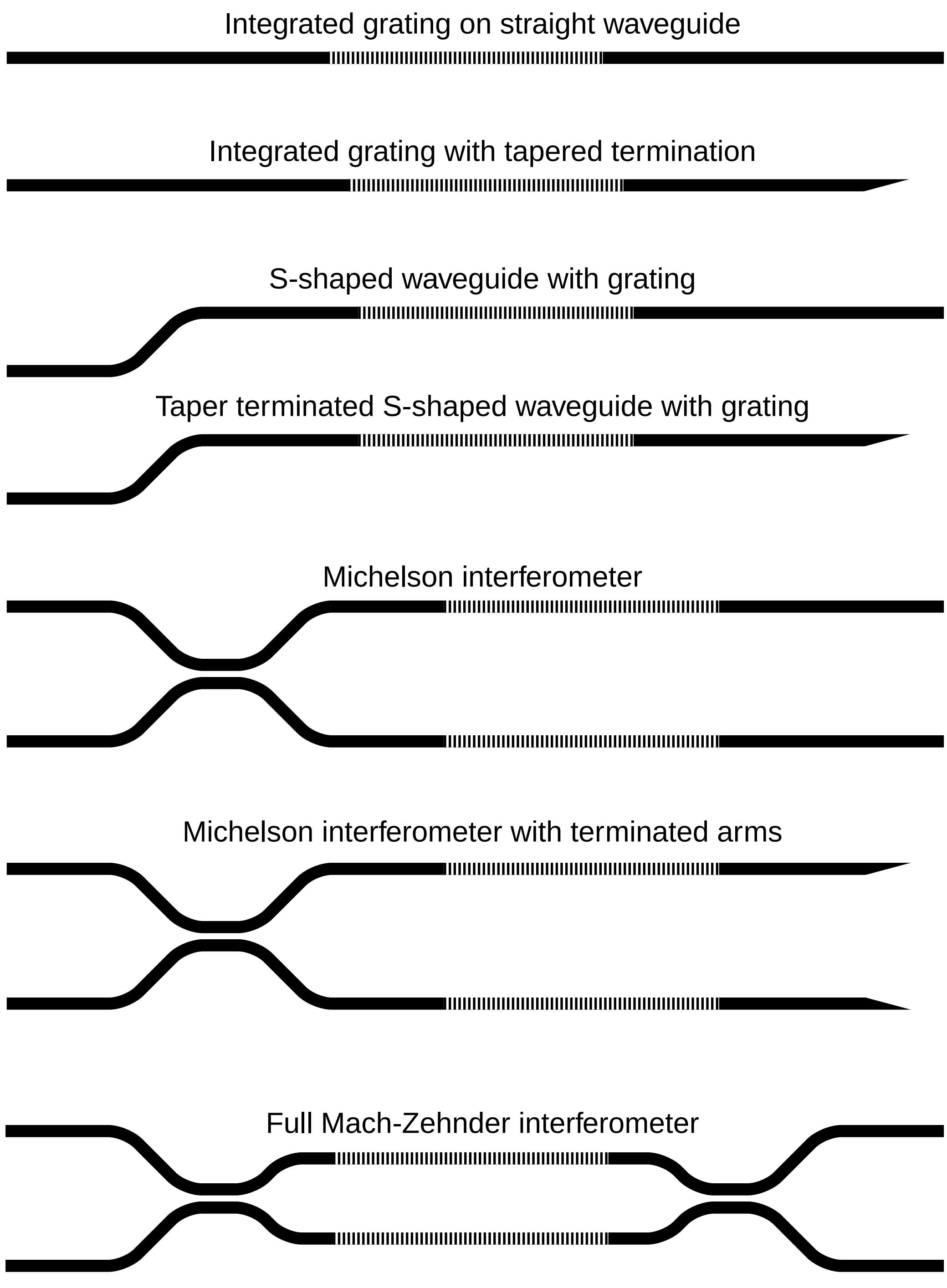


Figure 7.14 The types of waveguide devices included on the optical mask set.

The remaining devices on the chip are interferometer devices consisting of pairs of waveguides with identical gratings. As with the straight waveguides, some of the devices are terminated on-chip by gradually tapering the waveguides, whereas other devices are designed to extend from facet to facet. The geometry of the coupler is described elsewhere in this thesis. Several different coupler lengths are included in the mask, in order to correct for possible deviations in the fabrication process. Some of the devices have a second 50% coupler at the output, forming a Mach-Zehnder interferometer. Without the second coupler at the output, the light transmitted through the gratings is divided equally between the upper and lower waveguides. The second coupler recombines the transmitted light in the lower waveguide (or in the upper waveguide, if light is initially launched into the lower waveguide). For all of the interferometer devices, the waveguide center-to-center separation at the chip facets is 250 μm. The minimum separation between adjacent devices is 75 μm.

For diagnostic purposes, the waveguide mask also contains a series of devices that do not have grating patterns. These include straight waveguides, s-bend waveguides, couplers, and paired couplers.

The devices are separated into three chips, of different length. The longest chip is 30 mm, which is just long enough to contain a full-length Mach-Zehnder interferometer with gratings in the arms. The second chip, which contains Michelson interferometer devices and straight waveguides, is 23 mm in length. The third chip, which contains waveguides without gratings, is 20 mm in length. Figure 7.15 illustrates how these three chips are arranged on the mask set. An outline of the perimeter of the x-ray mask membrane is superimposed for reference.

The optical masks were written with electron-beam-lithography by an outside photomask supplier, using patterns provided by us. One of the challenges in constructing this type of integrated optical device is to translate complex geometrical shapes (including curves and tapers) into a sequence of simple shapes that can be written by an electron-beam-lithography system. Patterns which are sent to the e-beam system must be organized into rectangular shapes which lie on a Cartesian grid. The minimum addressable size is called the grid size or address unit. In designing an optical waveguide mask, the complicated curved and sloped waveguides must be approximated by a series of rectangular shapes lying on a Cartesian grid. To this end, we have developed specialized object-oriented software to translate an abstract description of a device into a series of primitive pixelated rectangles.

For the optical waveguide mask, wave have selected a grid size of 250/2048 = 0.1220703125 μm. The justification for this unconventional choice of grid size is twofold. First, the grid size for the optical waveguide mask is selected to be small enough that when the curved and sloped waveguide sections are rounded to a rectangular grid, the stair-stepping effect is sufficiently small. If the fabricated devices have sidewalls that exhibit stair-stepping, the waveguides could be subject to excess losses caused by scattering from the non-smooth boundary. Because the waveguide mask patterns are to be transferred to the substrate with optical lithography, a stair-stepping on the order of 0.12 μm should be outside of the resolution limits of the lithography. The second justification for the selected grid size is that it allows a single

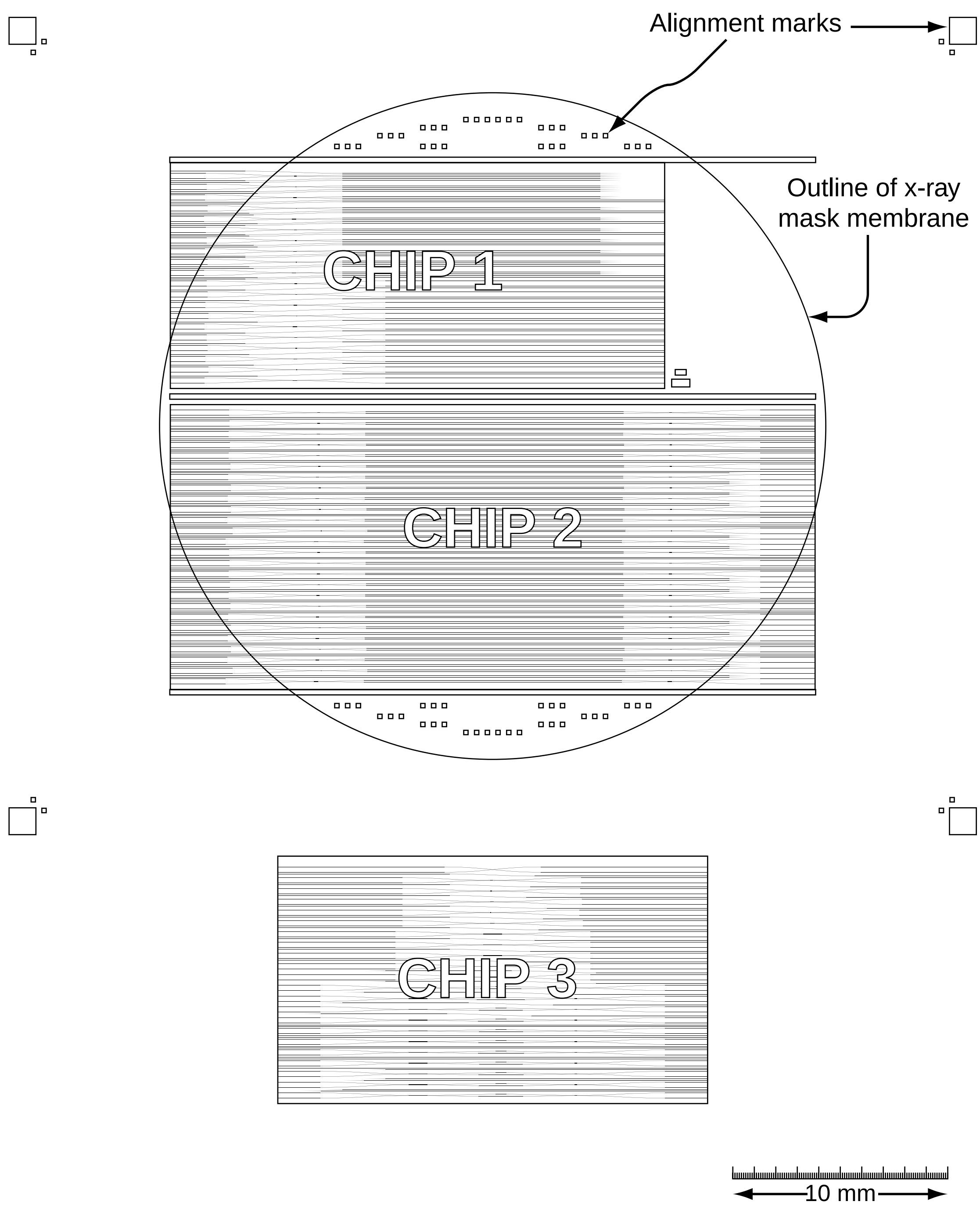


Figure 7.15 CAD diagram of the optical mask set. The mask contains three sets (chips) of optical waveguide devices, each with different length. The circle represents the perimeter of the x-ray mask which is to be used for patterning the Bragg gratings. The devices in chip 3 are waveguide devices without integrated Bragg gratings. Alignment marks are placed at the periphery of each mask, to allow alignment of the three lithography steps. Several sets of redundant marks have been included on the optical mask to allow for flexibility in the placement of marks on the x-ray mask. Note that the print resolution is insufficient to discern the individual waveguide patterns in the above figure.

250 μm wide device to fit exactly within one continuous e-beam field. The commonly used MEBES electron-beam-lithography systems compose patterns by stitching together fields (also called stripes) which are at most 2048 steps across. Thus, by selecting a grid size of 250/2048 we can ensure that an interferometer device (which is 250 μm wide) can be written within a single stripe. This choice of grid size allows us to ensure that the arms of the interferometers do not cross stripe boundaries by judiciously choosing the positions of the devices on the mask.

The grating window mask contains only relatively simple rectangular patterns. The size of these patterns is not as critical as for the waveguide patterns. Therefore, for the grating window mask a grid size of 250/512 = 0.48828125 μm was used. This grid size is a factor of four coarser than the waveguide mask.

7.5.2 Lithography and Materials Testing

Our preliminary tests use the fabrication process described in Section 7.4, with the exception that we have not yet implemented the grating alignment scheme described in Section 7.4.4. The fabrication steps were carried out on thick oxide monitors, which are less expensive and more readily available than the doped glass substrates fabricated by flame hydrolysis. The oxide monitor samples consist of a 4 μm thick oxide layer deposited by CVD onto 4 inch silicon substrates. The tungsten hard mask was patterned by wet chemical etching using a photoresist layer as a wet etch mask. As we will show, the wet chemical etching process causes the tungsten pattern to have rough edges characterized by the coarse grain structure of the thick tungsten layer. This roughness translates into striations in the walls of the waveguide.

Figure 7.16 illustrates the result of the double exposure process described in Section 7.4.2. Recall that the 500 nm-pattern is produced over long rectangular regions by performing a two-step expose and develop sequence in PMMA, followed by a chrome liftoff. In order to obtain accurate alignment of the grating to the waveguides, it is necessary for these two exposures to be aligned, which requires alignment marks on the grating x-ray mask. However, for these tests, only a coarse visual alignment was performed.

Having patterned the grating regions in chrome on the substrate, we sputter-coated a 500 nm thick tungsten layer on top of the chrome pattern, and spun a 420 nm layer of Shipley 1813 photoresist over the tungsten. We then performed an aligned optical exposure defining the waveguide patterns in the photoresist. The exposed tungsten regions were removed by wet chemical etching and the photoresist was stripped. Note that in wet etching of the tungsten layer, the underlying chrome pattern is unaffected. Figure 7.17 is a micrograph depicting the patterned tungsten layer covering a thin chrome grating pattern. Notice in Fig. 7.17 that the tungsten pattern is not perfectly centered within the grating region. However, because the grating region is wider than the waveguide pattern, the vertical misalignment is within tolerance. Even though the tungsten layer is a factor of 16 thicker than the underlying chrome

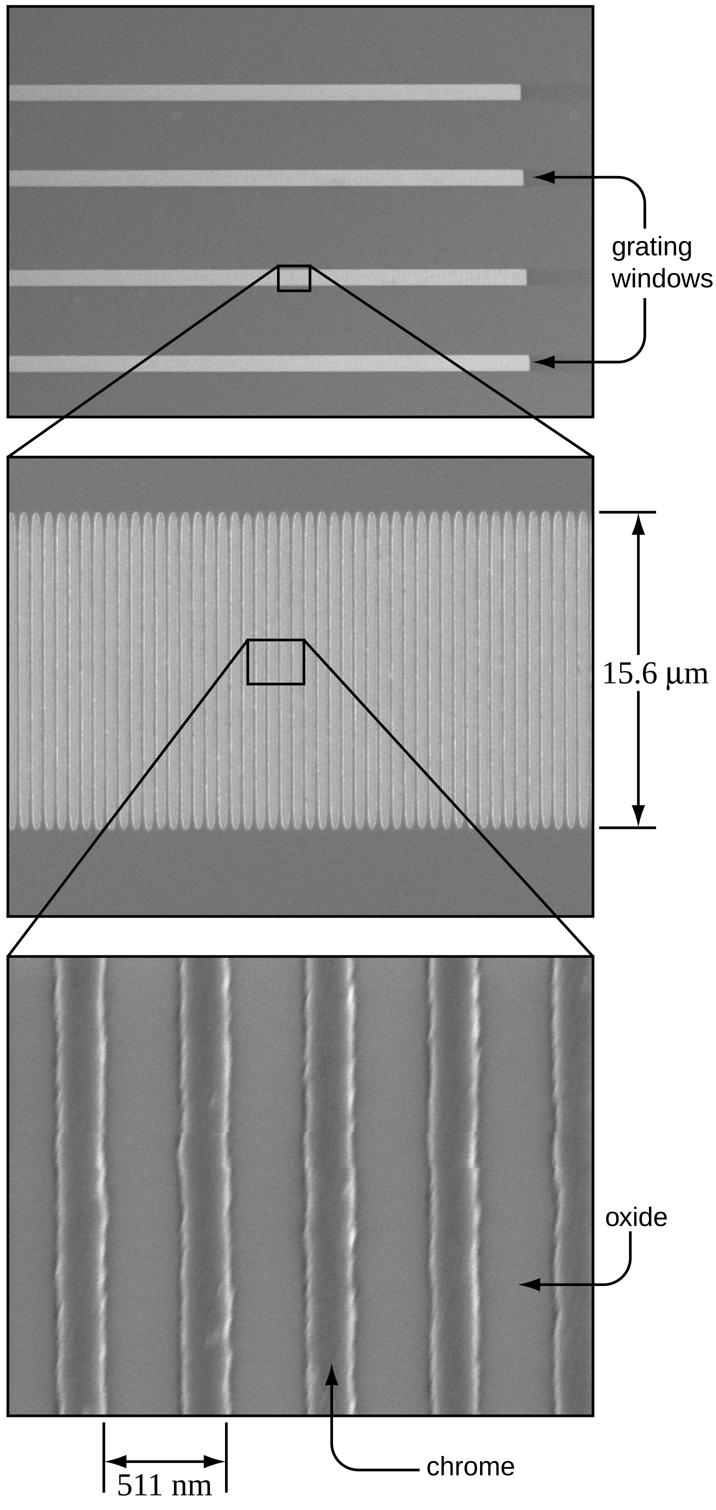


Figure 7.16 Scanning electron micrograph illustrating a chrome grating pattern produced by a two-step exposure sequence. The first exposure in the sequence defines the large rectangular grating windows, and the second exposure patterns the sub-micron grating pattern within these windows.

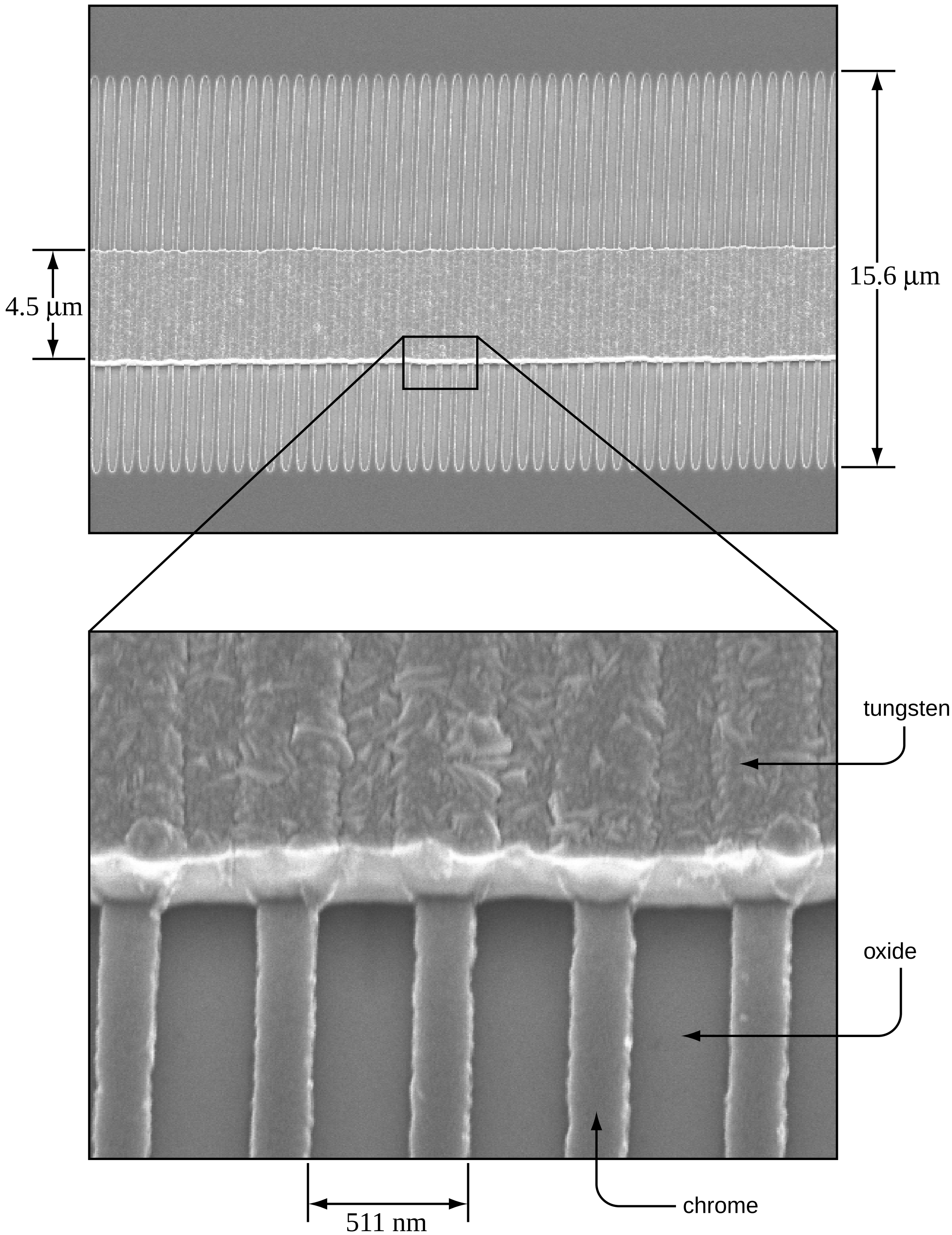


Figure 7.17 Micrographs depicting a patterned tungsten layer on top of a thin chrome grating pattern. The tungsten layer was patterned with conventional photolithography and wet chemical etching. Note that the etching of the tungsten does not disturb the underlying chrome pattern.

pattern, the features of the chrome pattern are still delineated on the surface of the thick tungsten layer, perhaps because of the selective nucleation of tungsten at the edges of the chrome pattern. Because of this characteristic of the tungsten deposition, we were able to perform the alignment of the waveguide mask to the substrate without removing the tungsten from the alignment mark regions on the substrate.

As depicted in Fig. 7.17, after patterning the tungsten hard mask, parts of the underlying chrome pattern are exposed. We then selectively wet-etched the exposed chrome patterns, using the tungsten layer as an etch mask. Those parts of the chrome pattern which are buried beneath the thick tungsten layer are protected from the wet etch, and will later form a hard mask for etching the grating. We then transferred the waveguide patterns into the SiO_2 substrate with reactive-ion-etching in CHF_3. We used a modified Perkins Elmer system, a CHF_3 flow rate of 15 sccm, a chamber pressure of 10 mTorr, and an RF power of 100 W. The measured oxide etch rate under these conditions was 39 nm/min. The tungsten etch rate under the same conditions is approximately 3 nm/min, giving an etch selectivity of 15:1. Figure 7.18 is a micrograph of the etched waveguide pattern, before (A) and after (B) the tungsten mask has been removed by a wet chemical etch. The underlying chrome grating pattern remains intact after etching the waveguide and stripping the tungsten mask. For these tests, the waveguides were etched to a depth of 1.5 μm.

Finally, we etched the gratings to a depth of 250 nm, using the chrome pattern depicted in Fig. 7.18B as a hard mask. The same etching parameters given above were used for this step, with a shorter etch time. Figure 7.19 is a micrograph depicting the resulting grating structures, after the chrome was stripped.

The results presented in Fig. 7.19 demonstrate the feasibility of the fabrication process outlined in Section 7.4. We have employed a scheme where all of the exposure steps are performed before any features are etched into the optical material. Tungsten and chrome form a suitable pair of etch-mask materials because they can be selectively removed with wet chemical etches, as demonstrated. This scheme allows the patterning of relatively shallow sub-micron period gratings on top of tall waveguide structures.

7.6 Conclusions

Design of an integrated optical device such as the matched filters described in this work must be intimately coordinated with the available fabrication technology. The limitations of the fabrication technology often place constraints on the device design, and conversely the goals of the device design can often motivate advances in fabrication technology. This chapter describes the relevant fabrication technologies involved in constructing integrated grating devices, which include interferometric lithography, x-ray lithography, electron-beam-lithography, and conventional photolithography. We also consider how the device is predicted to perform under deviations from the ideal design conditions, devoting particular

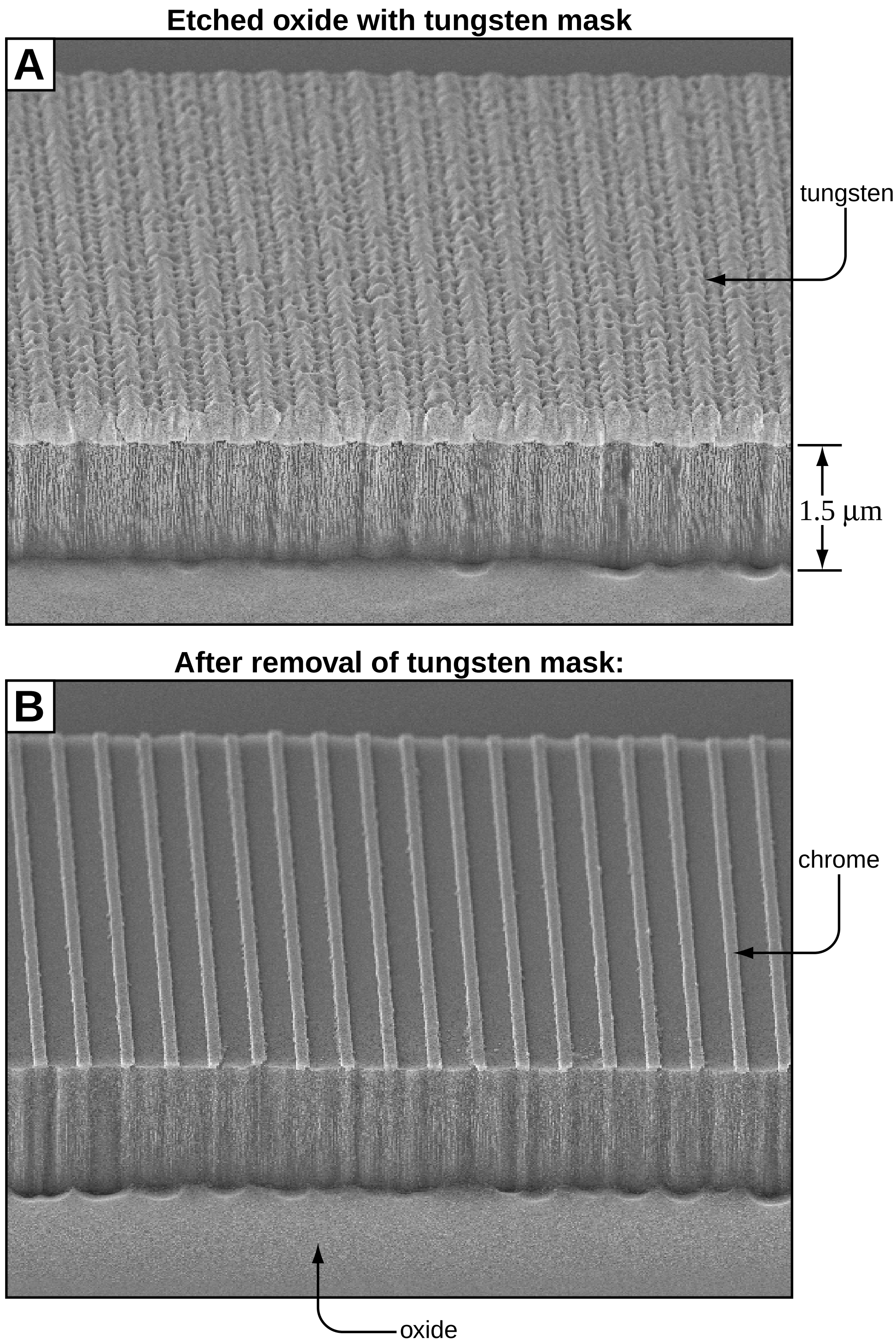


Figure 7.18 Micrograph illustrating etched oxide layer. The thick tungsten pattern (A) acts as an etch mask for reactive ion etching of the oxide layer. After etching to a depth of 1.5 μm, the tungsten is removed with a wet chemical etch, revealing the underlying chrome grating pattern (B).

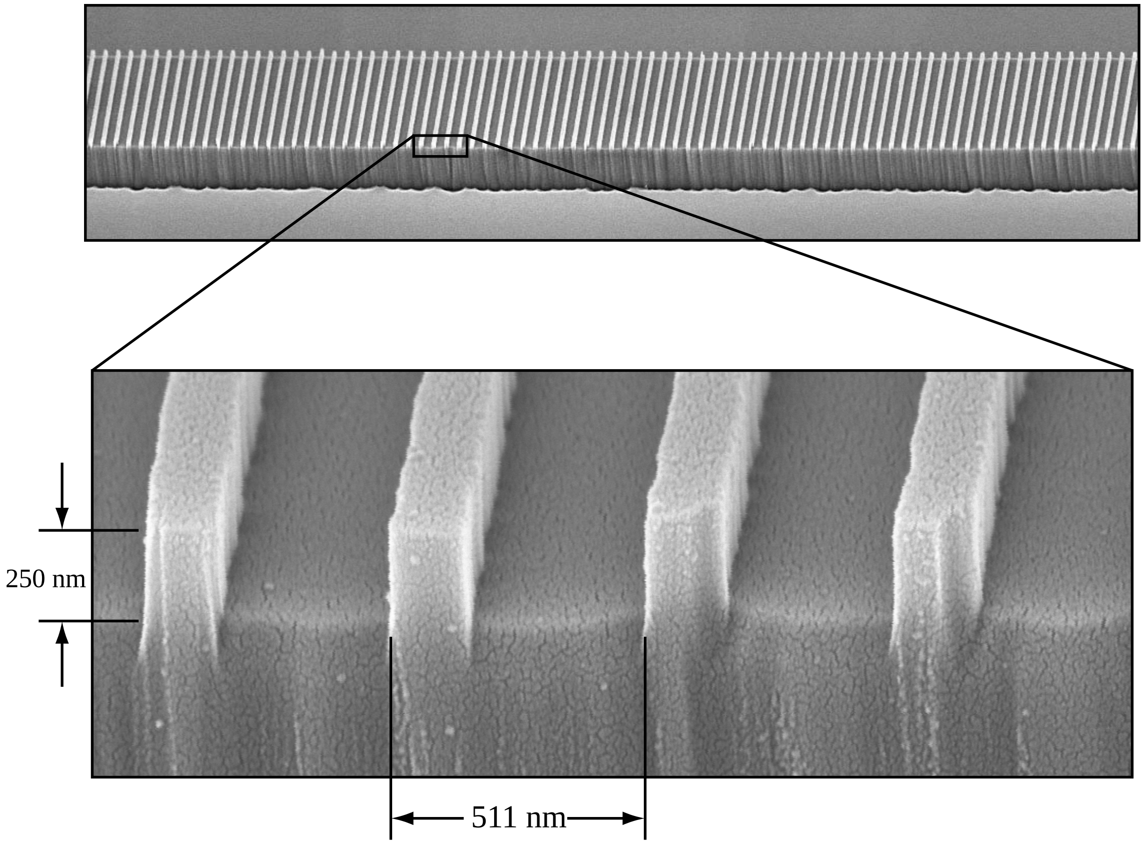


Figure 7.19 Micrograph depicting a shallow Bragg grating etched into the top surface of a tall waveguide

attention to alignment tolerances and waveguide dimension tolerances. These considerations motivate both the design of the devices (e.g. the minimum waveguide separation is selected to achieve a robust design) and the fabrication techniques employed (e.g. we described a novel alignment system that allows angular alignment of the grating with the waveguides.) Finally, we presented a fabrication sequence which addresses some of these issues, and we provided experimental results which demonstrate the effectiveness of the fabrication process.

Chapter 8: Conclusions

Filters are needed in any optical communications system in which noise is included with the signal to be detected. In modern communications systems, the limiting noise source is the amplified spontaneous emission noise generated in optical amplifiers. By filtering noise from the signal to be detected, the communications system becomes more *sensitive*, meaning that less signal power is required to achieve the same error-rate performance. The filters described in this thesis seek to push the limit of receiver sensitivity, by optimally filtering the noisy communications signal. The matched filter is predicted to yield the highest attainable signal-to-noise ratio for the filtered signal. In this work, we have concentrated on providing a thorough description of the theory and design of integrated matched filters, and we have presented results of our initial efforts to fabricate integrated grating structures.

Although this thesis seeks to provide a complete and comprehensive explanation of integrated matched filter devices, it must be pointed out that the research described herein is truly work in progress. In the final portion of this chapter, we outline the work yet to be completed.

Work in the immediate future will focus on fabrication of integrated grating structures. Much of the work to be done was mentioned or implied in the discussions presented in chapter 7.

One aspect of the fabrication that has not yet been implemented is the patterning of alignment marks on the interferometrically generated x-ray mask. Chapter 7 proposes a scheme that would place alignment marks on an x-ray mask in a way that allows the grating to be properly aligned in angle to the waveguides.

Another aspect of the fabrication that requires more investigation is the lithography procedure used to define the optical waveguides. The waveguide structures presented in Section 7.5 were patterned using photolithography and subsequent wet-etching of a tungsten layer. The tungsten layer must be relatively thick in order to withstand a deep oxide etch. Therefore, patterning the tungsten layer by wet-etching is prone to edge granularity and undercutting. In order to obtain smoother tungsten patterns and better control of the waveguide width, it will be necessary to use dry-etching techniques rather than wet-etching techniques. There might also be other pairs of etch mask materials that are more suitable than chrome and tungsten.

Once the fabrication of the gratings is complete, it will be necessary to have a top cladding layer of SiO_2 deposited over the devices, completing the waveguide structures.

Testing of the devices will begin with spectral measurements of the device response, and comparison with the theoretical results predicted in this work. The true performance measure of the devices will be their effect on the sensitivity of a preamplified optical receiver. For these measurements, we plan to incorporate the matched filters into a 10 GB/s optical communications system at Lincoln Laboratory, and characterize the bit-error rate performance of the receiver with and without the matched filter present.

REFERENCES

[1] John E. Midwinter. *Optical Fibers for Transmission*. John Wiley & Sons, NY, 1979.

[2] Paul S. Henry. Error-rate performance of optical amplifiers. *Optical Fiber Communications Conference*, paper THK3, 1989.

[3] S. R. Chinn. Error-rate performance of optical amplifiers with Fabry-Perot filters. *Electronics Letters*, vol. 31, no. 9, pp. 756–757, April 1995.

[4] T. Okoshi and K. Kikuchi. *Coherent Optical Fiber Communications*. KTK Scientific Publishers, Tokyo, 1988.

[5] Gunnar Jacobsen. *Noise in Digital Optical Transmission Systems*. Artech House, Boston, MA, 1994.

[6] Charles H. Henry. Theory of Spontaneous Emission Noise in Open Resonators and its Applications to Lasers and Optical Amplifiers. *Journal of Lightwave Technology*. vol. LT-4, no. 3, pp. 288–297, March 1986.

[7] Peyton Z. Peebles, Jr.. *Digital Communications Systems*. Prentice-Hall, Englewood Cliffs, NJ., 1987.

[8] J. Stone and L. W. Stulz. High-Performance Fibre Fabry-Perot Filters. *Electronics Letters*, vol. 27, no. 24, pp. 2239–2240, November 1991.

[9] J. C. Livas. High Sensitivity Optically Preamplified 10 Gb/s Receivers. Optical Fiber Communication, 1996.

[10] J. C. Livas. A One-Watt, 10-Gbps High-Sensitivity Optical Communication System. *Photonics Technology Letters*, vol. 7, no. 5, pp. 579–581, May 1995.

[11] Charles Kittel. *Introduction to Solid State Physics, 6th ed*. John Wiley & Sons, New York, 1986.

[12] D. K. W. Lam and B. K. Garside. Characterization of single-mode optical fiber filters. *Applied Optics*, vol. 20, no. 3, pp. 440–445, February 1981.

[13] V. V. Wong, J. Ferrera, J. N. Damask, T. E. Murphy, Henry I. Smith, and H. A. Haus. Distributed Bragg grating integrated-optical filters: Synthesis and fabrication. *Journal of Vacuum Science and Technology B*, vol. 13, no. 6, pp. 2859–2864. Nov/Dec 1995.

[14] Amnon Yariv. *Optical Electronics*. Holt, Rinehart, and Winston. New York, 1985.

[15] Alan W. Snyder. *Optical Waveguide Theory*. Chapman & Hall, 1983.

[16] Donald Lee. *Electromagnetic Principles of Integrated Optics.* John Wiley & Sons, NY, 1986.

[17] Dietrich Marcuse. *Light transmission optics*. Van Nostrand Reinhold. New York, 1982.

[18] Hermann A. Haus. *Waves and Fields in Optoelectronics*. Prentice-Hall, Englewood Cliffs, NJ, 1984.

[19] J. D. Jackson. *Classical Electrodynamics, 2nd ed*. John Wiley & Sons, New York, 1975.

[20] David K. Cheng. *Field and Wave Electromagnetics, 2nd. ed*. Addison-Wesley, Reading, MA, 1990.

[21] David J. Griffiths. *Introduction to Electrodynamics, 2nd ed*. Prentice-Hall, Englewood Cliffs, NJ, 1989.

[22] H. Kogelnik, Theodore Tamir (ed.). *Guided-Wave Optoelectronics, 2nd ed.* (2nd chapter) Springer-Verlag, Berlin, 1990.

[23] Haruhito Noro and Tsuneyoshi Nakayama. A New Approach to Scalar and Semivector Mode Analysis of Optical Waveguides. *Journal of Lightwave Technology*, vol. 14, no. 6, pp. 1546–1556, June, 1996.

[24] J. A. Kong,. *Electromagnetic Wave Theory, 2nd ed.* John Wiley & Sons, New York, 1990.

[25] E. C. M. Pennings. *Bends in Optical Ridge Waveguides, Modeling and Experiments.* Delft University of Technology, PhD Thesis, 1990.

[26] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics*. John Wiley & Sons, New York, 1977.

[27] C. L. Xu, W. P. Huang, M. S. Stern, S. K. Chaudhuri. Full Vectorial Mode Calculations by Finite Difference Method. *IEE Proceedings-Optoelectroncs*, vol. 141, no. 5, pp. 281–286, October 1994.

[28] M. S. Stern. Semivectorial polarised finite difference method for optical waveguides with arbitrary index profiles. *IEE Proceedings*, vol. 135, part J, no. 1, pp. 56–63, February, 1988.

[29] C. Yeh, K. Ha, S. B. Dong, and W. P. Brown. Single-mode optical waveguides. *Applied Optics*. Vol. 18, no. 10, pp. 1490–1504. May 1979.

[30] Bernice M. Dillon and Jon P. Webb. A Comparison of Formulations for the Vector Finite Element Analysis of Waveguides. *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 2, pp. 308–316, February, 1994.

[31] Zine-Eddine Abid, Klein L. Johnson, and Anand Gopinath. Analysis of Dielectric Guides by Vector Transverse Magnetic Finite Elements. *Journal of LightwaveTechnology*, vol. 11, no. 10, pp. 1545–1549, October, 1994.

[32] John M. Senior. *Optical Fiber Communications*. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[33] Fred S. Hickernell. Optical Waveguides on Silicon. *Solid State Technology*, pp. 83–87, November 1988.

[34] S. Valette, S. Renard, H. Denis, J. P. Jadot, A. Fournier, P. Philippe, P. Gidon, A. M. Grouillet, and E. Desgranges. Si-Based Integrated Optics Technologies. *Solid State Technology*, pp. 69–75, February 1989.

[35] C. H. Henry, G. E. Blonder, and R. F. Kazarinov. Glass Waveguides on Silicon for Hybrid Optical Packaging. *Journal of Lightwave Technology*, vol. 7, no. 10, pp. 1530–1539, October 1989.

[36] B. H. Verbeek, C. H. Henry, N. A. Olsson, K. J. Orlowsky, R. F. Kazarinov, and B. H. Johnson. Integrated Four-Channel Mach-Zehnder Multi/Demultiplexer Fabricated with Phosphorous Doped SiO_2 Waveguides on Si. *Journal of Lightwave Technology*, vol. 6, no. 6, pp. 1011–1015, June 1988.

[37] Masao Kawachi. Silica waveguides on silicon and their application to integrated-optic components, *Optical and Quantum Electronics*, vol. 22, pp. 391–416, 1990.

[38] Tatsuo Izawa, Hidefumi Mori, Yasuji Murakami, and Nobuo Shimizu. Deposited silica waveguide for integrated optical circuits. *Applied Physics Letters*, vol. 38, no. 7, pp. 483–485, April 1981.

[39] T. Kominato, Y. Ohmori, H. Okazaki, and M. Yasu. Very low-loss GeO_2-doped silica waveguides fabricated by flame hydrolysis deposition method. *Electronics Letters*, vol. 26, pp. 327–328, March, 1990.

[40] N. Takato, M. Yasu, and M, Kawachi. Low-Loss High-Silica Single-Mode Channel Waveguides. *Electronics Letters*, vol. 22, no. 6, pp. 321–322, 1986.

[41] Norio Takato, Kaname Jinguji, Mitsuho Yasu, Hiromu Toba, and Masao Kawachi. Silica-Based Single-Mode Waveguides on Silicon and their Application to Guided-Wave Optical Interferometers. *Journal of Lightwave Technology*, vol. 6, no. 6, pp. 1003–1010, June 1988.

[42] Hideaki Okayama, Masato Kawahara, and Takeshi Kamijoh. Reflective Waveguide Array Demultiplexter in $LiNbO_3$. *Journal of Lightwave Technology*, vol. 14, no. 6, p. 985. June 1996.

[43] Louay Eldada, Chengzeng Xu, Kelly M. T. Stengel, Lawrence W. Shacklette, and James T. Yardley. Laser-Fabricated Low-Loss Single-Mode Raised-Rib Waveguiding Devices in Polymers. *Journal of Lightwave Technology*, vol. 14, no. 7, July 1996.

[44] Amnon Yariv and Michiharu Nakamura. Periodic Structures of Integrated Optics. *Journal of Quantum Electronics*, vol. QE-13, no. 4, pp. 233–253, April, 1977.

[45] William Streifer, Donald R. Scifres, and Robert Burnham. Coupled Wave Analysis of DFB and DBR Lasers. *Journal of Quantum Electronics*, vol. QE-13, no. 4, pp. 134–141, April 1977.

[46] A. W. Snyder. Coupled-mode theory for optical fibers. *Journal of the Optical Society of America*, vol. 65, pp. 810–813, 1972.

[47] Yoshihisa Yamamoto, Takeshi Kamiya, and Hisayoshi Yanai. Improved Coupled Mode Analysis of Corrugated Waveguides and Lasers. *IEEE Journal of Quantum Electronics*, vol. QE-14, no. 4, pp. 245–258, April 1978.

[48] Jay N. Damask. *Integrated-Optic Grating-Based Filters For Optical Communication Systems*. PhD Thesis, Dept. of Electrical Engineering, MIT, May 1996.

[49] William Streifer, Don. R. Scifres, and Robert D. Burnham. Analsysis of Grating-Coupled Radiation in GaAs:GaAlAs Lasers and Waveguides. *IEEE Journal of Quantum Electronics*, vol. QE-12, no. 7, pp. 422–428, July 1976.

[50] Eli Kapon, A. Hardy, and Abraham Katzir. The Effect of Complex Coupling Coefficients on Distributed Feedback Lasers. *IEEE Journal of Quantum Electronics*, vol. QE-18, no. 1, pp. 66–71, January 1982.

[51] Jay S. Van Delden. Optical circulators improve bidirectional fiber systems. *Laser Focus World*, pp. 109–112, November 1995.

[52] Optics for Research, product information, 1995.

[53] R. Kashyap, G. D. Maxwell, and B. J. Ainslie. Laser-Trimmed Four-Port Bandpass Filter Fabricated in Single-Mode Photosensitive Ge-Doped Planar Waveguide. *IEEE Photonics Technology Letters*, vol. 5, no. 2, pp. 191–194, February, 1993.

[54] F. Bilodeau, D. C. Johnson, S. Theriault, B. Malo, J. Albert, and K. O. Hill. An All-Fiber Dense-Wavelength-Division Multiplexer/Demultiplexer Using Photoimprinted Bragg Gratings. *IEEE Photonics Technology Letters*, vol. 7, no. 4, pp. 388–390, April, 1993.

[55] Eugene Hecht. *Optics, 2nd ed*. Addison-Wesley, Reading, MA, 1987.

[56] K.-L. Chen and S. Wang. Cross-Talk Problems in Optical Directional Couplers. *Applied Physics Letters*, vol. 44, pp. 166–168, 1984.

[57] H. A. Haus. Coupled-mode theory revisited. *SPIE - Integrated Optical Circuit Engineering IV*, vol. 704, pp. 68–78.

[58] H. A. Haus, W. P. Huang, S. Kawakami, and N. A. Whitaker. Coupled Mode Theory of Optical Waveguides. *Journal of Lightwave Technology*, vol. LT-5, no. 1, pp. 16–23, January 1987.

[59] Hermann A. Haus and Weiping Huang. Coupled-Mode Theory. *Proceedings of the IEEE*, vol. 79, no. 10, pp. 1505–1518, October 1991.

[60] H. A. Haus and N. A. Whitaker, Jr. Elimination of cross talk in optical directional couplers. *Applied Physics Letters*, vol. 46, no. 1, pp. 1–3, January 1985.

[61] Z. Weissman, A. Hardy and E. Marom. On the Applicability of the Coupled Mode Theory to Non-parallel Waveguide Systems. *Optics Communications*, vol. 71, no. 6, pp. 341–344, June 1989.

[62] Weiping Huang and Herman A. Haus. Self-Consistent Vector Coupled-Mode Theory for Tapered Optical Waveguides. *Journal of Lightwave Technology*, vol. 8, no. 6, pp. 922–926, June 1990.

[63] H. A. Haus and W.-P. Huang. Mode Coupling in Tapered Structures. *Journal of Lightwave Technology*, vol. 7, no. 4, pp. 729–730, April 1989.

[64] J. Goell. A circular-harmonic computer analysis of rectangular dielectric waveguides. *Bell Systems Technical Journal*, vol. 48, pp. 2071–2102, 1969.

[65] Yoshinori Hibino, Fumiaki Hanawa, Hiroshi Nakagome, Motohaya Ishii, and Norio Takato. High Reliability Optical Splitters Composed of Silica-Based Planar Lightwave Circuits. *Journal of Lightwave Technology*, vol. 13, no. 8, pp. 1728–1735, August 1995.

[66] C. M. Weinert. Design of Fiber-Matched Uncladded Rib Waveguides on InP with Polarization-Independent Mode Matching Loss of 1 dB. *Photonics Technology Letters*, vol. 8, no. 8, pp. 1049–1051, August 1996.

[67] Osamu Mitomi, Kazuo Kasaya, Yuichi Tohmori, Yasumasa Suzaki, Hideki Fukano, Yoshihisa Sakai, Minoru Okamoto, and Shin-ichi Matsumoto. Optical Spot-Size Converters for Low-Loss Coupling Between Fibers and Optoelectronic Semiconductor Devices. *Journal of Lightwave Technology*, vol. 14, no. 7, pp. 1714–1719, July 1996.

[68] R. E. Smith, C. T. Sullivan, G. A. Vawter, G. R. Hadley, J. R. Wendt, M. B. Snipes, J. F. Klein. Reduced Coupling Loss Using a Tapered-Rib Adiabatic-Following Fiber Coupler. *Photonics Technology Letters*, vol. 8, no. 8, pp. 1052–1054, August 1996.

[69] Z. Weissman and I. Hendel. Analysis of Periodically Segmented Waveguide Mode Expanders. *Journal of Lightwave Technology*, vol. 13, no. 10, pp. 2053–2058, October 1995.

[70] T. E. Rozzi. Rigorous Analysis of the Step Discontinuity in a Planar Dielectric Waveguide. *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-26, no. 10, pp. 738–746, October 1978.

[71] H. M. de Ruiter. Radiation from the junction of two open ended waveguide sections. *Radio Science*, vol. 19, no. 5, pp. 1295–1300, 1984.

[72] Mordehai Heiblum and Jay H. Harris. Analysis of Curved Optical Waveguides by Conformal Transformation. *IEEE Journal of Quantum Electronics*, vol. QE-11, no. 2, pp. 75–83, February 1995.

[73] E. A. J. Marcatili. Dielectric rectangular waveguide and directional coupler for integrated optics. *Bell Systems Technical Journal*, vol. 48, pp. 2071–2102, 1969.

[74] Hermann A. Haus, Wei-Ping Huang, and N. M. Whitaker. Optical Waveguide Dispersion Characteristics from the Scalar Wave Equation. *Journal of Lightwave Technology*, vol. LT-5, no. 12, pp. 1748–1754, 1987.

[75] Corning Product Information, (http://www.usa.net/corning-fiber/products:), 1996.

[76] Meint K. Smit, Erik C. M. Pennings, and Hans Blok. A Normalized Approach to the Design of Los-Loss Optical Waveguide Bends. *Journal of Lightwave Technology*, vol. 11, no. 11, pp. 1737–1742, 1993.

[77] Chunsheng Ma and Shiyong Liu. Optical characteristics of bent dielectric rectangular waveguides. *Optical and Quantum Electronics*, vol. 19, pp. 83–92, 1987.

[78] Ajoy K. Ghatak. Leaky modes in optical waveguides. *Optical and Quantum Electronics*, vol. 17, pp. 311–321, 1987.

[79] K. Thyagarajan, M. R. Shenoy, and A. K. Ghatak. Accurate numerical method for the calculation of bending loss in optical waveguides using a matrix approach. *Optics Letters*, vol. 12, no. 4, pp. 296–298, 1987.

[80] Ajoy K. Ghatak, K. Thyagarajan, and M. R. Shenoy. Numerical Analysis of Planar Optical Waveguides Using Matrix Approach. *Journal of Lightwave Technology*, vol. LT-5, no. 5, pp. 660–667, May 1987.

[81] I. C. Goyal, R. L. Gallawa, and A. K. Ghatak. Bent Planar Waveguides and Whispering Gallery Modes: A New Method of Analysis. *Journal of Lightwave Technology*, vol. 8, no. 5, pp. 768–773, May 1987. (Correction in vol. 11, no. 12, p. 2206, December 1987.)

[82] M. Abramowitz and I. A. Stegun (Eds.). *Handboook of Mathematical Functions*. Dover publications, New York, 1965.

[83] S. J. Garth. Perturbation analysis of bent waveguides. *IEE Proceedings*, vol. 134, pt. J, no. 4, pp. 221–229, August, 1987.

[84] Yuehue Cheng, Weigan Lin, and Yoichi Fujii. Local Field Analysis of Bent Graded-Index Planar Waveguides. *Journal of Lightwave Technology*, vol. 8, no. 10, pp. 1461–1469, October 1990.

[85] Ernst-Georg Neumann. Curved dielectric optical waveguides with reduced transition losses. *IEE Proceedings*, vol. 129, pt. H, no. 5, pp. 278–280.

[86] Tsutomu Kitoh, Norio Takato, Mitsuho Yasu, and Maseo Kawachi. Bending Loss Reduction in Silica-Based Waveguides by Using Lateral Offsets. *Journal of Lightwave Technology*, vol. 13, no. 4, pp. 555–562, April 1995.

[87] Per Inge Jensen and Aasmund Subdo. Bragg Gratings for 1.55-mm Wavelength Fabricated on Semiconductor Material by Grating-Period Doubling Using a Phase Mask. *Photonics Technology Letters*, vol. 7, no. 7, pp. 783–785, July 1995.

[88] W. Chu. *Inorganic X-ray Mask Technology for Quantum-Effect Devices*. PhD Thesis, Dept. of Electrical Engineering, MIT, 1993.

[89] Vincent V. Wong. *Fabrication of Distributed Feedback Devices Using X-Ray Lithography*. PhD Thesis, Dept. of Electrical Engineering, MIT, August 1995.

[90] E. H. Anderson. *Fabrication and Electromagnetic Applications of Periodic Nanostructures*. PhD Thesis, Dept. of Electrical Engineering, MIT, 1988.

[91] Y. Hibino, T. Kitagawa, K. O. Hill, F. Bilodeau, B. Malo, J. Albert, and D. C. Johnson. Wavelength Division Multiplexer with Photoinduced Bragg Gratings Fabricated in a Planar-Lightwave-Circuit-Type Asymmetric Mach-Zehnder Interferometer on Si. *IEEE Photonics Technology Letters*, vol. 8, no. 1, pp. 84–86, January 1996.

[92] J. Albert, B. Malo, K. O. Hill, F. Bilodeau, D. C. Johnson, and S. Theriault. Comparison of one-photon and two-photon effects in the photosensitivity of germanium-doped silica optical fibers exposed to intense ArF excimer laser pulses. *Applied Physics Letters*, vol. 67, no. 24, pp. 3529–3531, 1995.

[93] Yan Duval, Raman Kashyap, and Simon Fleming. Correlation between ultriviolet-induced refractive index change and photoluminescence in Ge-doped fiber. *Applied Physics Letters*, vol. 61, no. 25, pp. 2955–2957, 1995.

INDEX

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

V

W

X