Design of a High Index Contrast Arrayed Waveguide Grating

by

Ryan N. Lang

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

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2003

Author .. ..
Department of Electrical Engineering and Computer Science
May 21, 2003

Certified by.....
Hermann A. Haus
Institute Professor Emeritus
Thesis Supervisor

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

Design of a High Index Contrast Arrayed Waveguide Grating

by

Ryan N. Lang

Submitted to the Department of Electrical Engineering and Computer Science
on May 21, 2003, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Arrayed waveguide gratings (AWGs) are useful structures for the implementation of wavelength division multiplexing. The AWG consists of an input splitter, a dispersive waveguide array which creates the wavelength demultiplexing and multiplexing effects, and an output coupler. Because the dispersive waveguide array consists of bent waveguides, the size of an AWG is limited by the light loss in the bends. In their current form, silica-based gratings are too large to be made cheaply or to use as an integrated component. The proposed solution is to redesign the AWG using high index contrast materials for tight confinement of the waveguide modes and, consequently, low bend loss. A rough design is presented for a high index contrast AWG using multimode interference couplers as the coupling stages. The major components were simulated using finite difference time domain (FDTD) techniques to find low loss but rather high crosstalk. A second possible design is also presented, making use of a coupled waveguide array as the input element. The coupling coefficients of as many as 41 coupled waveguides were adjusted to create a Gaussian profile as an input to the dispersive section of the AWG. The output coupler, however, will make use of more standard free space diffraction techniques, making the overall concept a unique mixture of waveguide and free space optical elements.

Thesis Supervisor: Hermann A. Haus
Title: Institute Professor Emeritus

Acknowledgments

To begin, I would like to thank my advisors, Professor Hermann Haus and Dr. Kazumi Wada. When I started this project, I was practically clueless about the field of integrated optics, but their faith in my abilities never wavered. It has been a delight to work with and learn from them.

I would also like to thank my officemate Milos Popović. Milos is a very busy guy, but he always had time to answer all of my questions, however stupid they might have been. In addition, his friendly company made working much more enjoyable.

I must also thank Christina Manolatou, whose FDTD code I used. The code was critical to all of my work in the first half of this project. Christina also helped me to modify the code slightly beyond its original functionality.

While on the subject of the FDTD code, I am grateful for the use of the longhorn supercomputer at the Texas Advanced Computing Center, without which I could have never finished any of my simulations before graduation!

I would not have been able to finish as much work as I did were it not for the generous support of a Siebel Scholarship provided by Siebel Systems, which paid my tuition and stipend for one semester and saved me from a busy schedule of teaching. In addition, I have to thank Professor Charles Sodini who not only helped me find my project, but also provided me with support over the summer.

Despite all of this help, I would have never made it through five years of MIT without all of the friends I have made here. It has been a delight to know every one of them. I would like to especially thank my best friend, Becky Hitchcock, for always putting up with me and making me smile even on the worst days.

I would never have even made it to MIT if it weren't for my family, especially my mom, dad, sister, and grandparents. Their love, support, and sacrifices made me who I am today and allowed me to follow my dreams. I love them all so much.

Finally, I'd like to thank God for giving me all of these gifts and helping me to make something good come from them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is concerned with the branch of electrical engineering known as photonics. Most students and professionals in electrical engineering are familiar with electronics, in which electric currents and voltages are used to represent information. The invention of the transistor in 1947, followed by the development of electronic integrated circuits, has led to revolutions in information technology, including computation and communication. Less known is the field of photonics, which uses electromagnetic radiation (light) as a medium for representing information. However, the development of optical technology, including such milestones as the invention of the laser in the late 1950s and the development of optical fibers in the 1970s, has been arguably just as revolutionary. Nevertheless, optical technology has normally been used only for large scale telecommunications, while electronics continues to be the technology of choice for localized, integrated applications such as signal processing and computation. It is true that photonics technology is limited; for instance, we may never develop nonlinear switches, such as transistors, using optics. But silicon-based electronics face their own problems in the near future because of fundamental limitations in speed and bandwidth. Many researchers hope to develop the field of photonics so that it can help solve these problems, eventually creating an entire toolbox of integrated components, known as photonic integrated circuits (PICs), as versatile and diverse as the electronic ones [23]. By combining photonic and electronic components, they hope to utilize the strengths of each to offset the weaknesses of the other.

In terms of speed, photonic components already have a great advantage over electronic ones, since light can move as much as one thousand times faster than electrons in semiconductors. This speed is already being utilized in simple ways, such as optical clock distribution, which uses a laser to transport the clock signal to different parts of an electronic chip [43]. Photonics also offers a second advantage over electronics by providing larger bandwidth on a single channel. This effect can be achieved by the technique known as wavelength division multiplexing (WDM). With WDM, light of several different wavelengths can be sent along a channel (fiber-optic cable or integrated waveguide) simultaneously. Each wavelength can carry a distinct signal from the others. In order to successfully apply the concept of wavelength division mutliplexing, some means must be found to multiplex different wavelengths into one channel and then demultiplex them back into separate channels. Such multiplexing and demultiplexing devices have applications in many parts of a larger photonic system. For example, a multiplexer would be used at the input to any communications channel or interconnect, while a demultiplexer would be necessary at the end of such channels to recover the individual signals for further processing [43].

The focus of this thesis is one popular solution to the MUX/DEMUX problem, known as an arrayed waveguide grating (AWG), waveguide grating router (WGR), or phased-array (PHASAR) device. The rest of this chapter describes the AWG in general terms, including previous work on the device, before discussing the problem of bend loss and the proposed solution of high index contrast. Chapter 2 gives a review of basic theory needed for this thesis, including Maxwell's equations, electromagnetic waves, slab waveguides, and bend loss. It also describes the finite difference time domain (FDTD) algorithm used for simulations. Chapter 3 describes our initial design for the AWG, using two multimode interference couplers (MMIs) instead of the more traditional free space coupling regions. After some theory, a design procedure and results for this design are presented. Chapter 4 considers a new design, based partly on the standard AWG but with a new type of input composed of many coupled waveguides. Finally, Chapter 5 summarizes the work and gives ideas for future directions of study.

1.1 Basics of the arrayed waveguide grating

There are several different ways to implement a WDM multiplexer/demultiplexer. These methods can be separated into two categories: resonator-based filters and interference-based filters. Resonator-based filters make use of microring and microcavity structures but ultimately cannot achieve the channel spacing required in many WDM systems. Interference-based filters can meet the channel spacing requirements but with the side effect of increased crosstalk [13]. The simplest type of interference filter uses an actual reflection grating to produce interference effects. However, a more robust and fabrication tolerant method is to use an arrayed waveguide grating (AWG) [34]. A typical AWG is pictured in Figure 1-1 [43]. The structure is made up of three main sections, a splitter, a dispersive waveguide array, and a combiner. An input signal is coupled by the splitter into the array waveguides. These guides all have different lengths, so they introduce wavelength-dependent phase shifts into the signals. The array waveguide outputs are then recombined so that constructive interference will occur at one unique output for each wavelength in the WDM system. In this way the AWG acts as a wavelength demultiplexer. The device can also serve

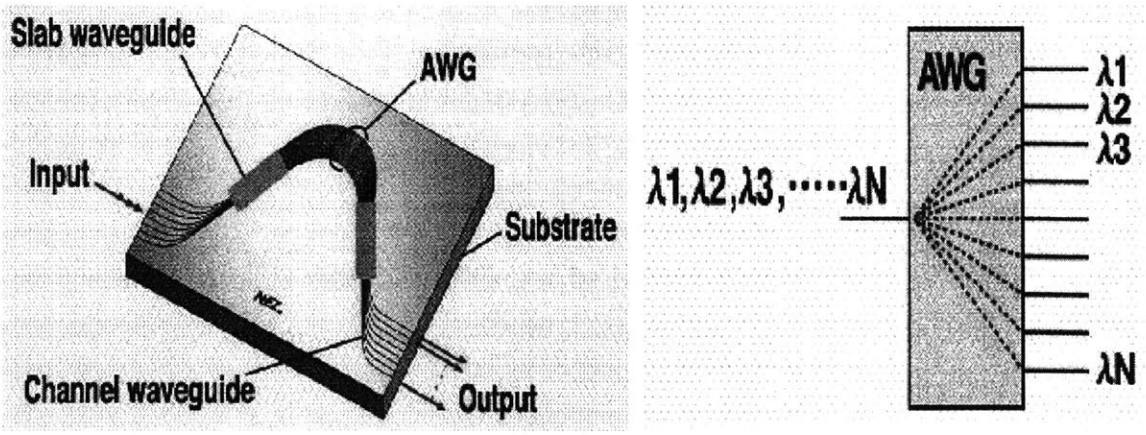

Figure 1-1: Typical AWG (left) and schematic description of WDM application (right) [43]

as a multiplexer when used in the reverse configuration [43].

1.2 Previous work and motivation

The arrayed waveguide grating was first proposed a solution to the WDM problem by Smit in 1988 [33] and was further developed in the following years by Takahashi [40] and Dragone [8], who extended the concept from 1 x N demultiplexers to N x N wavelength routers. Since then, researchers have designed many AWGs seeking to improve them by increasing the number of channels, decreasing the wavelength spacing, increasing transmission, lowering crosstalk, and reducing the size of the device. These AWGs have many applications in addition to simple demultiplexing applications, including add/drop filters [39], cross-connects [15], channel equalization [45], and multifrequency lasers [46].

AWG devices can be mostly divided into two categories: silica-based devices and indium phosphide-based devices [34]. Normally, silica-based waveguides are constructed of silica (SiO_2) waveguides on silica cladding, where doping is used to provide a very small index contrast. The problem with such low index contrast devices is their size; the smallest structures have footprints on the order of square centimeters. This large size is caused by the bend loss in the array waveguides. At a waveguide bend, the phase front velocity is smaller closer to the center of curvature and larger farther from the center of curvature in order to maintain a coherent phase front around the bend. However, at a certain radius in the outer cladding, the required phase front velocity is greater than the velocity of plane waves in the cladding. The exponentially decaying evanescent tail couples to radiation modes, causing radiation loss [21]. This smallest feasible bending radius in standard silica waveguides is on the order of 1 cm, leading to the large device footprints discussed above [13]. Indium phosphide based devices, on the other hand, have higher contrast and can be made smaller, but their size is still limited by the bend loss.

Several ideas have been developed to counteract this problem of large waveguide bends. Asymmetric claddings or air trenches can be used to reduce bend loss by

artificially increasing the index contrast at the outside of the bend. This idea was used with InP ridge waveguides to make a low loss, approximately 1 mm x 1 mm AWG [7]. Air trenches have also been explored in silica buried waveguides [30]. Another idea is to use a completely different, high index contrast (HIC) material system in which the core and cladding are chosen to have very different indices of refraction. This, in turn, causes the electric field to be strongly confined within the core, substantially reducing radiation loss for sharp bends and allowing smaller structures to be produced [43]. For example, silicon ($n = 3.5$) waveguides on silica ($n = 1.46$) cladding allows for 98% transmission bends of 1 μm radius [13]. Unfortunately, such high index contrast waveguides have increased losses due to scattering and are difficult to couple into because of their small size. Furthermore, their performance is polarization dependent. Nevertheless, the field of high index contrast optics is promising [19].

This thesis concerns the design of a small arrayed waveguide grating using a high index contrast material system. This device should have a much smaller footprint than previous AWGs, allowing for better integration capability. The size reduction also allows large numbers of AWGs to be made more cheaply and with greater thermal stability. In addition, the silica material system allows the use of CMOS and LSI processing technology. In fact, this AWG may be eventually integrated onto a chip with a fiber coupler, photodetectors, and electrical amplifiers [13]. This thesis, however, only considers the AWG. It has eight wavelength channels, spaced by 0.8 nm, or 100 GHz. The center wavelength is 1.55 μm, which has the minimum attenuation in optical fibers. The waveguides are assumed to be buried channels of silicon with silica cladding.

The primary concerns of the AWG design are the loss of the device and the crosstalk into neighboring channels. Scattering loss due to the high index contrast can be solved in fabrication [13]. Polarization dependence will be overcome by splitting the incoming signal into transverse electric (TE) and transverse magnetic (TM) components, rotating the TM into TE, propagating both components through identical AWGs, rotating the second component back to TM, and then recombining the polarizations [11]. We present a rough design for this AWG using multimode inter-

ference couplers as the input and output stages. Unlike the traditional free space couplers, MMIs are easy to fabricate and have relatively low loss. However, an MMI-based AWG is not new; others have implemented them before (e.g. [16], [27], [28], [32], [24]). The design presented here follows the basic steps laid out by these authors, but since the size is smaller, some new problems were encountered. Rather than adjust this design, H. Haus proposed a second design, using an array of coupled waveguides as the input stage for maximum control while reverting to the traditional free space coupling region for the output. Preliminary results are obtained for the input stage of this device, and the rest of the design is outlined.

Chapter 2

Theoretical and Computational Tools for Analysis of Integrated Optics

This chapter reviews the basic theoretical and computational tools needed for the design of an arrayed waveguide grating, leaving more specific theory for later chapters. We first present Maxwell's equations, the starting point for all optics theory. After briefly considering plane wave solutions to Maxwell's equations and the behavior of these fields at dielectric interfaces, we move on to the more interesting and relevant problem of slab dielectric waveguides. We find a dispersion relation for the modes of these waveguides using two different analytic techniques and discuss some properties of these modes. Then we detour briefly to look at the problem of bend loss before describing the computational tool used for this project, finite difference time domain (FDTD) simulation. Chapters 3 and 4 will provide more detailed theory and adjustments to the FDTD code specifically relevant to each of the two AWG designs we present.

2.1 Maxwell's equations [10], [18]

Maxwell's equations in a dielectric medium are:

$$\nabla \cdot \mathbf{D} = \rho \tag{2.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.3}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \tag{2.4}$$

Here ρ is the charge density, $\mathbf{J}$ is the current density, $\mathbf{E}$ is the electric field, $\mathbf{H}$ is the magnetic field, $\mathbf{D}$ is the electric displacement, and $\mathbf{B}$ is the magnetic flux density.[1]

We can express the equations in terms of the electric and magnetic fields only if we write

$$\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0(1 + \chi_e)\mathbf{E} \tag{2.5}$$

$$\mathbf{B} = \mu \mathbf{H} = \mu_0(1 + \chi_m)\mathbf{H} \tag{2.6}$$

where ϵ is the electric permittivity of the medium, ϵ_0 is the permittivity of free space, χ_e is the electric susceptibility, μ is the magnetic permeability of the medium, μ_0 is the permeability of free space, and χ_m is the magnetic susceptibility. Maxwell's equations become:

$$\nabla \cdot \epsilon \mathbf{E} = \rho \tag{2.7}$$

$$\nabla \cdot \mu \mathbf{H} = 0 \tag{2.8}$$

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \tag{2.9}$$

$$\nabla \times \mathbf{H} = \epsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \tag{2.10}$$

[1] We call $\mathbf{H}$ the magnetic field instead of $\mathbf{B}$ since $\mathbf{H}$ will be the quantity of most interest in a dielectric.

In optics, we usually consider a source-free medium ($\rho = 0$ and $\mathbf{J} = 0$). We also ignore the effect of magnetization ($\mu = \mu_0$). The index of refraction is defined as:

$$n = \sqrt{\frac{\epsilon}{\epsilon_0}} \tag{2.11}$$

We can construct a wave equation by taking the curl of equation (2.9) and substituting in equation (2.10). Using the vector identity $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A}$, and assuming a uniform medium, we obtain the familiar Helmholtz wave equation:

$$\nabla^2\mathbf{E} - \mu\epsilon\frac{\partial^2\mathbf{E}}{\partial t^2} = 0 \tag{2.12}$$

A similar equation holds for $\mathbf{H}$. We now restrict ourselves to time harmonic functions ($\mathbf{E} = \mathbf{E}(\mathbf{r})e^{j\omega t}$, $\mathbf{H} = \mathbf{H}(\mathbf{r})e^{j\omega t}$, $\mathbf{J} = \mathbf{J}(\mathbf{r})e^{j\omega t}$). Then Maxwell's equations become:

$$\nabla \cdot \epsilon\mathbf{E} = \rho \tag{2.13}$$

$$\nabla \cdot \mu\mathbf{H} = 0 \tag{2.14}$$

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \tag{2.15}$$

$$\nabla \times \mathbf{H} = j\omega\epsilon\mathbf{E} + \mathbf{J} \tag{2.16}$$

The wave equation becomes

$$\nabla^2\mathbf{E} + \omega^2\mu\epsilon\mathbf{E} = 0 \tag{2.17}$$

With this formalism, we can now consider dispersive media. In general, $\epsilon = \epsilon(\omega)$ and $\mu = \mu(\omega)$. For optics, this means that the index of refraction n is a function of frequency.

2.2 Plane waves [10], [18]

Plane wave solutions can be found by assuming a spatial dependence $\mathbf{E}(\mathbf{r}) = \mathbf{E}_{0+}e^{-j\mathbf{k}\cdot\mathbf{r}}$ and $\mathbf{H}(\mathbf{r}) = \mathbf{H}_{0+}e^{-j\mathbf{k}\cdot\mathbf{r}}$. (The label "0+" indicates that the solution is forward propagating in the direction $\mathbf{k}$; a negative propagating solution also exists with the reverse sign of $\mathbf{k}$. Both are needed for a complete solution to the differential equation.) Here $\mathbf{k}$ is the propagation vector, or wave vector. Its direction is the direction of propagation of the wave, and its magnitude is equal to $2\pi/\lambda$, where λ is the wavelength of the light. By plugging back into the wave equation (2.17), we find that

$$k^2 = \omega^2\mu\epsilon \tag{2.18}$$

This is the dispersion relation for a plane wave. Ignoring magnetic media, we see that in a medium other than free space, k differs from its free space value k_0 by a factor of $\sqrt{\epsilon/\epsilon_0}$, or $k = nk_0$. In other terms, this means that the wavelength of light in a medium is reduced from its free space value by a factor of n.

Equation (2.13) shows that

$$\nabla \cdot \mathbf{E} = -j\mathbf{k} \cdot \mathbf{E} = 0 \tag{2.19}$$

so $\mathbf{k}$ and $\mathbf{E}$ are orthogonal. The direction of $\mathbf{E}$ is defined as the direction of polarization. Information about $\mathbf{H}$ can be found from equation (2.15):

$$\mathbf{H} = \frac{j}{\omega\mu}\nabla \times \mathbf{E} = \frac{j}{\omega\mu}(-j\mathbf{k} \times \mathbf{E}) = \sqrt{\frac{\epsilon}{\mu}}\hat{\mathbf{k}} \times \mathbf{E} \tag{2.20}$$

where $\hat{\mathbf{k}}$ is a unit vector. So the magnetic field is orthogonal to both the electric field and the propagation vector such that $\mathbf{E} \times \mathbf{H}$ is in the propagation direction. Also, $\mathbf{H}$ differs from $\mathbf{E}$ in magnitude by the characteristic admittance $\sqrt{\epsilon/\mu}$.

The phase velocity of light is defined as the velocity of a plane of constant phase. For a plane wave, this plane is $\omega t - \mathbf{k} \cdot \mathbf{r} =$ constant, and the phase velocity is

$$v_p = \frac{\omega}{k} = \frac{1}{\sqrt{\mu\epsilon}} \tag{2.21}$$

For $\mu = \mu_0$ and $\epsilon = \epsilon_0$, $v_p = 1/\sqrt{\mu_0\epsilon_0} = c$, the speed of light in a vacuum. Ignoring magnetic materials, the phase velocity in another medium is $c\sqrt{\epsilon_0/\epsilon} = c/n$.

For the purposes of conveying information with optical signals, we often send a carrier signal which is amplitude modulated with pulse envelopes that represent the data. The group velocity of the signal is defined as the velocity of the envelope and is given by:

$$v_g = \frac{d\omega}{dk} \tag{2.22}$$

The group velocity is the velocity at which information is transmitted. In free space, we have $v_p = v_g$, but in dispersive media with permittivity $\epsilon(\omega)$, we have

$$v_g = \frac{1}{dk/d\omega} = \frac{1}{d(\omega n/c)/d\omega} = \frac{1}{n/c + (\omega/c)dn/d\omega} = \frac{c}{n - \lambda dn/d\lambda} = \frac{c}{n_g} \tag{2.23}$$

n_g is called the group index.

2.3 Reflection and refraction at dielectric interfaces [10], [18]

Consider a plane wave incident on the (infinite planar) interface between two different materials of permittivity ϵ_1 and ϵ_2. The reflection and transmission of the wave is dependent on the boundary conditions at the interface, which from Maxwell's equations (2.1) - (2.4) we find to be

$$\mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) = \sigma = 0 \tag{2.24}$$

$$\mathbf{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) = 0 \tag{2.25}$$

$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 \tag{2.26}$$

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{K} = 0 \tag{2.27}$$

Here $\hat{\mathbf{n}}$ is a unit normal vector pointing from medium 1 to medium 2, σ is the surface charge density, and $\mathbf{K}$ is the surface current density at the interface. We consider only interfaces with no surface charge or current. It turns out that these equations are not independent; the last two will suffice. Thus, we must only consider the fact that tangential $\mathbf{E}$ and $\mathbf{H}$ are continuous across the boundary. The incident, reflected, and transmitted fields are written as

$$\mathbf{E}_{incident} = \mathbf{E}_{+}^{(1)} e^{j(\omega_{+}^{(1)} t - \mathbf{k}_{+}^{(1)} \cdot \mathbf{r})} \tag{2.28}$$

$$\mathbf{E}_{reflected} = \mathbf{E}_{-}^{(1)} e^{j(\omega_{-}^{(1)} t - \mathbf{k}_{-}^{(1)} \cdot \mathbf{r})} \tag{2.29}$$

$$\mathbf{E}_{transmitted} = \mathbf{E}_{+}^{(2)} e^{j(\omega_{+}^{(2)} t - \mathbf{k}_{+}^{(2)} \cdot \mathbf{r})} \tag{2.30}$$

We define the coordinate system so that the plane of the interface has equation $z = 0$. Then in order for the boundary conditions to hold at all points on this plane, we must have

$$\omega_+^{(1)} = \omega_-^{(1)} = \omega_+^{(2)} = \omega \tag{2.31}$$

$$k_{x+}^{(1)} = k_{x-}^{(1)} = k_{x+}^{(2)} = k_x \tag{2.32}$$

$$k_{y+}^{(1)} = k_{y-}^{(1)} = k_{y+}^{(2)} = k_y \tag{2.33}$$

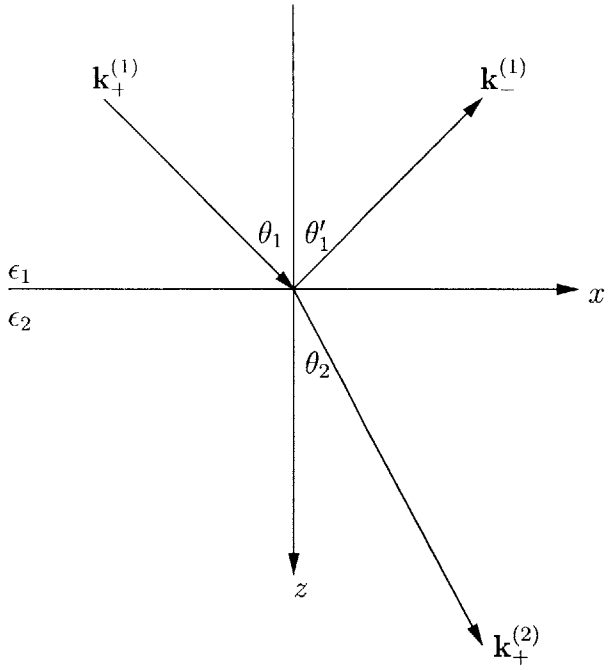

Figure 2-1: Plane wave at a dielectric interface

To simplify the problem, we place the incident, reflected, and transmitted rays into the xz plane and set $k_y = 0$. With the angles θ_1, θ_1', and θ_2 defined as in Figure 2-1, the condition on k_x implies that $\theta_1 = \theta_1'$, i.e., that the angles of incidence and reflection are equal. It also implies the well-known Snell's law for the angles of incidence and transmission:

$$k^{(1)} \sin\theta_1 = k^{(2)} \sin\theta_2 \tag{2.34}$$

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{2.35}$$

where we assume $\mu = \mu_0$. The next step in solving the interface problem is to apply the boundary conditions at the plane $z = 0$ to solve for the magnitudes of the fields. We consider only the transverse electric (TE) polarization, in which the electric field is perpendicular to the plane of incidence (here, the xz plane). The Fresnel formulas which result are:

$$E_-^{(1)} = \Gamma E_+^{(1)} \tag{2.36}$$

$$E_+^{(2)} = T E_+^{(1)} \tag{2.37}$$

where the reflection coefficient Γ and the transmission coefficient T are defined as

$$\Gamma = \frac{Z_0^{(2)} - Z_0^{(1)}}{Z_0^{(2)} + Z_0^{(1)}} \tag{2.38}$$

$$T = \frac{2Z_0^{(2)}}{Z_0^{(2)} + Z_0^{(1)}} \tag{2.39}$$

and the characteristic impedance in a medium is defined as

$$Z_0^{(i)} = \frac{\omega \mu_i}{k_z^{(i)}} \tag{2.40}$$

This definition relates reflection at dielectric interfaces to reflection at transmission line interfaces. Just as in the transmission line case, impedance matching leads to $\Gamma = 0$ and $T = 1$.

2.3.1 Total internal reflection

A special case of this problem, and one which makes the waveguide possible, is the phenomenon of total internal reflection. If $\sin\theta_1 > n_2/n_1$, then Snell's law (2.35) shows that θ_2 becomes imaginary. In other words, while k_x must remain constant across the interface, $k_x^{(2)} > k^{(2)}$, where $k^{(2)2} = k_x^{(2)2} + k_z^{(2)2}$. In order for this to hold,

$$k_z^{(2)} = -j\alpha_z^{(2)} \tag{2.41}$$

The field exponentially decays into medium 2 in the z direction. Furthermore, no power is transmitted into medium 2. Hence the wave is "totally reflected." Using the Fresnel relation (2.36) with this definition of $k_z^{(2)}$ gives

$$\Gamma = \omega\mu_0 \frac{\frac{j}{\alpha_z^{(2)}} - \frac{1}{k_z^{(1)}}}{\frac{j}{\alpha_z^{(2)}} + \frac{1}{k_z^{(1)}}} = \omega\mu_0 \frac{k_z^{(1)} + j\alpha_z^{(2)}}{k_z^{(1)} - j\alpha_z^{(2)}} \tag{2.42}$$

So the incident and reflected waves have the same magnitude, but there is a phase difference:

$$E_-^{(1)} = e^{-j2\phi} E_+^{(1)} \tag{2.43}$$

where

$$\tan\phi = -\frac{\sqrt{\sin^2\theta - n_2^2/n_1^2}}{\cos\theta} \tag{2.44}$$

This phase shift is known as the Goos-Hänchen shift.

2.4 Slab waveguides

We shall consider only the simplest form of waveguide in this thesis, the two dimensional slab waveguide. In reality, all the waveguides in the final product will be buried channel waveguides. However, this 3D structure can be reduced to a 2D structure via the effective index method, so we need only study the 2D solution [18]. (In this thesis, we actually did not use the effective index method to reduce the structure to two dimensions; instead, we just assumed it already was two dimensional. This choice undoubtedly causes errors.)

Generally, a 2D slab waveguide consists of a higher index core material (index n_1) sandwiched between cladding materials of indices n_2 and n_3, where, for sake of definition, $n_1 > n_2 > n_3$. In this thesis, however, we only consider symmetric waveguides with $n_2 = n_3$. We define our axes such that z is the direction of propagation, parallel to the dielectric interfaces, x is the direction in which the index changes, and y is orthogonal to x and z, as shown in Figure 2-2. We will analyze the slab waveguide in two ways, making use of ray optics and wave optics principles.

2.4.1 Ray optics derivation of modes [22], [25]

A waveguide operates on the principle of total internal reflection, described above. When light is coupled into the waveguide so that the angle of incidence is greater than the critical angle for total internal reflection, the light will stay confined within the waveguide. Previously, we considered θ, the angle between the light ray and the normal to the interface, but in this section, we will consider α, the angle between the

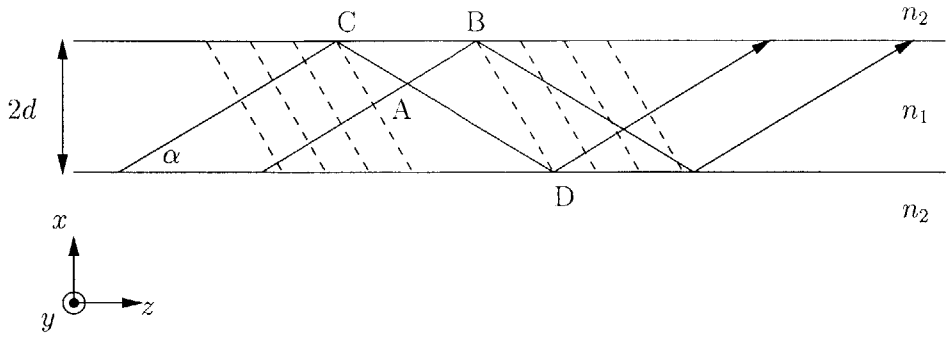

Figure 2-2: Plane wave propagating in a slab waveguide, with dashed lines representing phase fronts

light ray and the interface itself. So the critical angle which α must be less than is $\cos^{-1} n_2/n_1$.

As the ray travels in the waveguide, it has propagation vector of magnitude $k = n_1 k_0$, where k_0 is the magnitude in free space. This vector can be decomposed into two components along the x and z axes:

$$\kappa = k_x = n_1 k_0 \sin\alpha \tag{2.45}$$

$$\beta = k_z = n_1 k_0 \cos\alpha \tag{2.46}$$

Since the mode propagates along the z direction, β is commonly referred to as the mode propagation constant. Surprisingly, it turns out that not all rays with an angle less than the critical angle can propagate in the waveguide. In fact, only a discrete set is allowed. These are the waveguide modes. We solve for the modes by considering points on the same phase fronts. In Figure 2-2, points A and C are on the same phase front, and points B and D are, respectively, the same points some time later. Since they lie on the same phase front, the phase accumulated from A to B and from C to D must be the same. The distances from A to B and C to D can be found to be [22]:

$$s_{AB} = \left(\frac{1}{\tan\alpha} - \tan\alpha\right) 2d\cos\alpha = (\cos^2\alpha - \sin^2\alpha)\frac{2d}{\sin\alpha} \tag{2.47}$$

$$s_{CD} = \frac{2d}{\sin\alpha} \tag{2.48}$$

where d is the half width of the waveguide. To set the phases equal, we note that the ray from C to D encounters two total internal reflections. So the total equation for phase matching is

$$n_1 k_0 s_{CD} + 2\phi + 2\phi = n_1 k_0 s_{AB} + 2m\pi \tag{2.49}$$

where ϕ is the Goos-Hänchen angle defined earlier. We can also express it in terms of propagation constants and indices of refraction:

$$\tan\phi = -\frac{\sqrt{\cos^2\alpha - n_2^2/n_1^2}}{\sin\alpha} = -\frac{\sqrt{\beta^2 - n_2^2 k_0^2}}{\kappa} = -\frac{\gamma}{\kappa} \tag{2.50}$$

where

$$\gamma = \sqrt{\beta^2 - n_2^2 k_0^2} \tag{2.51}$$

is an abbreviation which will have more meaning later.

Thus the eigenvalue equation for the acceptable modes of the waveguide turns out to be

$$\tan\left(\kappa d - \frac{m\pi}{2}\right) = \frac{\gamma}{\kappa} \tag{2.52}$$

This equation is also called the dispersion relation for the TE modes of the slab waveguide. Each value of m produces one mode. $m = 0$ is the fundamental mode with the smallest possible value of α, while modes with higher m propagate at larger angles.

2.4.2 Wave optics derivation of modes [10], [18], [22], [25]

We can also solve the slab waveguide problem by solving a wave equation. Unlike the plane wave case, however, the slab waveguide has a varying permittivity, $\epsilon(x)$. We proceed by looking separately at the wave equation in each dielectric and then matching boundary conditions at the interfaces. The fields have the general form:

$$\mathbf{E} = \mathbf{E}_0(x, y)e^{j(\omega t - \beta z)} \tag{2.53}$$

$$\mathbf{H} = \mathbf{H}_0(x, y)e^{j(\omega t - \beta z)} \tag{2.54}$$

Note that propagation is in the z direction with the propagation constant β, described earlier as the z component of the wavevector in the core. In addition, a key property of a waveguide is that the field's shape remains constant as it propagates down the waveguide, so the fields are written without any z dependence.

We can use Maxwell's equations (2.15) and (2.16) to obtain the following equations for the field components in the medium (assuming the forms above):

$$\frac{\partial E_z}{\partial y} + j\beta E_y = -j\omega\mu_0 H_x \tag{2.55}$$

$$-j\beta E_x - \frac{\partial E_z}{\partial x} = -j\omega\mu_0 H_y \tag{2.56}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -j\omega\mu_0 H_z \tag{2.57}$$

$$\frac{\partial H_z}{\partial y} + j\beta H_y = j\omega\epsilon E_x \tag{2.58}$$

$$-j\beta H_x - \frac{\partial H_z}{\partial x} = j\omega\epsilon E_y \tag{2.59}$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = j\omega\epsilon E_z \tag{2.60}$$

For the slab guide, we assume that the fields have no y dependence, as the slab is considered to continue infinitely in that direction. To find a wave equation, we must choose the polarization. For transverse electric (TE) waves, $\mathbf{E}_0$ is in the y direction, and $\mathbf{H}_0$ is in the xz plane. For transverse magnetic (TM) waves, $\mathbf{H}_0$ is in the y direction, and $\mathbf{E}_0$ is in the xz plane. As before, we shall consider only the TE solution, but the TM solution is similar.

The remaining Maxwell's equations for a TE polarized wave are:

$$j\beta E_y = -j\omega\mu_0 H_x \tag{2.61}$$

$$\frac{\partial E_y}{\partial x} = -j\omega\mu_0 H_z \tag{2.62}$$

$$-j\beta H_x - \frac{\partial H_z}{\partial x} = j\omega\epsilon E_y \tag{2.63}$$

We can then find a TE wave equation:

$$\frac{d^2E_y}{dx^2} = -j\omega\mu_0\frac{dH_z}{dx} \tag{2.64}$$

$$= -j\omega\mu_0(-j\beta H_x - j\omega\epsilon E_y) \tag{2.65}$$

$$= -\omega\mu_0\beta(-\frac{\beta}{\omega\mu_0}E_y) - \omega^2\mu\epsilon E_y \tag{2.66}$$

$$= (\beta^2 - n^2k_0^2)E_y \tag{2.67}$$

$$\frac{d^2E_y}{dx^2} + (n^2k_0^2 - \beta^2)E_y = 0 \tag{2.68}$$

We will now solve for the electric field in each of the three regions of the slab. The magnetic field components can then be found by plugging into (2.61) and (2.62). The general solutions to this equation are:

$$E_y(x) = E_0e^{\pm\sqrt{\beta^2 - n^2k_0^2}x} \tag{2.69}$$

If $\beta^2 > n^2k_0^2$, the solutions are exponential, while if $\beta^2 < n^2k_0^2$, the solutions are oscillatory. For a guided mode, the field is exponentially decaying in the cladding and oscillatory in the core. The modes look like:

$$E_y(x) = \begin{cases} Ae^{\gamma(x+d)} & , x < -d \\ B\cos(\kappa x - \phi) & , -d < x < d \\ Ce^{-\gamma(x-d)} & , x > d \end{cases} \tag{2.70}$$

Here d is the half width of the waveguide; A, B, C, and ϕ are unknown constants; and

$$\gamma = \sqrt{\beta^2 - n_2^2k_0^2} \tag{2.71}$$

$$\kappa = \sqrt{n_1^2k_0^2 - \beta^2} \tag{2.72}$$

We have seen κ before in the ray optics solution as the x component of the wavevector in the medium. Now we see that γ is actually the coefficient of exponential decay in the cladding.

The boundary conditions (2.26) and (2.27) imply that tangential **E** and **H** should be continuous at the boundary between media. For E_y, this implies that $A = B\cos(-\kappa d - \phi)$ and $C = B\cos(\kappa d - \phi)$. The other boundary condition is for H_z:

$$H_z = \frac{j}{\omega\mu_0}\frac{\partial E_y}{\partial x} \tag{2.73}$$

$$= \begin{cases} B\cos(-\kappa d - \phi)\frac{j\gamma}{\omega\mu_0}e^{\gamma(x+d)} & , x < -d \\ -B\frac{j\kappa}{\omega\mu_0}\sin(\kappa x - \phi) & , -d < x < d \\ -B\cos(\kappa d - \phi)\frac{j\gamma}{\omega\mu_0}e^{-\gamma(x-d)} & , x > d \end{cases} \tag{2.74}$$

Matching at $x = \pm d$ gives

$$\cos(-\kappa d - \phi)\gamma = -\kappa\sin(-\kappa d - \phi) \tag{2.75}$$

$$\cos(\kappa d - \phi)\gamma = \kappa\sin(\kappa d - \phi) \tag{2.76}$$

The solutions are $\phi = m\pi/2$. For m even,

$$\tan\kappa d = \frac{\gamma}{\kappa} \tag{2.77}$$

This corresponds to m even in equation (2.52). These modes are cosinusoidal and symmetric. For m odd,

$$\tan \kappa d = -\frac{\kappa}{\gamma} \tag{2.78}$$

This corresponds to m odd in equation (2.52). These modes are sinusoidal and antisymmetric.

The first and only mode of a single mode, high index contrast waveguide is pictured in Figure 2-3.

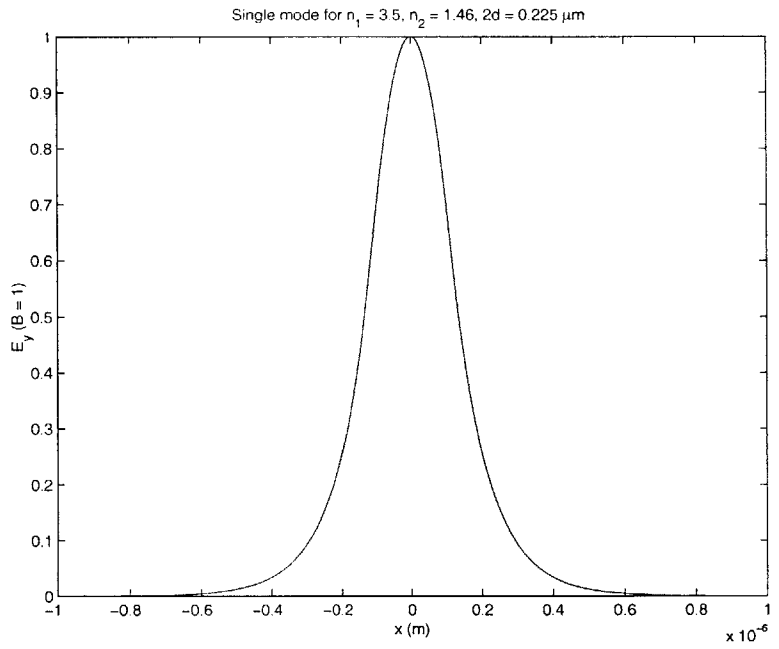

Figure 2-3: Mode of a single mode, high index contrast waveguide

2.4.3 Properties of modes [10], [18], [22], [25]

The propagation constants for the modes of a waveguide can be found by plotting the function $\tan \kappa d$ on the same axes as γ/κ and $-\kappa/\gamma$ and seeing where they intersect. These points are the eigenvalues for κ for a given value of d. From equation (2.72), the propagation constant β can be found. The phase velocity of the mode is ω/β,

and the group velocity is $d\omega/d\beta$. We also define the effective index n_{eff} of a mode as

$$n_{eff} = \frac{\beta}{k_0} \tag{2.79}$$

As the effective index increases, β becomes larger relative to κ, and the mode travels at a smaller angle α. It is clear that the effective index can be no larger than the core index, n_1. In the other direction, mode cutoff occurs when $n_{eff} = n_2$, or $\beta = n_2 k_0$, at which point $\gamma = 0$. When γ becomes imaginary, the field will radiate, not decay, in the cladding. The condition turns out to be equivalent to the cutoff for total internal reflection. From the dispersion relations (2.77) and (2.78), we see that at cutoff,

$$\tan \kappa d = 0 \tag{2.80}$$

$$\tan \kappa d = \infty \tag{2.81}$$

which correspond to the condition $\kappa d = m\pi/2$ at cutoff. Substituting in for κ at cutoff, we get the condition

$$V = k_0 d\sqrt{n_1^2 - n_2^2} = \frac{m\pi}{2} \tag{2.82}$$

V is called the normalized frequency. Calculating the normalized frequency for a particular design shows how many modes there are. Note that the cutoff for the fundamental mode is $V = 0$, so that the fundamental mode always exists in a symmetric waveguide. We can also see some trends in the number of modes. As k_0 increases, and thus ω increases, the number of modes increases. As d increases, the number of modes increases. Finally, as the index contrast increases, the number of modes

increases.

An alternative viewpoint is that the single mode width for a high index contrast waveguide is smaller than the single mode width for a low index contrast waveguide. It can be shown that the product κd is the same for waveguides of all index contrasts when d is chosen so that the second mode is just cut off ($V = \pi/2$). Thus the basic shape of the first mode at this point is the same for all index contrasts. However, for a high index contrast waveguide, the critical value of d is small, and the mode is located within a smaller distance. The exponential decay coefficient γ is necessarily larger to match the slope at the core/cladding boundary, and thus the field decays over a much smaller distance than for a low index contrast waveguide. The fact that high index contrast waveguides provide better confinement allows us to use them as a solution to the bend loss problem.

Another interesting property of waveguide modes is their orthogonality and completeness. Orthogonality means that the following relation holds:

$$\int (\mathbf{E}_m \times \mathbf{H}_\ell^*) \cdot d\mathbf{A} = 0, m \neq \ell \tag{2.83}$$

Here the integral is taken over the cross section of the waveguide. This is shown to be a consequence of power conservation in [18]. Completeness implies that any arbitrary field distribution can be expressed as a superposition of weighted modes of the waveguide. However, completeness requires another set of modes known as the radiation modes. These modes correspond to $\beta < n_2 k_0$, which have imaginary γ. Oscillation occurs in the cladding, and the mode is not confined. Radiation modes are rarely of interest to us; they are not useful to excite in our waveguides, since they do not actually guide the field! However, guided modes can sometimes couple into radiation modes, as when a guided mode traverses a bend. It may then be important to think of the field as a superposition of both types of modes.

2.5 Bend loss

When light passes through a bent waveguide, loss inevitably occurs due to radiation. Figure 2-4 shows the planes of constant phase around a bend. Clearly, the tangential speed of the mode must increase with the radius from the center of the bend. Eventually, at a radius known as the radiation caustic, the required tangential speed exceeds the speed of light in the cladding, c/n_2. Any field beyond the caustic is thus coupled to radiation modes, which move at c/n_2, and the power is lost. Since all slab waveguide modes have field extending off to infinity, radiation loss is unavoidable when the waveguide reaches a bend.

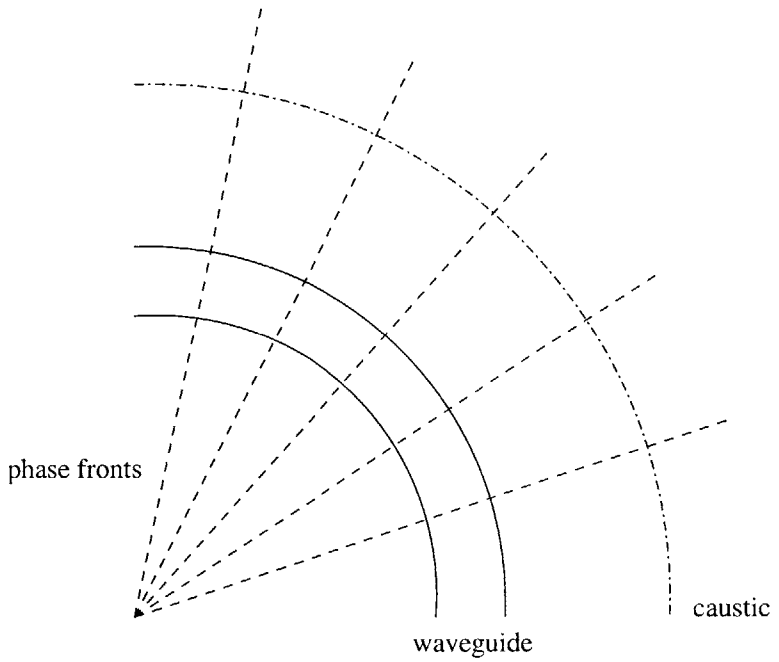

Figure 2-4: Waveguide bend, showing planes of constant phase and radiation caustic

Marcuse [21] derived the following relation for the exponential attenuation coefficient of the power over a distance z ($I = I_0 \exp(-\alpha z)$):

$$\alpha = \frac{\gamma^2\kappa^2}{\beta(1+\gamma d)(n_1^2-n_2^2)k_0^2} e^{2\gamma d} e^{-2(\beta \tanh^{-1}(\gamma/\beta)-\gamma)R} \tag{2.84}$$

This equation clearly shows the exponential dependence of the attenuation on R. For fixed waveguide structure, the bend loss goes up exponentially with decreasing R, thus making it difficult to create integrated optics structures with small bends.

The solution, as mentioned in Chapter 1, is high index contrast. As we saw earlier, high index contrast modes are more confined in space. Thus there is less overlap with the radiation modes outside the caustic, decreasing the loss. Figure 2-5

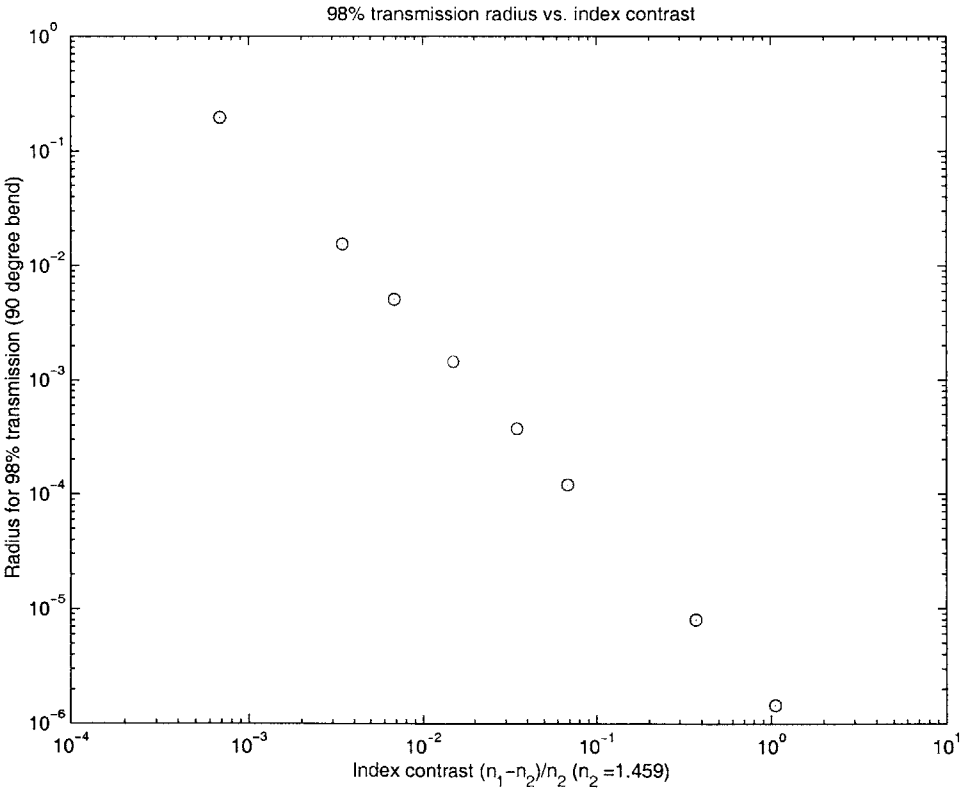

Figure 2-5: Radius required for 98% transmission through a 90 degree bend vs. index contrast [31]

plots the radius required for 98 percent transmission through a 90 degree bend versus the index contrast. The width of each waveguide is the critical one where the second mode is cut off for that particular index contrast [31].

2.6 Finite difference time domain (FDTD) methods

There exist many different methods to simulate the performance of an optical device, each with its own strengths and weaknesses. One popular method is the beam propagation method (BPM), which relies on a paraxial approximation and does not take reflections into account very well. Because of these restrictions, BPM turns out to not be very useful for high index contrast devices, which feature sharply inclined k vectors and many reflections.

Other methods are exact. Most of these methods use some sort of spatial discretization, in which the structure under study is divided into a grid. The one which we used for this project is the so-called finite difference time domain (FDTD) method, first developed by K.S. Yee in 1966 [44]. The "finite difference" part of the name implies that the derivatives in Maxwell's equations (2.1) - (2.4) are approximated as finite differences, while "time domain" indicates that the time variable is discretized in the same way as the spatial variables. The fact that FDTD operates in the time domain allows for many frequencies to be studied by sending in an input pulse of sufficient frequency content [4]; we will make use of this fact in our simulations.

2.6.1 Principles of FDTD

The spatial and time discretization of FDTD is well explained by W.C. Chew's theory of electromagnetism on a discrete lattice [6]. In a discrete space, derivatives can be replaced by finite differences, either forward or backward. For example, if $g(x) = \frac{\partial}{\partial x} f(x)$, the analog in discrete space is

$$g_{m+1/2} = \tilde{\partial}_x f_m = \frac{1}{\Delta_x}(f_{m+1} - f_m) \tag{2.85}$$

or

$$g_{m-1/2} = \hat{\partial}_x f_m = \frac{1}{\Delta_x}(f_m - f_{m-1}) \tag{2.86}$$

Here Δ_x is the discretization size in the x direction, and $f_m = f(m\Delta_x)$. Note that $g_{m+1/2}$ and $g_{m-1/2}$ are defined at half grid points. Finally, the two different difference operators are distinguished by a tilde on the forward difference, $\tilde{\partial}$, and a hat on the backward difference, $\hat{\partial}$.

This theory can be extended to vector calculus, including the gradient, divergence, curl, Gauss' theorem, and Stokes' theorem. With these tools, a discrete form of Maxwell's equations (2.1) - (2.4) can be expressed as:

$$\tilde{\nabla} \times \tilde{\mathbf{E}}^{\ell}_{\mathbf{m}} = -\tilde{\partial}_t \hat{\mathbf{B}}^{\ell-1/2}_{\mathbf{m}+1/2} \tag{2.87}$$

$$\hat{\nabla} \times \hat{\mathbf{H}}^{\ell-1/2}_{\mathbf{m}+1/2} = \hat{\partial}_t \tilde{\mathbf{D}}^{\ell}_{\mathbf{m}} + \tilde{\mathbf{J}}^{\ell-1/2}_{\mathbf{m}} \tag{2.88}$$

$$\tilde{\nabla} \cdot \hat{\mathbf{B}}^{\ell-1/2}_{\mathbf{m}+1/2} = 0 \tag{2.89}$$

$$\hat{\nabla} \cdot \tilde{\mathbf{D}}^{\ell}_{\mathbf{m}} = \rho^{\ell}_{\mathbf{m}} \tag{2.90}$$

Here the subscripts represent the spatial variable, where $\mathbf{m} + 1/2 = (m + 1/2, n + 1/2, p + 1/2)$. The superscripts represent the time variable, which is discretized as $t = \ell\Delta t$. Finally, the tildes and hats on the vectors denote whether they are so-called "forevectors" or "backvectors," respectively. A forevector associated with the point $\mathbf{m}$ (as a subscript) actually has its x component defined at the point $(m + 1/2, n, p)$,

its y component defined at the point $(m, n+1/2, p)$, and its z component defined at the point $(m, n, p+1/2)$. For example,

$$\tilde{\mathbf{E}}_{\mathbf{m}} = E^x_{m+1/2,n,p}\hat{x} + E^y_{m,n+1/2,p}\hat{y} + E^z_{m,n,p+1/2}\hat{z} \tag{2.91}$$

Similarly, a back vector associated with the point $\mathbf{m}$ actually has its x component defined at the point $(m-1/2, n, p)$, its y component defined at the point $(m, n-1/2, p)$, and its z component defined at the point $(m, n, p-1/2)$. The last thing we need are the relations between $\tilde{\mathbf{E}}$, $\tilde{\mathbf{D}}$, $\hat{\mathbf{H}}$, and $\hat{\mathbf{B}}$:

$$\hat{\mathbf{B}}_{\mathbf{m}+1/2} = \mu_{\mathbf{m}+1/2} \cdot \hat{\mathbf{H}}_{\mathbf{m}+1/2} \tag{2.92}$$

$$\tilde{\mathbf{D}}_{\mathbf{m}} = \epsilon_{\mathbf{m}} \cdot \tilde{\mathbf{E}}_{\mathbf{m}} \tag{2.93}$$

where μ is the permeability tensor and ϵ is the permittivity tensor. There is now enough information to solve for the fields numerically; this is the basis of the FDTD algorithm.

One important result found in [6] is the fact that the propagation constant of a discrete plane wave is not the same as that of the equivalent continuous plane wave. Similarly, the propagation constant β for a mode in FDTD is not quite the same as the propagation constant found by a modesolver. Specifically,

$$\beta^2_{real} = \beta^2_{FDTD} \frac{\sin^2(\beta_{FDTD}\Delta z/2)}{(\beta_{FDTD}\Delta z/2)^2} \tag{2.94}$$

Here Δz is the discretization in the direction of propagation. This correction, though slight, turns out to be critically important when attempting to calculate phases in the AWG.

Another important issue for FDTD simulations is the choice of a boundary condition. Since the computational domain cannot be infinitely large, and, in fact, should be as small as possible, we need some sort of boundary condition on the outside of this domain to ensure the fields are not reflected back like at metal walls. Various boundary conditions exist, but most are imperfect in that they only suppress reflections which meet certain criteria, such as normal incidence. This deficiency limits the structures which can be simulated or forces the use of a large computational domain to keep the interesting elements away from the boundary. The problem was solved by the development of the perfectly matched layer (PML) [3]. The PML perfectly absorbs, without reflection, all waves traveling towards the computational domain boundaries, regardless of frequency or angle of incidence. PML is therefore the most common boundary condition used today; indeed, the FDTD code used in this thesis features a PML boundary condition.

2.6.2 FDTD in this thesis

The FDTD code used in this thesis was developed by C. Manolatou [19] in Fortran. Some supplementary files were also necessary to change the output dielectric profiles and field profiles to graphic files, as well as to use the discrete Fourier transform output to estimate a mode overlap. Finally, we made some modifications to the code in order to easily test the AWG. These modifications are discussed in Chapter 3.

The input files to the FDTD code consist of sequences of numbers which define the dielectric structures, sources, outputs, and much more. A sample input file is given in Appendix A. Some of the highlights are described below.

Discretization The discretization should be chosen to be 10-20 pixels per wavelength in the chosen material. Since silicon has $n = 3.5$, the center wavelength of the WDM system, 1.55 μm in free space, is actually about 0.443 μm in the waveguide. Therefore, we should have pixels which are, at the most, 0.0443 μm across. Usually 0.025 μm spacing was used. Earlier in the project, 0.05 μm was used, until it was determined that the results were inaccurate. Unfortunately, cutting the spacing in half

makes the simulation eight times slower, four times for the two spatial dimensions, and two times more for the time step, which is calculated based on the grid spacing.

Sources The first important parameter we must choose for the sources is the time type. An input of 0, or "no time type," is used to define dummy sources to be used in mode overlap calculations. The input 1 gives a continuous wave (CW) source, which we used for testing sometimes. The input 2 gives a Gaussian pulse around a carrier. We used this for most of our real sources, in order to easily probe the frequency response of a device. The carrier wavelength was chosen to be 1.55 μm, and the pulse width was usually chosen to give a broad frequency spectrum.

The other important parameter we need to decide is the spatial structure of the excitation across a cross section. Usually, we used a modesolver and fit the excitation to the mode of the input waveguide. However, we later made modifications to allow spatial distributions from previous simulations to be input as well. These modifications will be discussed in Chapter 3.

Mode overlap Mode overlaps can be calculated by placing mode overlap cross sections across an output waveguide and slightly ahead of the input source. A dummy source is also required on the output waveguide. The program "proc" then uses the spatial profiles of the (real and dummy) sources and the Fourier output at the overlap cross sections to calculate the power reflection and transmission coefficients.

Power flux Power fluxes can also be obtained at arbitrary cross sections using Fourier methods. Power fluxes are less accurate than mode overlaps, since some of the power may not match the mode of the output waveguide, causing it to eventually be lost. However, while mode overlaps only operate at a single wavelength, power fluxes give results for a range of wavelengths, making them useful for analyzing the wavelength dependence of a structure.

The FDTD code used in this thesis gave accurate results, but at a price. When the simulations became too large for a standard desktop machine, we had to continue the simulations on a supercomputer. Unfortunately, the quirks of the supercomputer caused the existing code not to work due to a problem with the precision of certain variables. The so-called "hybrid precision" code that was originally provided failed, so we had to use double precision throughout, increasing accuracy but slowing the code.

Hopefully, other users of this FDTD code will one day standardize it and improve its speed somehow. Another attractive possibility is the use of a different simulation method based on expansion of the field into the eigenmodes of the structure, whose propagation is well known. An example of this method is P. Bienstman's software CAMFR [4]. Such a method is perfect for the MMIs described in Chapter 3, where the index profile is easily divided into three sections: the inputs, the multimode section, and the outputs. Unfortunately, there was not enough time to explore the potential of this technique.

Chapter 3

Multimode Interference Coupler AWG

The traditional design for an arrayed waveguide grating makes use of free space couplers, also known as star couplers or focusing slab regions, as the input splitting and output combining elements. In the input, a single waveguide enters a free space region, causing the waveguide mode to diffract outward. The far field, which is the Fourier transform of the input, is sampled by the array waveguides. The array adds a wavelength-dependent phase, which is linear across the array. At the second free space region, the light is focused to an output waveguide. For the center wavelength, which has no extra phase from the array, the focal point will be in the center output waveguide, thus reversing the effects of the input coupler. For all other wavelengths, the wavelength-dependent linear phase shift will focus the light to a different output waveguide. This structure is easily thought of in terms of more conventional optical components, including a lens (first coupler), mask (sampling of the array), prism (dispersion of the array), and a second lens (second coupler). Note that the sampling of the array creates multiple diffraction orders in the output, an unwanted effect which we will discuss later [18].

This basic AWG structure is the one proposed by Smit [33], Takahashi [40], and Dragone [8] and is the model used by most researchers and manufacturers of AWGs today. It is also the type of AWG pictured in Figure 1-1. However, several researchers

([16], [27], [28], [32], [24]) have proposed making AWG devices using different structures for the input splitting and output combining elements. Specifically, they suggest using a multimode interference coupler (MMI), which consists of input waveguides, a large multimode section, and output waveguides. A typical MMI is shown in Figure 3-1. The MMI has many advantages, including large bandwidth, low polarization sensitivity, low loss, and most importantly for this project, compact size, simple design procedures, and relaxed fabrication tolerances [16], [27]. Because of these advantages, we decided to design the AWG using MMI couplers for the splitting and combining stages.

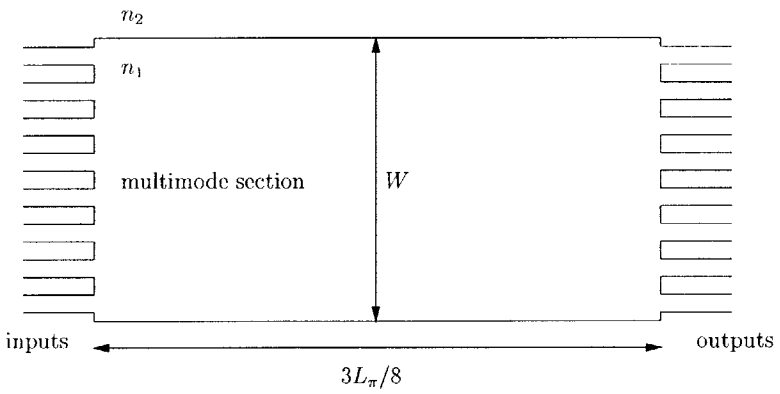

Figure 3-1: Structure of 8 × 8 multimode interference coupler

In this chapter, we will first present the basic principles of multimode interference couplers, including the primary equations used in their design. Then we will discuss our efforts to design MMI couplers for the given material system and desired size. Next, we will present the design of the array waveguides in the AWG. Unlike the traditional AWG, where the array waveguides have a constant spacing ΔL in order to create a linear phase across the array, the array waveguides in an MMI AWG are not simple to design. The MMI couplers have phase transfer relations which must be taken into account when designing the array waveguide lengths for proper phase matching and interference. We discuss a technique for choosing these lengths. We then describe some slight modifications to the FDTD code which enabled us to

effectively simulate the entire AWG structure. Finally, we discuss the results and their implications.

3.1 Principles of multimode interference couplers [35], [41], [1]

Multimode interference couplers operate on the principle of self-imaging, which is defined by Soldano and Pennings [35] as "a property of multimode waveguides by which an input field profile is reproduced in single or multiple images at periodic intervals along the propagation direction of the guide." Self-imaging in multimode waveguides was first investigated by Bryngdahl [5] and studied further by Ulrich and Ankele [42]. To describe the operation of an MMI, we use a modal decomposition technique.

3.1.1 Modal decomposition of input field

Consider a multimode waveguide of width W, core index n_1, and cladding index n_2. The guide supports m modes, numbered $i = 0, 1, ..., m-1$, where m is determined by calculating the normalized frequency V from (2.82). We know from (2.72) that

$$\kappa_i^2 + \beta_i^2 = n_1^2 k_0^2 \tag{3.1}$$

We can approximate the value of κ_i by assuming that the field goes completely to zero at some point instead of exponentially decaying:

$$\kappa_i = \frac{(i+1)\pi}{W_{ei}} \tag{3.2}$$

The effective width W_{ei} takes into account the penetration depth of the field into the cladding:

$$W_{ei} \approx W_e = W + \frac{\lambda_0}{\pi}(n_1^2 - n_2^2)^{-1/2} \tag{3.3}$$

For high index contrast waveguides like the ones we use, $W_e \approx W$. We can now expand (3.1) as:

$$\begin{aligned} \beta_i &= n_1 k_0 \sqrt{1 - \frac{\kappa^2}{n_1^2 k_0^2}} && (3.4) \\ &\approx n_1 k_0 \left(1 - \frac{1}{2}\frac{(i+1)^2\pi^2}{W_e^2 n_1^2 k_0^2}\right) && (3.5) \\ &= n_1 k_0 - \frac{(i+1)^2 \pi \lambda_0}{4 n_1 W_e^2} && (3.6) \end{aligned}$$

This approximate quadratic dependence of β on the mode number is critical to the derivation of the self-imaging property. A key parameter is the beat length L_π:

$$L_\pi = \frac{\pi}{\beta_0 - \beta_1} = \frac{4 n_1 W_e^2}{3\lambda_0} \tag{3.7}$$

Some more algebra shows the propagation constants are spaced by

$$\beta_0 - \beta_i = \frac{i(i+2)\pi}{3L_\pi} \tag{3.8}$$

Suppose we excite the MMI with an input field profile $E(x, 0)$ at $z = 0$. From the

principle of completeness, we can express the profile as a sum of all the modes of the multimode section:

$$E(x,0) = \sum_i c_i E_i(x) \tag{3.9}$$

Here, the E_i include the radiation modes as well as the guided modes, but we can assume that the radiation modes are not excited for reasonable input fields. Using the principle of orthogonality (2.83), we can express the expansion coefficients by an overlap integral:

$$c_i = \frac{\int E(x,0)E_i(x)dx}{\sqrt{\int E_i^2(x)dx}} \tag{3.10}$$

With this decomposition, we can write the field at some z as

$$E(x,z) = \sum_{i=0}^{m-1} c_i E_i(x) e^{j(\omega t - \beta_i z)} \tag{3.11}$$

$$= e^{j(\omega t - \beta_0 z)} \sum_{i=0}^{m-1} c_i E_i(x) e^{j(\beta_0 - \beta_i)z} \tag{3.12}$$

$$= e^{j(\omega t - \beta_0 z)} \sum_{i=0}^{m-1} c_i E_i(x) \exp\left(j\frac{i(i+2)\pi}{3L_\pi}z\right) \tag{3.13}$$

3.1.2 Single images

Note that if the phase factor $\exp\left(j\frac{i(i+2)\pi}{3L_\pi}z\right) = 1$, the initial field profile will be reproduced at z. To find when this occurs, note that $i(i+2)$ is even if i is even and odd if i is odd. To ensure the condition, therefore, we must have $z = p(3L_\pi)$, where p is even.

Another self image occurs if $\exp\left(j\frac{i(i+2)\pi}{3L_\pi}z\right) = (-1)^i$. In this case, the odd modes are 180 degrees out of phase, so the field will be reproduced on the opposite side of the waveguide, a mirror image of the original. Some consideration shows that this condition occurs for $z = p(3L_\pi)$, where p is odd. Thus, the condition for a single self image is

$$z = p(3L_\pi), p = 0, 1, 2, ... \tag{3.14}$$

3.1.3 Multiple images

Consider the field at distances $z = p(3L_\pi)/2$, $p = 1, 3, 5, ...$ halfway in between the single images. The field is given by

$$\begin{aligned} E(x, p(3L_\pi)/2) &= \sum_{i=0}^{m-1} c_i E_i(x) \exp\left(j\frac{i(i+2)p\pi}{2}\right) && (3.15) \\ &= \sum_{i=0,2,4,...} c_i E_i(x)(1)^p + \sum_{i=1,3,5,...} c_i E_i(x)(-j)^p && (3.16) \\ &= \frac{1+(-j)^p}{2}E(x,0) + \frac{1-(-j)^p}{2}E(-x,0) && (3.17) \end{aligned}$$

So there are two images with equal amplitude which are 90 degrees out of phase. Bachmann, Besse, and Melchior extend this analysis even further [1] to show that N images are formed at distances

$$z = \frac{p}{N}(3L_\pi) \tag{3.18}$$

Normally, we choose $p = 1$ to minimize the length of our couplers.

Although the arguments above are general for any input distribution, we usually

only care about inputs arising from N input waveguides, which then couple to N output waveguides. If we number the inputs $i = 1, ..., N$ from the bottom up and the outputs $j = 1, ..., N$ from the top down, we find the following phase relations for imaging from input i to output j (excluding a constant phase) [1]:

$$\phi_{ij} = \begin{cases} \pi + \frac{\pi}{4N}(j-i)(2N-j+i) & , i+j \text{ even} \\ \frac{\pi}{4N}(j+i-1)(2N-j-i+1) & , i+j \text{ odd} \end{cases} \tag{3.19}$$

3.1.4 Symmetric excitation of MMI coupler

The arguments above apply to MMI couplers with arbitrary input field profiles which can excite all of the modes in the multimode section. Consider now a symmetric input profile which only excites modes with even i. Then the phase factor $\exp\left(j\frac{i(i+2)\pi}{3L_\pi}z\right)$ will still give an output of 1 for lengths

$$z = p\left(\frac{3L_\pi}{4}\right), p = 0, 1, 2, ... \tag{3.20}$$

This property extends to N images. Thus MMI couplers can be made four times shorter if only the even modes are excited. We cannot use this fact to alter the 8×8 combiner at the end of the MMI. However, instead of making an 8×8 splitter, with only one of the inputs used, we can in fact make a shorter 1×8 splitter with the input waveguide centered on the multimode slab.

The phases of the images (excluding a constant phase) are given by [2]:

$$\phi_j = \frac{\pi}{N}(j-1)(N-j) \tag{3.21}$$

3.2 Design of multimode interference couplers

Designing the 1×8 and 8×8 MMI structures for the AWG was a slow procedure, due to the learning curve associated with both the physical principles and the simulation software. Many test runs were made over several months until the final designs were achieved.

The design procedure began simply. The width $W = 10$ μm was chosen for the multimode sections in order to achieve overall device size on the correct order (approximately 25 μm for the 1×8 and 100 μm for the 8×8). Then the beat length L_π was calculated according to (3.7). For the purposes of this calculation, the effective width W_e was assumed to be equal to the actual width W. This gives the length

$$L_\pi = 301.08 \ \mu\text{m} \tag{3.22}$$

Finally, the coupler lengths were calculated according to

$$L_{1\times 8} = \frac{3L_\pi}{32} = 28.23 \ \mu\text{m} \tag{3.23}$$

$$L_{8\times 8} = \frac{3L_\pi}{8} = 112.90 \ \mu\text{m} \tag{3.24}$$

Unfortunately, these lengths did not produce the desired results, which were only qualitative at this point. The first attempt to correct the problem involved replacing W with W_e, thus making the coupler somewhat longer. Unfortunately, this did not help. In fact, it made the problem worse! It was finally decided to investigate the couplers without output waveguides, to look at the imaging as the excitation propagated down the waveguide and attempt to see what was causing the deviations from theory. It turned out that the images were in fact forming more closely to the input than expected. This effect was difficult to see with the output waveguides in place

because the reflections at the interface between sections quickly obscured the results.

3.2.1 Phase errors in multimode interference couplers

The reason for the unexpected result above can be understood by considering the problem of phase errors in the MMI. Phase errors cause poor imaging and also may impact the overall crosstalk performance of an MMI AWG. In order to obtain the self-imaging condition in an MMI, we assumed that the propagation constant spacing was given by (3.8). Two approximations went into this derivation.

First, we assumed that the effective width was the same for all modes. (Actually, characterizing the mode distribution with an effective width was an approximation in itself, but it is less significant and not as easily quantified.) The correct relation is given by (3.2), with a different W_{ei} for each mode i. The error can be approximated by [12]:

$$(\beta_0 - \beta_i)_{error1} \approx -\frac{\pi^2(i+1)^2}{n_1 k_0 W_{e0}} \left(\frac{1}{W_{e0}} - \frac{1}{W_{ei}} \right) \tag{3.25}$$

The approximation of equal effective width, however, is acceptable for high index contrast and can be ignored in that case.

The second assumption we made is far more troubling. In (3.5), we expanded the square root in the expression using the approximation

$$\frac{\kappa^2}{n_1^2 k_0^2} \ll 1 \tag{3.26}$$

$$\kappa^2 \ll n_1^2 k_0^2 \tag{3.27}$$

This is the paraxial approximation, that the transverse wave vector is only a small part of the overall wave vector. In other words, the beam is propagating paraxially in the waveguide. However, this assumption is not true for the highest numbered modes

in the waveguide, which propagate at large angles. The phase error is given by the next term in the expansion:

$$(\beta_0 - \beta_i)_{error2} = \frac{\pi^4((i+1)^4 - 1)}{8n_1^3 k_0^3 W_{e0}^4} \tag{3.28}$$

The total phase error for a distance $z = 3L_\pi/N$ is approximately given by [12]:

$$\Delta\phi_i \approx \frac{\lambda_0^2(i+1)^4\pi}{2Nn_1^2W_{e0}^2}\left(\frac{1}{8} - \frac{\lambda_0 n_1^2}{6\pi W_{e0}(n_1^2 - n_2^2)^{3/2}}\right) \tag{3.29}$$

The first term in this sum is the error from the paraxial approximation, while the second is the error from the effective width approximation, which is negligible in a high index contrast system. For a $1 \times N$ symmetric excitation (images at $z = 3L_\pi/4N$), the phase error will obviously be 4 times smaller, which is evident in the performance of the designed structures.

Note that the first term in (3.29) gives too great a phase, while the second term gives too small a phase. The authors of [12] showed that by proper choice of index contrast, the two types of phase error could be made to (approximately) cancel, producing an optimal dispersion relation for many more modes than before. Strangely, this optimal solution is low index contrast (3.5/3.44). Previously it had been thought that high index contrast was best for MMIs because of the large number of modes supported. Higher order modes have sharper features and thus have larger c_i in the modal expansion (3.9) of small input distributions than do lower order modes. Thus, higher order modes allow, in theory, for sharper resolution images [35]. However, if these higher order modes suffer large phase errors in the MMI, then they will not image properly, and the replication will be very poor. While high index contrast may allow for smaller MMIs, it does not necessarily create better ones.

In our MMI structure, the problem is worse than for the high index contrast example in [12] because the smaller effective width W_{e0} increases the phase error.

In fact, phase errors may not have significantly affected earlier high index contrast MMIs which were not designed for small, integrated applications. The equation for phase error in our (8×8) MMI is

$$\Delta\phi_i = (4.6677 \times 10^{-5} - 1.1508 \times 10^{-6})(i+1)^4 = 4.5526 \times 10^{-5}(i+1)^4 \qquad (3.30)$$

Using the normalized frequency method, the multimode waveguide is found to support 42 modes. The phase errors for these modes are plotted in Figure 3-2. The greatest phase error is 141.6621 radians. Of course, to interpret this correctly, we take it modulo 2π, where we see it is 3.4321, almost completely out of phase! Even worse,

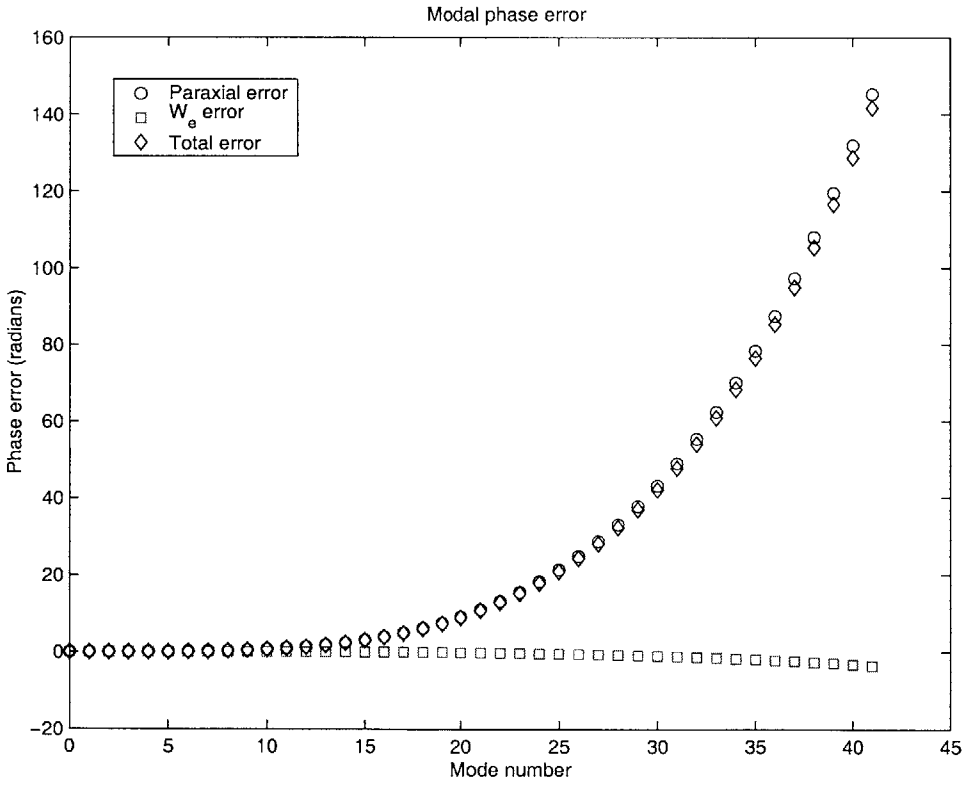

Figure 3-2: Modal phase error for our 8×8 MMI structure

the phase error has circled around 2π many different times. Following the criterion in [12], that we desire phase errors less than five degrees, we find that the sixth mode ($i = 5$) is the highest mode to meet this requirement. This analysis clearly shows the difficulty of making an effective MMI coupler in such a small, high index contrast device.

The highest order mode which contributes significantly to the imaging is approximately equal to the ratio W/W_{wg}, where W_{wg} is the width of the access waveguides. Thus, if the access waveguides are made larger relative to the multimode section, the higher order modes will be less important in the modal expansion of the input field, and the imaging should improve. This strategy was used to improve the performance of the MMI couplers. The original tests featured single mode input waveguides of width 0.2 μm, giving the ratio $W/W_{wg} = 50$. This means that the number of modes was probably not even big enough to achieve the desired resolution! But worst of all, the modes which were being most useful were plagued with large phase errors, causing the imaging to be poor and not in the correct location. Therefore, we also tried waveguide widths of 0.5 μm, 0.6 μm, and 1 μm. Figure 3-3 shows the FDTD results for these different widths on the 1×8 input structure. It is clear that the larger access waveguides are imaged more clearly. For the rest of the design, we chose a 0.6 μm access waveguide. The 1 μm case might image better, but directional coupling issues (discussed more in Chapter 4) would pose a problem at the inputs and outputs. 0.6 μm, while not optimized in any sense, provides a tradeoff between the poorly imaged single mode 0.2 μm waveguides and the very tightly spaced 1 μm waveguides. One problem that presents itself is that the 0.6 μm waveguides are not single mode. Since we want single mode waveguides in the AWG, the best solution is to taper the widths at the entrances and exits to the MMI. If we do this adiabatically, only one mode should be excited, and loss should be minimized.

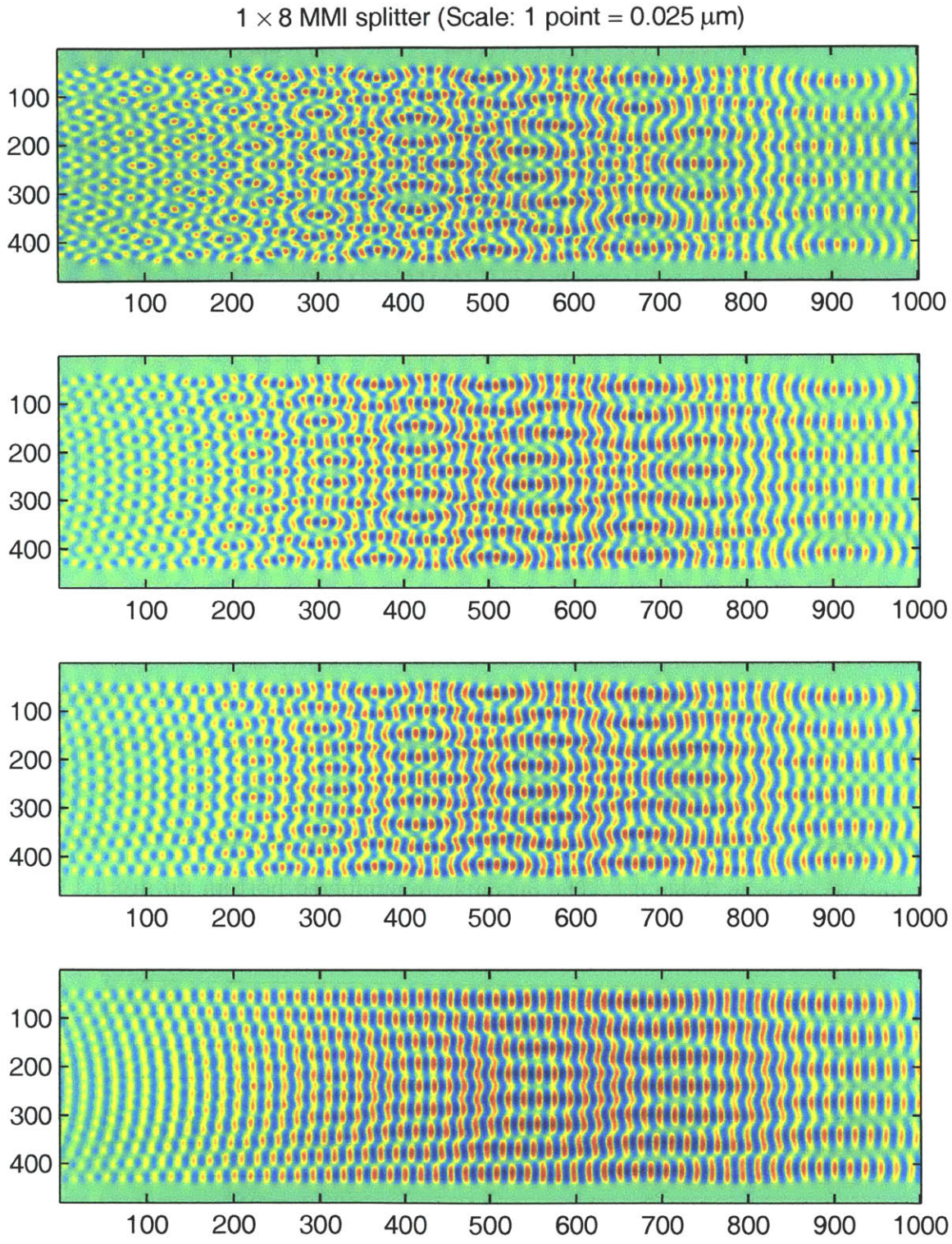

Figure 3-3: 1×8 splitter for different input waveguide widths (and no output waveguides). Top to bottom: 0.2 μm, 0.5 μm, 0.6 μm, 1 μm

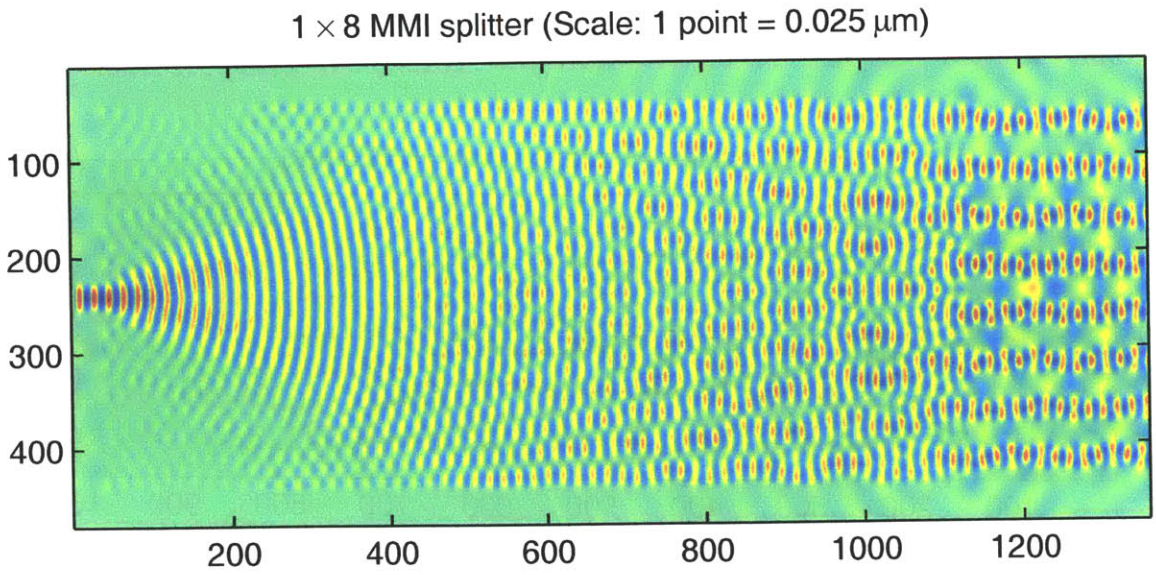

Figure 3-4: 1×8 multimode interference coupler, $W = 10$ μm, $L = 27.25$ μm, access waveguides $= 0.6$ μm

3.2.2 1×8 MMI splitter

Figure 3-4 shows the performance of the final 1×8 MMI design. It was found by assuming a 0.6 μm access waveguide, propagating the excitation down the multimode section, and picking the length where the eight outputs looked the best. Because of the phase errors, the images do not all appear at the same z, so an intermediate value had to be picked to optimize the results. The final length was chosen to be 27.25 μm.

The power transmission, found by both mode overlaps and power flux (for $\lambda = 1.55$ μm), is given in Table 3.1. The outputs are numbered bottom up, from 1 to 8. Since the structure is symmetric, we need only give the results for outputs 1-4.

Output number	Mode overlap transmission	Flux transmission
1	0.1014	0.1094
2	0.1014	0.1130
3	0.0961	0.1160
4	0.1158	0.1298
Total	0.8294	0.9364

Table 3.1: Power transmission of 1×8 MMI splitter

The mode overlap calculation is probably more correct; nevertheless, the design features fairly low loss and good uniformity.

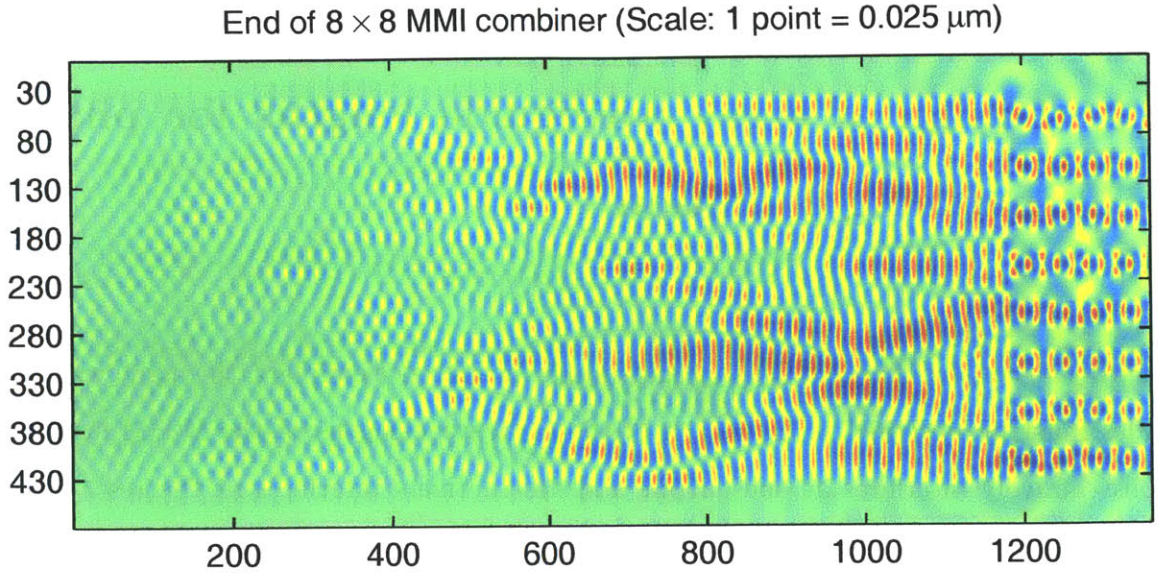

Figure 3-5: 8×8 multimode interference coupler, $W = 10$ μm, $L = 111.5$ μm, access waveguides $= 0.6$ μm

Output number	Mode overlap transmission (input 4)
1	0.139
2	0.049
3	0.076
4	0.109
5	0.043
6	0.112
7	0.104
8	0.041
Total	0.673

Table 3.2: Power transmission of 8×8 MMI combiner

3.2.3 8×8 MMI combiner

The 8×8 design was obtained in the same manner as the 1×8 design. The optimal length was $L = 111.5$ μm. Figure 3-5 shows the output when input number 4 is excited. Table 3.2 gives the mode overlap transmission for this input.

It is clear that the 8×8 MMI has much greater loss and nonuniformity than the 1×8 MMI. This result was expected because the 8×8 structure is longer and thus accumulates more phase error. Nevertheless, we decided to use these results and move on to the next step without optimizing.

3.3 Design of array waveguide lengths for phase matching condition

In order to construct an AWG using multimode interference couplers, the array waveguides must be designed so that each of the eight wavelengths of the WDM system constructively interferes at one output and destructively interferes at the rest. The total phase of the system is composed of the phase of the 1×8 splitter, given by (3.21), the phase from the array waveguides, and the phase of the 8×8 combiner, given by (3.19). Lierstuen and Sudbø [16] give a table showing the required phase differences between array waveguides to couple to each output. However, it is difficult to obtain a set of lengths which, when matched with the eight wavelengths in the WDM system, give these desired phases and correctly route the signals. We will outline a solution by Paiam and MacDonald [27].

Assume that the shortest arm in the array is labeled m. It is simple to choose the array guide lengths so that the center wavelength λ_0 is selected by output waveguide m, i.e., constructive interference occurs for λ_0 at output m. (From here on, we always consider the waveguides to be numbered bottom up at all points in the device.) However, there will be phase mismatch at all other outputs $k \neq m$. We consider the difference in phase at output k between a signal of wavelength λ_0 which originated from array guide $j \neq m$ and one which originated from array guide m:

$$\Delta\phi_{j,k} = (\phi_{i,j} + \phi_j + \phi_{j,k}) - (\phi_{i,m} + \phi_m + \phi_{m,k}) \tag{3.31}$$

where the first term in each sum is the transfer phase of the input MMI, the second is the phase $-\beta_0 L$ from the array guide, and the third is the transfer phase of the output MMI. In order to put λ_0 at output m, we have

$$\phi_{i,j} + \phi_j + \phi_{j,m} = \phi_{i,m} + \phi_m + \phi_{m,m} \tag{3.32}$$

for all j. So (3.31) becomes

$$\Delta\phi_{j,k} = (\phi_{j,k} - \phi_{m,k}) - (\phi_{j,m} - \phi_{m,m}) \tag{3.33}$$

Note that this result does not involve any transfer phase from the input coupler. We can substitute for the transfer phase of the 8×8 MMI, but we must recall that the relation (3.19) applies to waveguide indices going in opposite directions at the input and output. The authors of [27] rederive the phase for their notation and are able to show that $\Delta\phi_{j,k}$ is an integer multiple of $2\pi/N$. In order for a different wavelength $\lambda \neq \lambda_0$ to be constructively interfered at output k, it must counteract this phase with the phase in the array guides. Thus we have the condition

$$L_j - L_m \approx d_j \Delta L \tag{3.34}$$

where ΔL is defined as the distance which will produce a phase shift of $2\pi/N$ between two wavelengths in the WDM system:

$$\beta_0 \Delta L = \frac{2\pi n_{eff0}}{\lambda_0} \Delta L = 2\pi n \tag{3.35}$$

$$\beta_1 \Delta L = \frac{2\pi (n_{eff0} + (dn_{eff}/d\lambda)\Delta\lambda)}{\lambda_0 + \Delta\lambda} \Delta L = 2\pi n - \frac{2\pi}{N} \tag{3.36}$$

Solving, we obtain

$$\Delta L = \frac{\lambda_0(\lambda_0 + \Delta\lambda)}{N\Delta\lambda(n_{eff0} - \lambda_0(dn_{eff}/d\lambda))} \tag{3.37}$$

Paiam and MacDonald [27] spend a good deal of time deriving the smallest set of d_j which avoid waveguide crossings for different size AWGs. The d_j for $N = 8$ are given in Table 3.3.

What we have, therefore, is an easy method to obtain the lengths of the array guides. As long as we fine tune the lengths of the array waveguides to direct λ_0 to output m (in our case, $m = 4$), then the lengths given by (3.34) will ensure that the other wavelengths are uniquely directed to the other seven outputs. In practice, we do this in the reverse order. We first fix the shortest array guide, L_4. In this case we chose $L_4 = 5\mu$m, which is too small for a real device. In reality, all the lengths would have to be increased by a common amount. In any case, once L_4 is chosen, all the other lengths are approximately determined from (3.34) and Table 3.3. Then the lengths are fine tuned to ensure λ_0 ends up in output 4.

To find a set of lengths which will map λ_0 to output 4, we could in theory calculate the phases of the 1×8 and 8×8 couplers from (3.21) and (3.19), but Lierstuen [16] gives us the required phases in his chart. Table 3.4 lists them, as well as the phases needed to map to output 5, for reasons which will be made clear later. All that is required is to find the nearest lengths to the values found from (3.34) which satisfy these conditions for λ_0.

The lengths which were used are listed in Table 3.5. Note that these lengths were calculated using β_{FDTD} rather than β_{real}, so they would need to be adjusted in real life. However, assuming a perfect simulation in all other respects, the response should be the same for the equivalent lengths in real life. Also, it should be noted that the

Array guide number	1	2	3	4	5	6	7	8
d_j	6	3	1	0	4	5	7	10

Table 3.3: Optimum values of d_j for an AWG with $N = 8$ [27]

Output number	$\Delta\phi_{1,2}$	$\Delta\phi_{1,3}$	$\Delta\phi_{1,4}$	$\Delta\phi_{1,5}$	$\Delta\phi_{1,6}$	$\Delta\phi_{1,7}$	$\Delta\phi_{1,8}$
4	$7\pi/8$	$3\pi/8$	π	0	$3\pi/8$	$7\pi/8$	π
5	$-\pi/8$	$-5\pi/8$	π	0	$-5\pi/8$	$-\pi/8$	π

Table 3.4: Array guide phase differences $\Delta\phi_{1,n} = \beta(L_1 - L_n)$ needed to map to outputs 4 and 5 (Lierstuen's notation) [16]

	Length (μm)
L_1	621.0003
L_2	312.9698
L_3	108.0193
L_4	5.1120
L_5	415.9346
L_6	518.6115
L_7	723.1011
L_8	1031.3621

Table 3.5: Lengths chosen for array waveguides in MMI AWG

calculation of ΔL, as well as all the fine adjustments for phase matching, assumed that the array waveguides were 0.6 μm. This was useful because, as described below, we did not actually physically model the array. In real life, we would need to calculate the lengths for single mode waveguides. While these two assumptions seem troublesome, the lengths are relatively easy to choose. Picking those which should mathematically give us the correct output, even assuming 0.6 μm array guides, will provide insight on the quality of the phase transfer in the MMI couplers, the primary source of error.

3.4 Adjustments to FDTD code

Previously, we had only simulated the input and output MMI couplers. In order to measure the performance of the AWG, we needed to also include the effects of the array waveguides and then couple the three elements together. Unfortunately, the memory requirements made it impossible to create one simulation for the entire AWG structure. Thus, some method was needed to take the output of one simulation and use it as the input to another simulation. This was done by making use of the

mode overlap output files [20]. The main program was modified to output the magnitude and phase of the discrete Fourier transform of the electric field at the desired wavelength along the mode overlap cross section. These files are named "finmagX" and "finphsX," where X is the number of the overlap. The initialization program was then modified to allow a new option for spatial profiles of sources, represented in the input file as -3. When -3 is given as the spatial profile, the code reads in the values of the finmag and finphs files. (Note that at present, the code will associate source X in the listing with files "finmagX" and "finphsX," so care must be taken to make sure everything is ordered correctly in the input file.) The values in finmag are used to create the spatial profile across the source cross section. One problem which occurs is normalization. As originally written, the code automatically normalizes all the sources to the same power, so those lines are skipped in the case of a continuing simulation. However, the profile generated by finmag is not normalized correctly either. In order to calculate transfer functions from the input of one simulation to the output of a second, a renormalization step must be done. By looking at the flux at the output of the first simulation and the input of the second simulation, which should be identical, a scale factor can be approximated for the transition.

Meanwhile, the values in finphs are converted into a delay by dividing by $2\pi f_0$. (Note that in this code, $f_0 = 1/\lambda$, so the "delay," and all other time variables, actually have units of length.) This delay is then added to the time variables which are used in the source excitation.

We decided that it would be too time-consuming and difficult to simulate the array waveguides separately from the MMI structures. Therefore, we chose to represent the array waveguides as additional delays in the excitation of the output box. Since the inputs to our simulation are pulses, we have to find both the group delay and the phase delay of these pulses in order to obtain all the wavelength-dependent phase effects of the array guides. The first step is to calculate the group velocity and phase velocity for the 0.6 μm waveguides as described in Chapter 2, with the group velocity approximated by a first difference. We must remember again to use β_{FDTD} instead of β_{real} in these calculations. Then the group delay is L/v_g, and the phase delay is L/v_p.

We must also multiply these times by c to get them into the correct units for the code. These two delays are then added as new inputs to the FDTD code. The group delay is added to the expression for the Gaussian envelope in the source excitation code, while the phase delay is added to the carrier.

These additions to the code worked well, but they were based on some fundamental assumptions. To begin with, the method of using the Fourier magnitude to reproduce the amplitude profile relied on the first simulation having no reflections. For the MMI, this is pretty much true. In addition, the code only uses information at one wavelength, but again this works well for the MMI, which has a largely wavelength-independent performance. Another assumption is that the performance of the first simulation stays the same in time, since the finmag and finphs information is applied to a brand new Gaussian pulse in the second simulation. Again, with the MMI, there is no reason to believe this is not true. The only problem is that the Fourier transform operation changes the scale. The more serious problems come from the fact that we ignored the array waveguides. The field profile taken at the output of the $1\times$ 8 MMI is likely not to be the same as at the input to the 8 $\times$ 8 MMI for several reasons. The array guides themselves will have loss, even though high index contrast should hopefully reduce that problem. In addition, any part of the field at the 1 $\times$ 8 output which is not matched to the modes of the array guides will eventually be lost. It is clear, then, that this scheme is not perfect, and it should perhaps be improved if new simulations of this type are to be made.

3.5 Results and discussion

The results of the simulation are shown in Figure 3-6. The figure was created by taking the wavelength-dependent flux at the output waveguides and dividing by the wavelength-dependent flux (a perfect Gaussian) at the input waveguide. It is thus the transfer function for the MMI AWG. Some error occurs because of the flaws of the simulation continuation scheme described above. In addition, the use of flux measurements is perhaps inferior to mode overlap methods, providing a somewhat

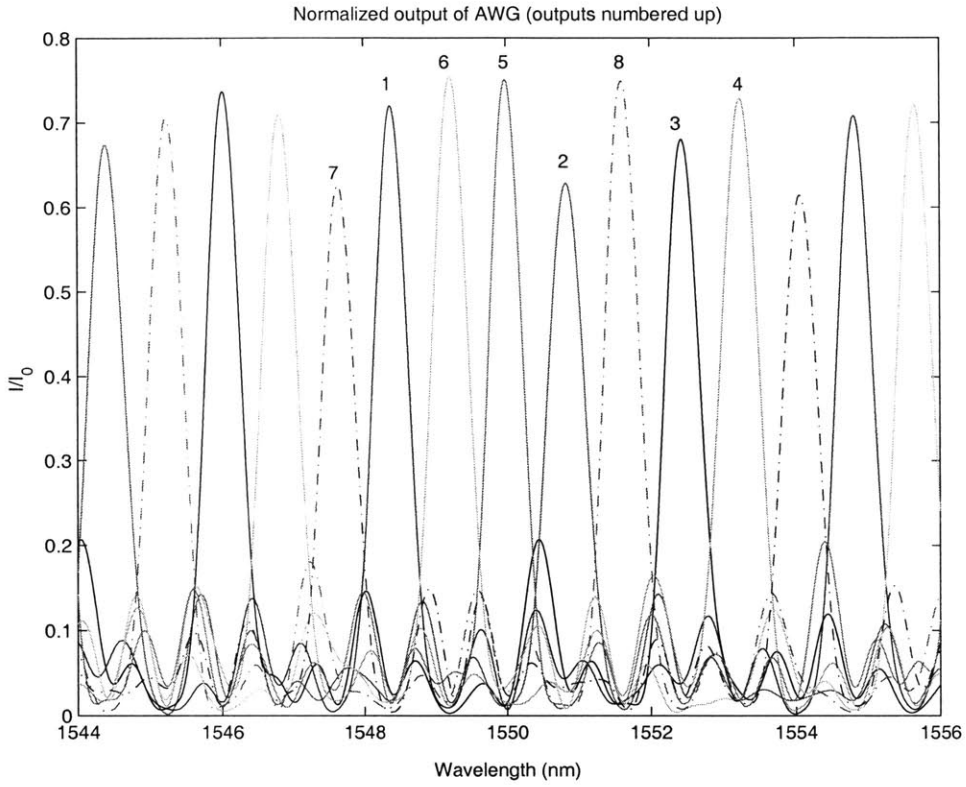

Figure 3-6: Transfer function for MMI AWG, with output waveguides numbered for one period

optimistic view of the loss. However, as implemented, the mode overlaps would never be able to generate a reasonable wavelength-dependent output.

A logarithmic plot is shown in Figure 3-7. It shows that the insertion loss ranges from -1.22 dB in the best channel to -2.01 dB in the worst channel, for a loss uniformity of 0.79 dB. The overall insertion loss is very good; however, we must remember that the flux calculations include some field which will not overlap with the modes of the output waveguides. Loss will occur due to mode mismatch in the array waveguides as well. The fact that the story presented here is incomplete is obvious from Tables 3.1 and 3.2, which predict a higher loss for the overall structure. The loss uniformity is also partially predicted by the results of Table 3.2, which shows that the 8×8 MMI

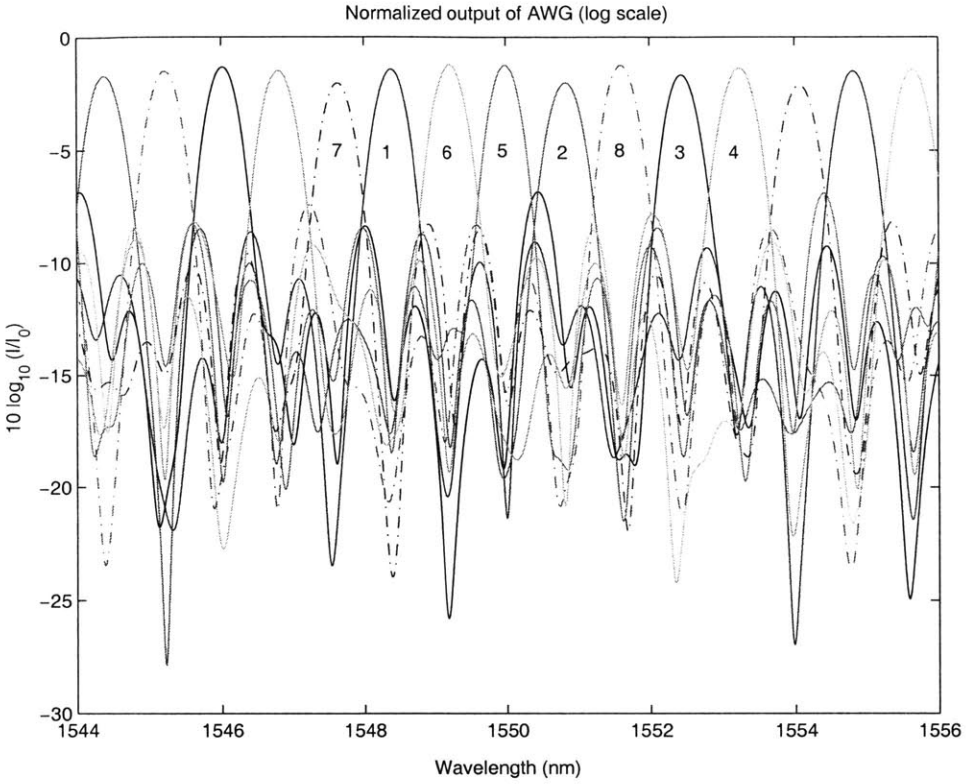

Figure 3-7: Transfer function for MMI AWG, logarithmic scale

we designed can be strongly nonuniform. Nevertheless, the results are on par with or better than many AWGs based on traditional star couplers.

The failure of the MMI AWG is the crosstalk. The worst nearest neighbor crosstalk is around -10 dB, for light at the peak wavelength of channel 7 found in channel 6. The nearest neighbor crosstalk may be partially caused by the flux cross sections, which catch part of the field that is not actually coupled to the mode of the associated waveguide. Again, it would be better to use mode overlaps or even to reposition the flux cross sections. Nevertheless, the crosstalk performance is still disappointing. The crosstalk never gets better than -15 dB at the peak wavelength of any channel, and we would like much better performance than this.

One source of error was discovered soon after the simulation. It turns out that the center wavelength 1.55 μm actually ends up in output waveguide 5, not 4! Obviously, Lierstuen [16] labels the waveguides differently than we did, but the paper never indicates his convention. Therefore, some other wavelength, not 1.55 μm, was (unknown to us at the time) the "center." However, since the structure is periodic, it should make little difference. The only problem should be the use of the wrong λ_0 (and n_{eff0}, etc.) in the expression for ΔL, which was already approximate. In addition, it was discovered that ΔL was calculated slightly incorrectly due to an error reading and transcribing from [27]. The group index for the array waveguides was used instead of the effective index. A second simulation was soon run to correct these problems. The lengths were found using the new ΔL and the phases to map to output "5" from Table 3.4. They are given in Table 3.6, and the results are pictured in Figure 3-8. ΔL changed only slightly, so little change in the results was expected. In fact, the loss and crosstalk were slightly worse on the second try. A slight improvement was noticed in the location of the peaks of the transmission with respect to the desired wavelengths. Still, the changes were minor and illustrate that the real problem is not in the array guides, but in the MMIs.

3.5.1 Analysis of errors

The two results given above in Figures 3-6 and 3-8 are very similar, despite the fact that one is based on a correct use of the formulas in [27], while the other is not. This

	Length (μm)
L_1	623.1618
L_2	313.9791
L_3	108.1070
L_4	4.9694
L_5	417.1744
L_6	520.0817
L_7	726.4146
L_8	1035.3668

Table 3.6: Lengths chosen for array waveguides in MMI AWG, second attempt

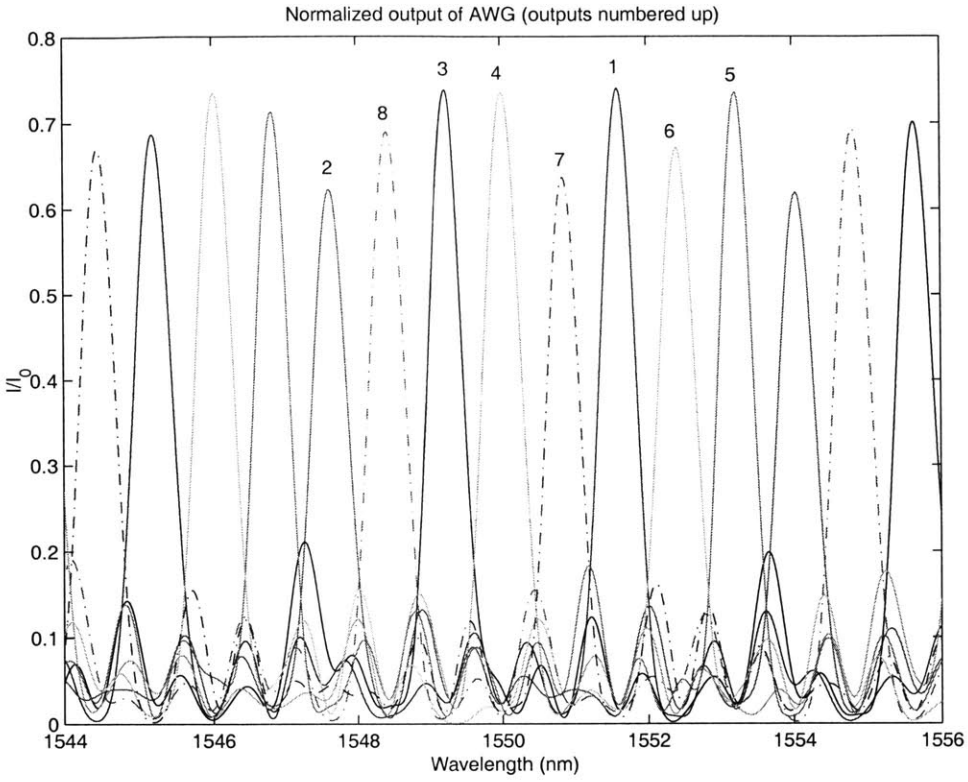

Figure 3-8: Transfer function for MMI AWG, second attempt

fact emphasizes that the problems with the MMI AWG are probably not in the array waveguides; instead, they are probably due to the MMIs.

One reason for the crosstalk of the output is almost certainly due to the MMI phase errors analyzed earlier. We based our derivation of the array guide lengths on the calculated phases of the MMI couplers. If these phases are incorrect, due to the rather sloppy imaging we expect for these small, high phase error couplers, then the array guide lengths are not optimized. The simplest solution to this problem is to find the real phases of the MMI couplers and adjust the array guide lengths accordingly. Other approaches would be to analyze the MMIs more carefully and try to find new ways to design them for increased accuracy. This problem will be discussed further

in Chapter 5.

A second reason for the crosstalk of the output is a fundamental problem of MMI-based AWGs. The reason can be understood by considering digital filter design [26]. In essence, we want to design a digital filter which implements a certain transfer function. For each output waveguide, the desired transfer function is a sharp bandpass filter around the wavelength assigned to that channel. Unfortunately, such a response can only be obtained by an infinite impulse response (IIR) digital filter. We have to approximate this filter with a finite impulse response (FIR), or transversal, filter. In general, the impulse response of this filter has the form

$$h[n] = a_0\delta[n] + a_1\delta[n-1] + a_2\delta[n-2] + \cdots + a_N\delta[n-N] \tag{3.38}$$

This is a series of delays, all spaced by integers. The a_n are called the tap coefficients. In the frequency domain, the response is

$$H(e^{j\omega}) = a_0 + a_1e^{-j\omega} + a_2e^{-2j\omega} + \cdots + a_Ne^{-Nj\omega} \tag{3.39}$$

So each term is a different phase shift, $-m\omega$ for integer m. We can see that for the AWG, the delays of (3.38) correspond to the array guide delays if the integer spacing is changed to ΔL. Similarly, the phase shifts of (3.39) are replaced by $-m\beta\Delta L$.

In order to approximate an IIR filter by an FIR filter, what we are really doing is multiplying the infinite impulse response by a windowing function which selects out only some of the values. In the frequency domain, this multiplication becomes a convolution. In the case of the MMI AWG, we are multiplying by a windowing function which is 1 for all values of m which are in our list of d_j (Table 3.3) and 0 for all other values of m. If we were to use a continuous set of d_j with no gaps, the windowing function would be a rectangle, which corresponds to a sinc in the frequency domain. The convolution of the sinc with the desired frequency response produces sidelobes, which we see in Figures 3-6 and 3-8. Even worse, our windowing

function actually has jumps in it, since d_j does not simply go from 0 to 7 without any gaps. This strange window makes the frequency response worse and is the primary reason Paiam and MacDonald were searching for the smallest set of d_j for a given number of channels [27]. By contrast, the traditional free space coupler excites the array waveguides not with equal amplitudes, but with a Gaussian distribution. Such a windowing function produces smaller sidelobes and better performance.

Despite this setback, it seems conceivable to somehow alter the windowing function of the MMI AWG by nonuniform splitting of the input. This idea will be discussed in Chapter 5. The more fundamental problem lies in the size of the window. Because of the way the MMI couplers work, we can only have as many array waveguides as outputs (8 in our case), whereas free space coupler AWGs can and do have many more array waveguides than channels. With more array waveguides, we can include more terms of the desired IIR filter and obtain a more exact response. Because the MMI AWG is limited in the number of array waveguides it can have, it will never compete with the free space coupler AWG in lowering crosstalk. All of these problems combine to make the MMI AWG suboptimal for our purposes.

Chapter 4

A New AWG Input Structure Using Coupled Mode Theory

The multimode interference coupler AWG described in the last chapter was successfully implemented for high index contrast at small size, but the design was not optimal. While the overall loss of the device was low, the transfer function contained undesirable sidelobes and crosstalk into neighboring channels. While there were possible ways to improve the MMI AWG, including both simple optimization and more complex enhancements, our attention soon turned to a different method of coupling to be used as the input to the AWG.

Since the mode of a waveguide extends past the core as an evanescent tail, the mode in one waveguide may actually excite a field in a neighboring waveguide. The mathematics which describes this coupling is known as coupled mode theory. The effects can be harmful, causing closely spaced waveguides to couple into one another and destroying the original intention of a device. However, the effects can also be useful. Coupled modes can be exploited to make so-called directional couplers for power transfer or division. With the application of a voltage, these couplers can be used as tunable filters or switches [10]. In addition, coupled modes have been used to optimize the uniformity of a star coupler [25], [38], [29], [37].

Our second AWG design uses an array of coupled waveguides as the input combiner instead of a free space region or 1×8 MMI coupler. The input signal excites one

of these coupled waveguides, and the power is then transferred to all of them before entering the dispersive array as usual. There are enough degrees of freedom in the system to achieve any desired distribution, providing a great deal of control over the final output of the AWG. Such large scale use of coupled waveguides has, to our knowledge, not been studied before, especially in the context of an AWG.

In this chapter, we review the basics of coupled mode theory, including an interesting solution to the excitation of one waveguide in a infinite, uniformly coupled array. Then we describe our efforts to determine coupling coefficients for uniform and Gaussian distributions in as many as 41 waveguides. Next we consider the problem of phase, critical to the operation of the rest of the AWG, and a simple solution to this problem. Finally, we briefly outline the design of the output star coupler which will be part of this new AWG.

4.1 Coupled mode theory [10], [18]

4.1.1 Derivation of coupled mode equations

Consider a system of N adjacent single mode waveguides aligned along the z direction. The evanescent tails of the modes of these waveguides interact with the other waveguides and create a coupling situation. When the waveguides are spaced far apart, this coupling is negligible, and the waveguide modes are those we found in Chapter 2, unaffected by the surrounding guides. However, when the waveguides are closely spaced, the coupling can be significant.

To study this coupling, we first define the mode of the nth waveguide to have profile $E_n(x)$ and propagation constant β_n:

$$E_n(x,z) = E_n(x)e^{-j\beta_n z} \tag{4.1}$$

We assume that these modes are normalized to have unit power. The following

expression calculates the time averaged power of one mode:

$$\frac{1}{2}\int E_{yn}H_{zn}^{*}dxdy = \frac{\beta_n}{2\omega\mu_0}\int E_n^{*}E_n dxdy = 1 \tag{4.2}$$

We also assume that the modes of different waveguides are orthogonal. Putting this fact together with (4.2), we get the condition

$$\int E_m^{*}E_n dxdy = \frac{2\omega\mu_0}{\beta_n}\delta_{mn} \tag{4.3}$$

The wave equation for the total field is

$$\nabla^2 E(x,z) + \omega^2\mu_0\epsilon E(x,z) = 0 \tag{4.4}$$

(Recall that we are using TE polarization throughout, so this field is in the y direction.) We can assume that the solution to the equation must be some superposition of the original modes:

$$E(x,z) = \sum_{n=1}^{N} a_n(z)E_n(x)e^{-j\beta_n z} \tag{4.5}$$

Plugging this trial solution into (4.4), we find:

$$\sum_{n=1}^{N}\left(\frac{d^2a_n}{dz^2} - 2j\beta_n\frac{da_n}{dz} + \omega^2\mu_0(\epsilon-\epsilon_n)a_n\right)E_n e^{-j\beta_n z} = 0 \tag{4.6}$$

$\epsilon = \epsilon(x)$ is the permittivity distribution for the whole structure, while ϵ_n is the permittivity for waveguide n only, as if the other waveguides were not present. We can ignore the second order derivative terms in this expansion if we assume that the fields are changing only slowly in z. We can now multiply this equation by E_m^*, the conjugate of the mth mode, and integrate over the entire structure:

$$\sum_{n=1}^{N} \left(-2j\beta_n \frac{da_n}{dz} \int E_m^* E_n dx dy + \omega^2 \mu_0 a_n \int (\epsilon - \epsilon_n) E_m^* E_n dx dy \right) e^{-j\beta_n z} = 0 \qquad (4.7)$$

From (4.3), we see that the first term in the sum is nonzero only if $m = n$. We obtain the equations

$$\frac{da_m}{dz} e^{-j\beta_m z} = \sum_{n=1}^{N} \left(-j\frac{\omega}{4} a_n \int (\epsilon - \epsilon_n) E_m^* E_n dx dy \right) e^{-j\beta_n z} \qquad (4.8)$$

$$\frac{da_m}{dz} = \sum_{n=1}^{N} \kappa_{mn} a_n e^{-j\Delta\beta_{nm}} \qquad (4.9)$$

Here κ_{mn} is the coupling coefficient between waveguides m and n.

$$\kappa_{mn} = -j\frac{\omega}{4} \int (\epsilon - \epsilon_n) E_m^* E_n dx dy \qquad (4.10)$$

We ignore the case when $m = n$, which describes a perturbation to the mode of a waveguide from the presence of other waveguide structures. We also usually ignore the coupling coefficients for all waveguides which are not nearest neighbors; that is, we assume $\kappa_{mn} = 0, m \neq n \pm 1$. For a TE slab waveguide, the coupling coefficient between two waveguides is [10]:

$$\kappa = -j\frac{\gamma\kappa_x^2}{(\kappa_x^2+\gamma^2)(\beta d+\beta/\gamma)}e^{-\gamma(D-2d)} \tag{4.11}$$

Here κ_x is the transverse wavevector, formerly known as κ but changed to reduce confusion. D is the separation between the centers of the two slabs. Note the exponential dependence of the coupling coefficient on separation distance, due to the exponentially decaying nature of the cladding field.

$\Delta\beta_{nm}$ is the relative detuning of propagation constants between waveguides. In particular, $\Delta\beta_{nm} = \beta_n - \beta_m$. In this thesis we only consider waveguide couplers with identical waveguides, so $\Delta\beta_{nm} = 0$. It would perhaps be interesting to investigate the effects of detuning in such large scale couplers, since it would provide an extra $N-1$ degrees of freedom. However, for simplicity, we ignore such effects.

4.1.2 Solving the coupled mode equations

The best way to solve the coupled mode problem is to use eigenvalue methods as applied to differential equations. The system of differential equations we are trying to solve is

$$\frac{d\mathbf{a}}{dz} = \mathbf{K}\mathbf{a} \tag{4.12}$$

Here $\mathbf{a}$ is the vector of mode amplitude coefficients for each waveguide. $\mathbf{K}$ is the matrix of coupling coefficients, which is nonzero only for elements with indices m and n satisfying $m = n \pm 1$:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \tag{4.13}$$

$$\mathbf{K} = \begin{bmatrix} 0 & \kappa_1 & \cdots & 0 & 0 \\ \kappa_1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \kappa_{N-1} \\ 0 & 0 & \cdots & \kappa_{N-1} & 0 \end{bmatrix} \tag{4.14}$$

The key to solving the system is to assume solutions with z dependence $e^{-j\beta z}$. Then the system of equations is

$$-j\beta\mathbf{a} = \mathbf{K}\mathbf{a} \tag{4.15}$$

Hence it is an eigenvalue problem with eigenvalue $-j\beta$. To solve it, we find the eigenvalues of the matrix $\mathbf{K}$ and the corresponding eigenvectors. These eigenvectors give the coefficients a_n for the expansion of the eigenmodes as a superposition of the individual waveguide modes. The original modes of the individual waveguides no longer propagate unchanged along z, but these new eigenmodes, often called supermodes, do. The complete solution to the field at some point z is given by a weighted sum of these supermodes, each multiplied by the corresponding phase factor $e^{-j\beta z}$:

$$E(x,z) = \sum_{i=1}^{N} c_i \phi_i(x) e^{-j\beta_i z} \tag{4.16}$$

Here, ϕ_i are the supermodes, and c_i are the weighting factors. These are determined by the initial conditions. With the right initial conditions, as many of the modes as desired can be excited.

4.2 The arbitrary profile problem

In Chapter 3, we designed an MMI AWG which exhibited low loss and qualitatively correct behavior, but which also showed unacceptable levels of crosstalk. We propose a new type of input structure for the AWG made up of an array of coupled waveguides. We believe such a structure can offer better performance than an MMI for several reasons. To begin with, one of the failures of the MMI is the phase error accumulated due to high numbered modes propagating at large angles. This effect causes increased loss and, more importantly, errors in the overall phase which show up again as crosstalk in the output. A system of coupled waveguides is a much more controllable system. Loss in such a system should be minimal, while the phase at the output will be well known. Any approximations made in the design of a coupled waveguide system should have only little effect on the output, unlike the large errors we found in the MMI. In addition, the MMI AWG also suffers from crosstalk because the number of array waveguides can only be as large as the number of outputs. As we saw in Chapter 3, the sampling of the field by the array can be considering a windowing of an infinite impulse response filter. To better approximate the desired frequency response, more array waveguides are needed. With a coupled waveguide system, we can distribute the incoming power over as many waveguides as we like.

It seems at first glance that many of the problems of the MMI could be solved by using the more traditional free space coupler design. Such a design naturally has less crosstalk due to the lack of large phase errors and the ability to use more array waveguides. Yet the free space coupler still features a great deal of loss and is difficult to control precisely. The coupled waveguide solution features low loss and extra degrees of freedom to fine tune the device response. Unfortunately, the coupled waveguides cannot be used as the output of the AWG, because we cannot engineer the

necessary interference condition at different output ports. Thus we will have to use a free space coupler as the output structure. However, if the response of the output is nonideal, simple adjustments can be made in the coupling coefficients of the input to attempt to fix the problem.

The essential problem of this chapter is simple. We wish to find the set of coupling coefficients κ which produce certain output distributions at a given z in a coupled waveguide array. Solving this problem directly is incredibly difficult. Examining (4.16), we see that it is nontrivial to solve for the matrix $\mathbf{K}$ given the field at $z = 0$ and at some other $z = z'$. Instead, we examine an interesting analytic solution to the coupled waveguide problem and then modify its results until we achieve the desired result.

4.2.1 Bessel function property of coupled waveguides [36]

Consider the Bessel function identity

$$2\frac{dJ_n}{dz} = J_{n-1} - J_{n+1} \tag{4.17}$$

This equation looks very similar to the coupled mode equations (for uniform nearest neighbor coupling):

$$\frac{da_n}{dz} = j\kappa a_{n-1} + j\kappa a_{n+1} \tag{4.18}$$

Note that here we have taken the complex factor j out of the definition of κ to better emphasize the nature of the phase shift between waveguides, an issue which will be discussed in more detail later[1]. If we make the substitution $a_n = j^n J_n$, we

[1]Unfortunately, for most of this chapter, this notation is used assuming positive κ, but equation (4.11) shows that the coupling coefficient between two slab guides is actually negative. We must remember that the phases we find mathematically in this chapter are actually the reverse of real life.

obtain

$$j^n \frac{dJ_n}{dz} = jj^{n-1}\kappa J_{n-1} + jj^{n+1}J_{n+1} \tag{4.19}$$

$$\frac{dJ_n}{dz} = \kappa J_{n-1} - \kappa J_{n+1} \tag{4.20}$$

If we substitute $\kappa = 1/2$, we recover the relation (4.17). This shows that for a constant coupling coefficient of $1/2$, if we excite waveguide 0 at $z = 0$, the mode amplitude for waveguide n at distance z is equal to $J_n(z)$. In fact, $\kappa = 1/2$ is not required; scaling κ just (inversely) scales the distance z over which coupling occurs. So any constant coupling coefficient produces amplitude distributions that go like Bessel functions in z.

A Matlab program was used to plot the Bessel relation for 99 waveguides, shown in Figure 4-1(a). Each row is a waveguide, and the colors represent the magnitude of the Bessel function squared.

A second Matlab program was used to calculate the response of 99 waveguides

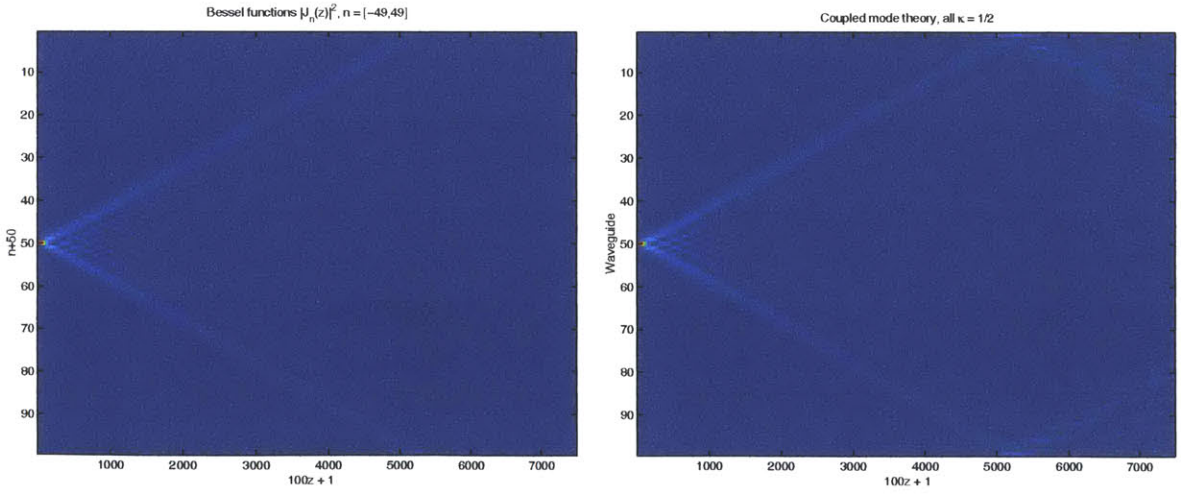

Figure 4-1: (a) Bessel functions $|J_n(z)|^2, n = [-49, 49]$, (b) Power in 99 uniformly coupled waveguides

coupled with $\kappa = 1/2$ when the center waveguide is excited. The result is shown in Figure 4-1(b). It is clear that the basic pattern is the same as the Bessel functions. Some slight differences do occur at the edges because the Bessel relation (4.17) requires an infinite set of Bessel functions. Thus, uniformly coupled waveguides only follow the Bessel function pattern perfectly if there is an infinite set of waveguides.

For a finite number of waveguides N, we can still find an analytic solution. We consider only the case of N odd so we can excite the structure from the center, but the argument can be generalized. We handle the finite number of waveguides by treating the ends of the coupler as mirrors. This can be simulated in a system of infinite waveguides by first placing a negative dummy source into waveguides $N+1$ and $-(N+1)$. These sources cancel the excitation of the real input in waveguides $(N+1)/2$ and $-(N+1)/2$, respectively, making it seem as if the waveguide array ends at those points with just N total guides. However, the new source in waveguide $N+1$ will also create an unwanted excitation in waveguide $-(N+1)/2$, so we need another positive source to cancel this contribution. In order to fully cancel the field in the $(N+1)/2$ and $-(N+1)/2$ waveguides, we need an infinite array of alternating positive and negative dummy sources. The field in any waveguide can then be written as the sum of the Bessel contributions from each of these sources. For example, the response of the center waveguide of a uniformly coupled three waveguide system (excited in the center) can be written as

$$a_0(z) = J_0(z) - J_4(z) - J_{-4}(z) + J_8(z) + J_{-8}(z) + \cdots \tag{4.21}$$

Figure 4-2 shows the actual response for this waveguide using coupled mode theory, as well as the Bessel approximation using the first five terms above. It is clear that the Bessel formula works well for a while, but at larger z, more Bessel functions need to be used to obtain an accurate solution.

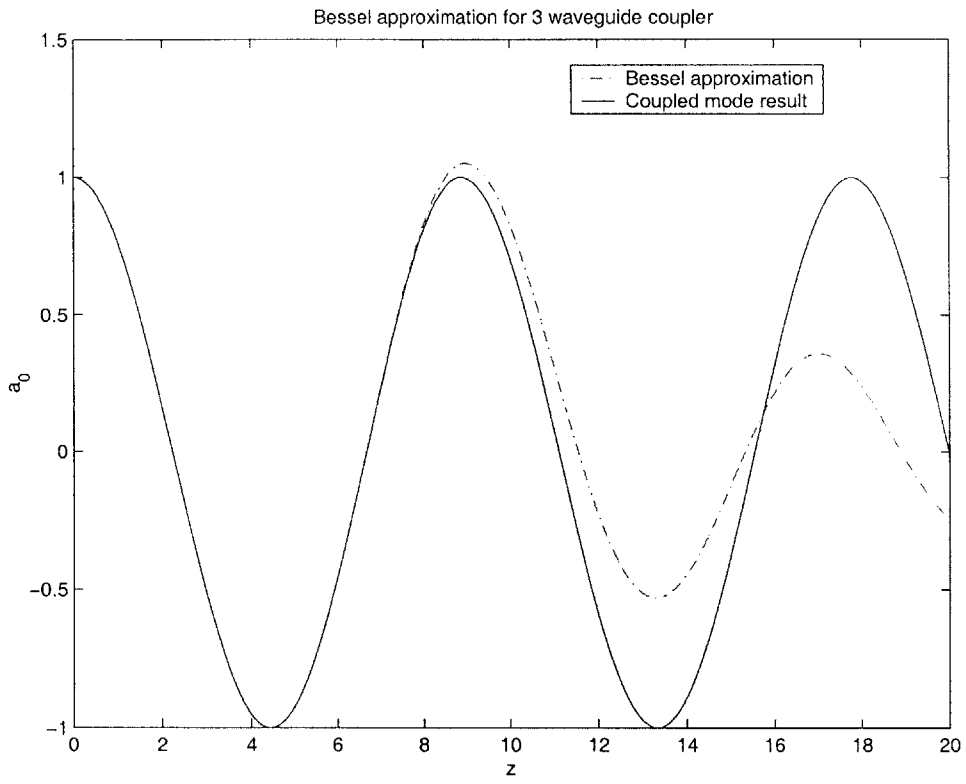

Figure 4-2: Coupled mode solution for a_0 in centrally excited three waveguide system, compared to approximation by sums of Bessel functions

4.2.2 Matlab tools for determining coupling coefficients

The coupled wave solution depicted in Figure 4-1(b) was generated using a Matlab program which makes use of the propagation of supermodes (4.16). Prior to running the program, the user must define a coupling matrix, find the eigenvalue matrix $\mathbf{D}$ and the eigenvector matrix $\mathbf{V}$, define an initial condition $\mathbf{a}(0)$, and then solve for the supermode expansion coefficients $\mathbf{c}$ using the expression

$$\mathbf{c} = \mathbf{V}^{-1}\mathbf{a}(0) \tag{4.22}$$

The program then takes as inputs $\mathbf{D}$, $\mathbf{V}$, $\mathbf{c}$, and a specific length or range of lengths z. The output $\mathbf{A}$ is a matrix of the waveguide mode amplitudes at each value of z, which can then be studied for a specific waveguide or distance or viewed as a 2D color image like the ones above.

To fit the coupling coefficients to the desired values, we began with uniform coupling between all the waveguides, $\kappa = 1/2$. Then the coupling coefficients were manually adjusted in order to match the desired profile. Obviously, such a process becomes increasingly more difficult as the number of waveguides in the coupler increases. Finally, once the output profile was close to the desired profile, Newton's method was used in order to make the results more accurate.

Newton's method [9]

Newton's method is used to calculate the coupling coefficients which will produce a desired profile, given an initial guess obtained through trial and error. Two different Matlab programs are used to perform this optimization. The first is designed for flat profiles, while the second is made for Gaussian profiles. The method begins by calculating the output at a given distance for an initial guess κ of the coupling coefficients. The resulting waveguide mode amplitudes are denoted a_n, but for the rest of the algorithm we are concerned only with the power in each waveguide, $p_n = |a_n|^2$. We calculate the vector function

$$\mathbf{f}(\kappa) = \mathbf{p}(\kappa) - \mathbf{p}_{desired} \tag{4.23}$$

where $\mathbf{p}(\kappa)$ is a vector of the calculated powers for the guess κ in all waveguides but the center. (Power conservation guarantees that if the other waveguides have the correct output, the center will as well. Said another way, there are not enough adjustable coefficients to set the amplitudes in each waveguide uniquely.) For a uniform profile,

$\mathbf{p}_{desired}$ is a vector with all elements equal to the value of p in the center waveguide. For the Gaussian profile, each element is equal to the value of a Gaussian at that point. (As seen below, the program treats each waveguide as an integer along x in the formulation of the Gaussian.) The goal is obviously to achieve $f(\kappa) = 0$, i.e., to find the roots of $f(\kappa)$. We can expand this equation as

$$\mathbf{f}(\kappa) = \mathbf{f}(\kappa^{(m)}) + \nabla\mathbf{f}(\kappa^{(m)}) \cdot d\kappa^{(m)} + \cdots = 0 \tag{4.24}$$

where m is the iteration number. Here $\nabla\mathbf{f}$ is the Jacobian matrix:

$$\nabla\mathbf{f} = \begin{bmatrix} \frac{\partial f_1}{\partial \kappa_1} & \frac{\partial f_1}{\partial \kappa_2} & \cdots & \frac{\partial f_1}{\partial \kappa_n} \\ \frac{\partial f_2}{\partial \kappa_1} & \frac{\partial f_2}{\partial \kappa_2} & \cdots & \frac{\partial f_2}{\partial \kappa_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial \kappa_1} & \frac{\partial f_n}{\partial \kappa_2} & \cdots & \frac{\partial f_n}{\partial \kappa_n} \end{bmatrix} \tag{4.25}$$

where the subscripts indicate elements of $\mathbf{f}$ and κ. In Matlab, this matrix can be calculated by changing each element of the current guess of κ by a small amount, recalculating $\mathbf{f}$, and then estimating the derivatives using a first difference.

The vector $d\kappa$ is the correction which should be applied to create the next guess for κ. It is calculated by solving the matrix equation (4.24):

$$\nabla\mathbf{f}(\kappa^{(m)}) \cdot d\kappa^{(m)} = -\mathbf{f}(\kappa^{(m)}) \tag{4.26}$$

Then the correction $d\kappa$ is added to κ to create a new guess:

$$\kappa^{(m+1)} = \kappa^{(m)} + d\kappa^{(m)} \tag{4.27}$$

and the process begins again. It is repeated until κ converges (if it does).

4.2.3 Uniform distribution

The first amplitude profiles we generated had constant power across all the waveguides. The coupling coefficients were found for 3, 5, 7, 21, 31, and 41 waveguides. (We always used an odd number so that we could launch from the center and take advantage of symmetry.) For 3 waveguides, the problem is trivial. The two coupling coefficients are the same, and constant power distribution always occurs at some z. For the other cases, many iterations were run by hand before turning to Newton's method for a more exact solution. For example, Figure 4-3(a) shows the best attempt

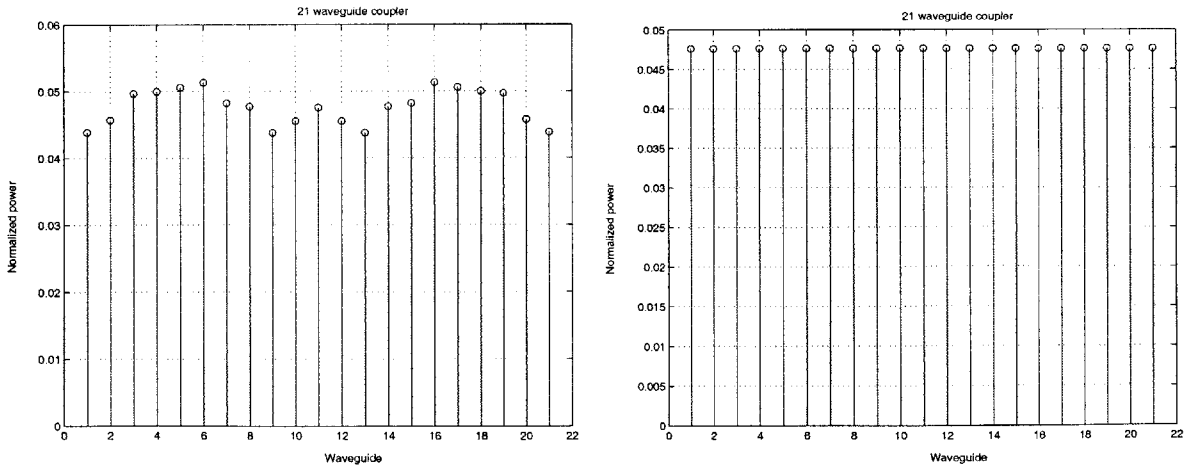

Figure 4-3: (a) Output of 21 waveguide coupler for uniform profile, κ chosen by inspection, (b) Output of 21 waveguide coupler for uniform profile, κ optimized by Newton's method

at a uniform power profile by trial and error adjustment of the coupling coefficients. Figure 4-3(b) shows the optimized profile, after Newton's method was applied to the best guess.

Figure 4-4 shows the power in all 41 waveguides of the 41 waveguide coupler from $z = 0$ to $z = 1$, where the uniform profile occurs. (Actually, in the original determination of the coupling coefficients, the uniform profile was found at a convenient z given that the trial and error process started at the Bessel function result. For purposes of later analysis, the coupling coefficients were then adjusted to place the desired output at $z = 1$.) Figure 4-5 shows the same picture, zoomed in on the area around $z = 1$.

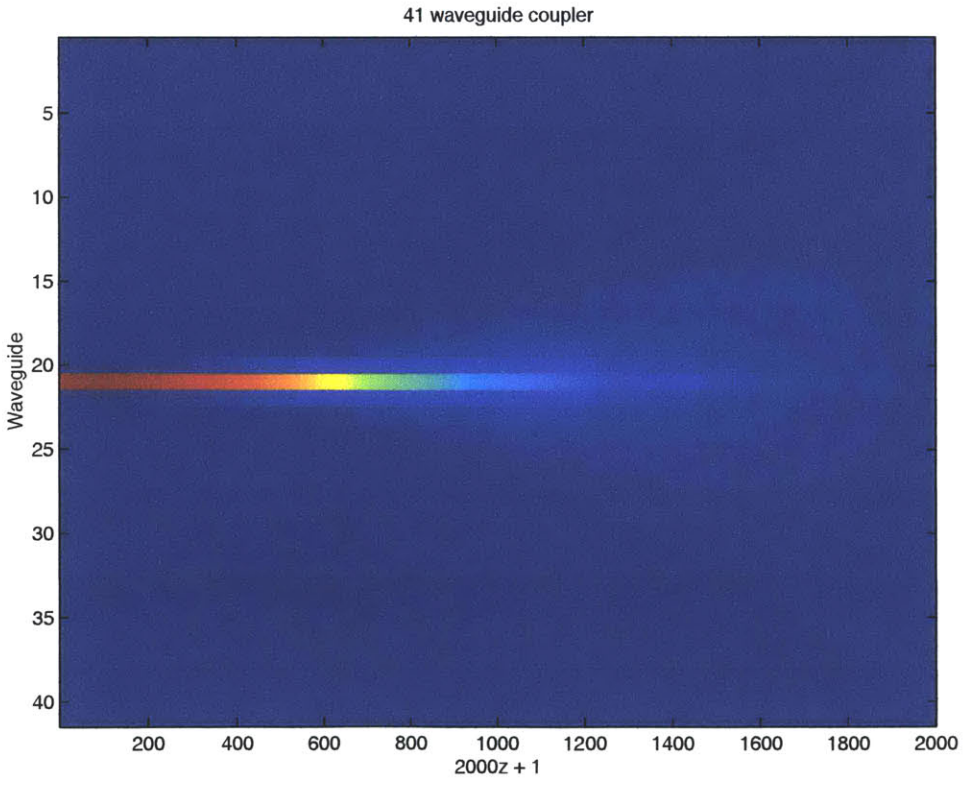

Figure 4-4: 41 waveguide coupler for uniform profile

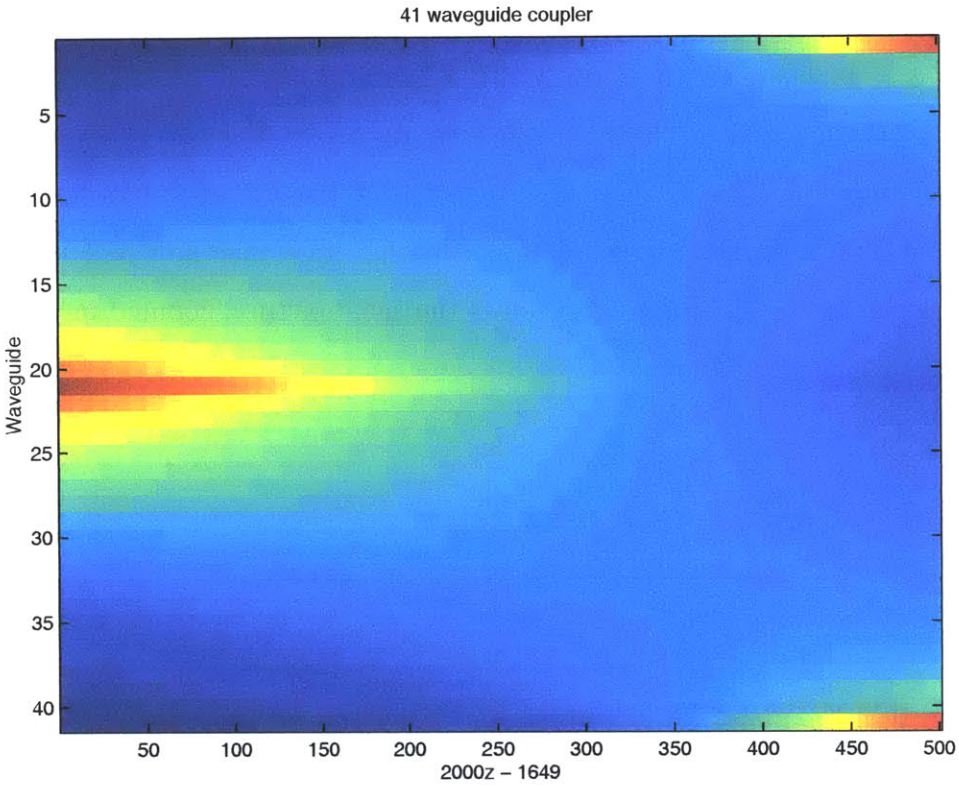

Figure 4-5: Close up image of 41 waveguide coupler for uniform profile

The figures show how the coupling is at first small near the center and then becomes larger away from the center. We expected this from the constant coupling case (see Figure 4-1(b)), which shows that power is coupled easily into waveguides near the center but takes a long time to reach the outside guides. Figure 4-6 shows the coupling coefficients for 3, 5, 7, 21, 31, and 41 waveguides with the output at $z = 1$. We see that the coupling coefficients do indeed increase as we move away from the center guide. Furthermore, the basic pattern seems to be the same for all the couplers. However, we also notice that the coupling coefficients eventually start to go back down. This can be explained by the reflection at the end of the coupler. We lower the outside coefficients so the field does not couple back towards the center too

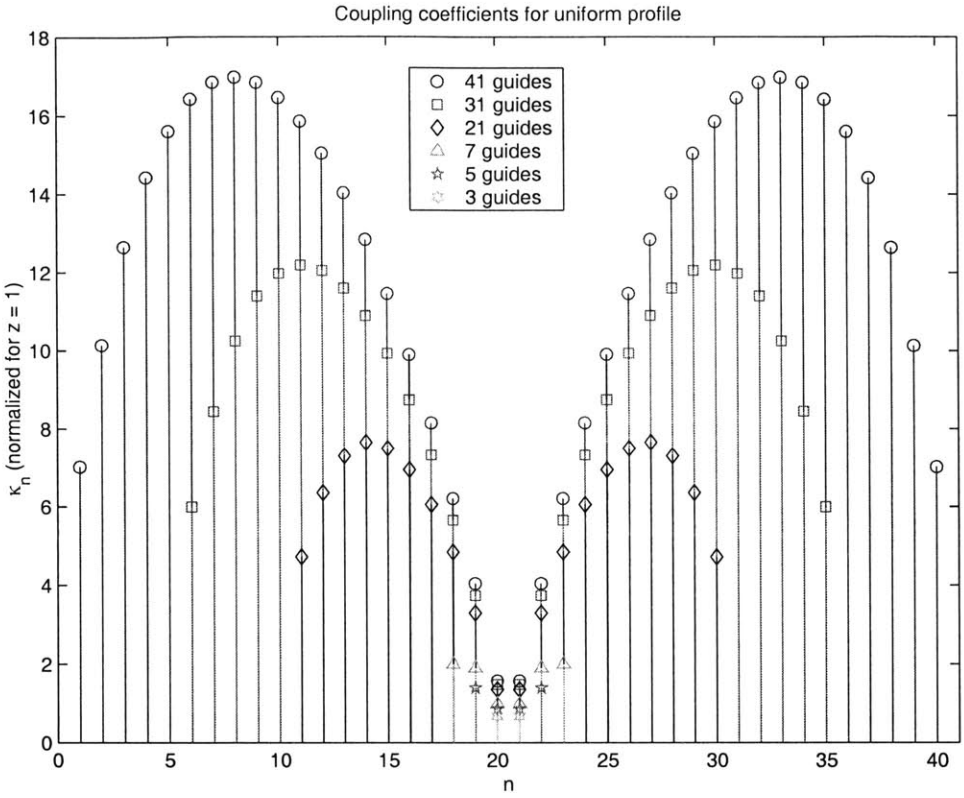

Figure 4-6: Coupling coefficients for uniform profile couplers

quickly.

A uniform profile coupler like these could potentially be very useful. Previously, waveguide couplers were generally used only for a few ports, while other devices were used to realize couplers requiring a larger number of ports. Free space star couplers are perhaps the most common devices, but much effort is required in order to obtain a uniform response in the output. MMIs often have excellent uniformity, but as we have seen, their performance breaks down for small, high index contrast devices. In addition, the number of ports is limited in an MMI. It is clear that large N waveguide couplers have the potential to be quite useful in their own right. Furthermore, an 8 waveguide uniform coupler could replace the input MMI of the MMI AWG, reducing

the phase errors and loss associated with that part of the device.

4.2.4 Gaussian distribution

In order to use a free space output coupler, we desire a Gaussian distribution, which will Fourier transform in the far field to another Gaussian. Figure 4-7 shows a Gaussian pattern generated for 41 waveguides following the equation

$$p = \frac{1}{17.6587} e^{-\frac{(i-21)^2}{100}} \tag{4.28}$$

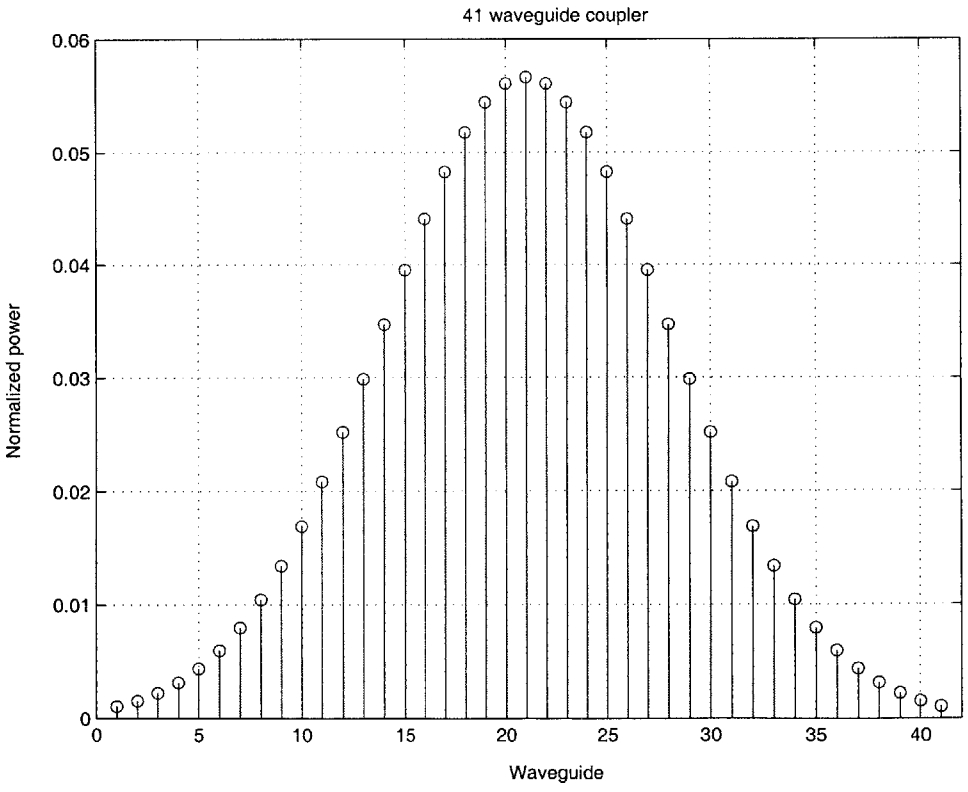

Figure 4-7: Output of 41 waveguide coupler for Gaussian profile

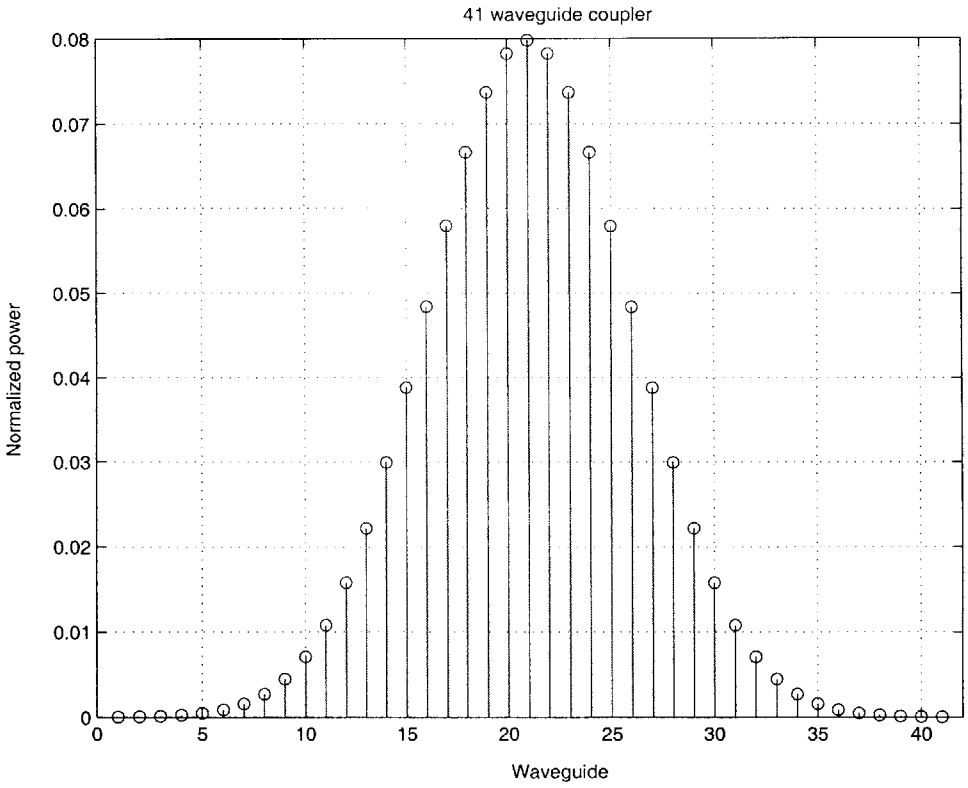

Figure 4-8: Output of 41 waveguide coupler for improved Gaussian profile

where i is the waveguide number from 1 to 41 and the numerical constant is for normalization (power conservation).

This profile was very promising; however, since it represents a power distribution, the associated field is perhaps a bit too wide. We want the electric field to follow a profile similar to the one in Figure 4-7 so that the wings are not too cut off. Therefore, we made a second Gaussian shape, shown in Figure 4-8, which is the square of the shape in Figure 4-7:

$$p = \frac{1}{12.5326} e^{-\frac{(i-21)^2}{50}} \tag{4.29}$$

Figure 4-9 shows the coupling coefficients for the 41 waveguide uniform coupler, as well as the two Gaussian couplers. We see that the Gaussian coefficients follow the same basic pattern as the uniform ones, increasing farther away from the center before finally dropping off slightly. As expected, the Gaussian pattern does not need as much coupling as the uniform pattern (and the thinner Gaussian requires less than the wider Gaussian), but it is interesting that the Gaussian coefficients are still more similar to the coefficients for a uniform profile than to the Bessel function case.

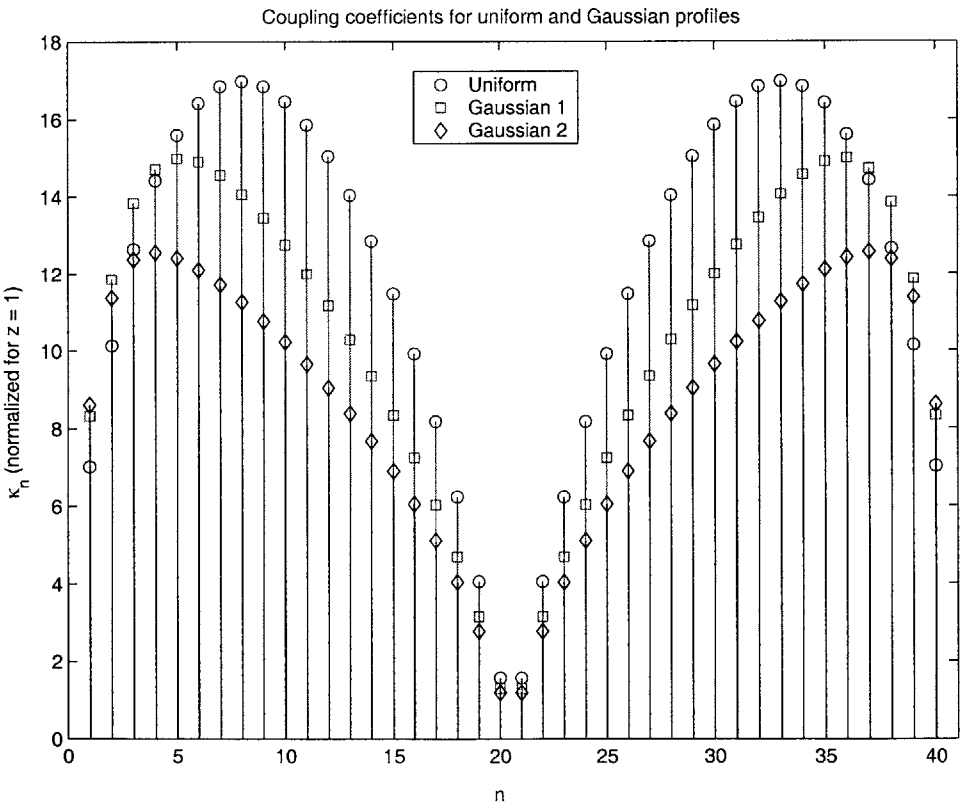

Figure 4-9: Coupling coefficients for 41 waveguide uniform and Gaussian couplers

4.2.5 Phase of the waveguide couplers

We must also consider the phase of the waveguide couplers. Equation (4.18) shows that there is a phase shift of j, or $\pi/2$ radians, between a guide and the neighboring guides which it excites. For the 41 waveguide Gaussian coupler of Figure 4-8, the phase is depicted in Figure 4-10. The slope of $\pi/2$ is apparent on both sides of the center waveguide. This phase profile unfortunately creates a problem for the design of an AWG. We want the center wavelength to be completely in phase at the entrance to the output coupler so that it will focus to the center output waveguide. To accomplish this, we could simply remove the phase of the coupler by propagating the light through different lengths in a uncoupled region. However, the other wavelengths

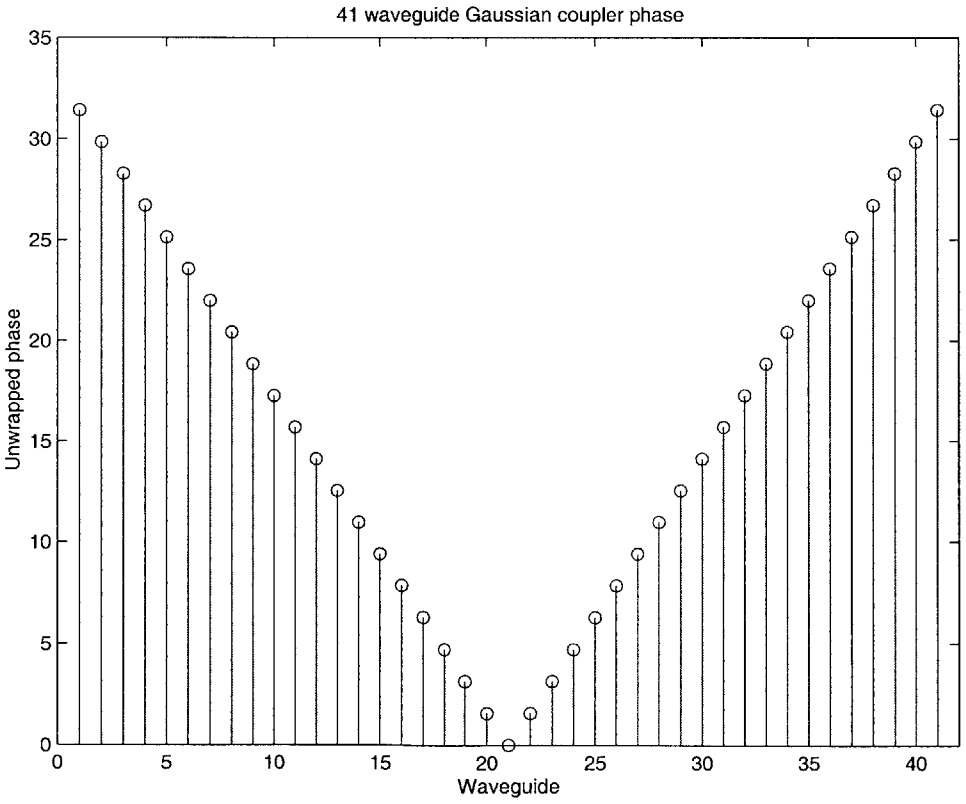

Figure 4-10: Phase of 41 waveguide Gaussian coupler

would exit with a phase that was still proportional to the phase in Figure 4-10. But we want these wavelengths to contain a linear (asymmetric) phase profile in order to focus them to different output waveguides.

One simple way to fix this problem is to excite the waveguide coupler not at the center waveguide, but at one of the end waveguides. In this case the phase will grow by $\pi/2$ in neighboring waveguides, but this effect will only propagate in one direction, creating a line antisymmetric about the center waveguide as desired. In order to achieve this phase, we must redesign the Gaussian coupler for excitation from one of the sides. This makes the choice of coupling coefficients more difficult because we can no longer take advantage of symmetry. Still, we were still able to

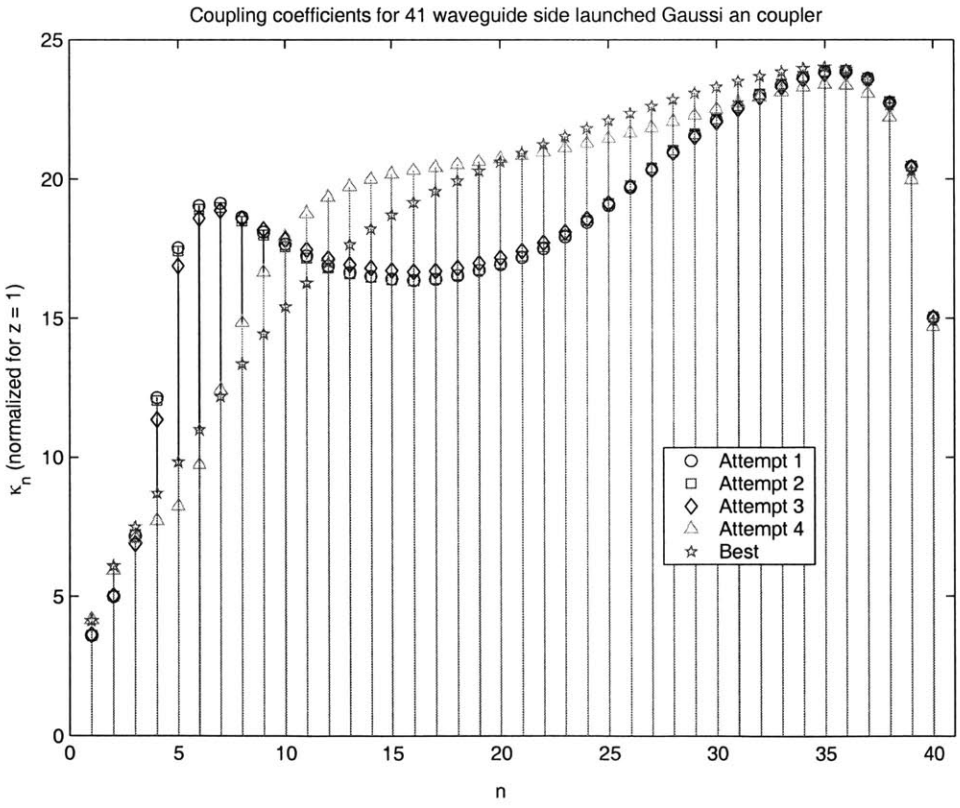

Figure 4-11: Coupling coefficients for 41 waveguide side launched Gaussian coupler

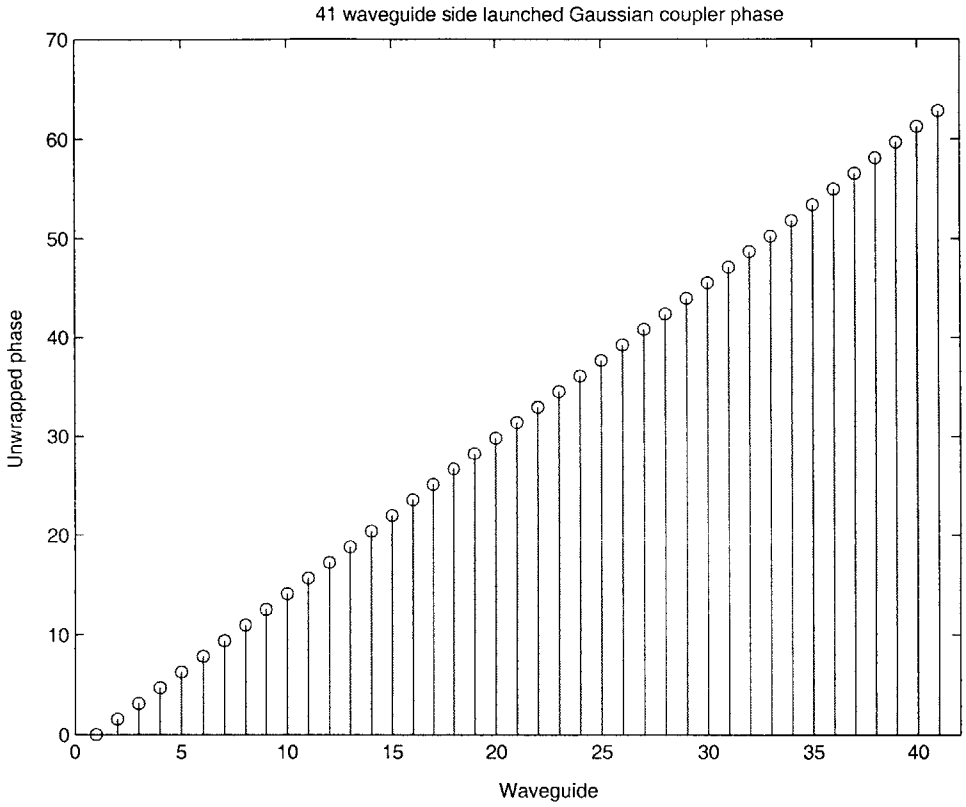

Figure 4-12: Phase of 41 waveguide side launched Gaussian coupler

achieve a 41 waveguide side launched Gaussian coupler. The coupling coefficients are pictured in Figure 4-11. Notice that many sets of coupling coefficients gave a Gaussian profile. However, only one gave the desired phase response, pictured in Figure 4-12. The others had slight problems in the small numbered waveguides, nearest the input, because the field amplitude had gone below zero, giving the desired power but the incorrect phase. The correct answer was found by tweaking the previous, slightly incorrect set of coupling coefficients and inputting them into the Newton's method routine until a new, better solution was found. We see that the correct solution looks qualitatively the best, as it is the smoothest and looks the most like the profiles of Figure 4-9.

4.2.6 Waveguide spacings required for coupling

Using (4.11), we calculated the distances D between the centers of neighboring waveguides required to achieve the 41 waveguide side launched Gaussian coupler. We assumed a 0.225 μm waveguide, which is single mode and as close to the second mode cutoff width of 0.24364 μm as we can get using 0.025 μm separation in FDTD. The calculation was performed for coupler lengths of 27.25 μm and 111.5 μm, the lengths we found for the 1×8 and 8×8 MMI couplers. The results are pictured in Figure 4-13. We see that for such high index contrast and confined modes, the spacing must be small. Even worse, though, are the very slight changes in spacing required from

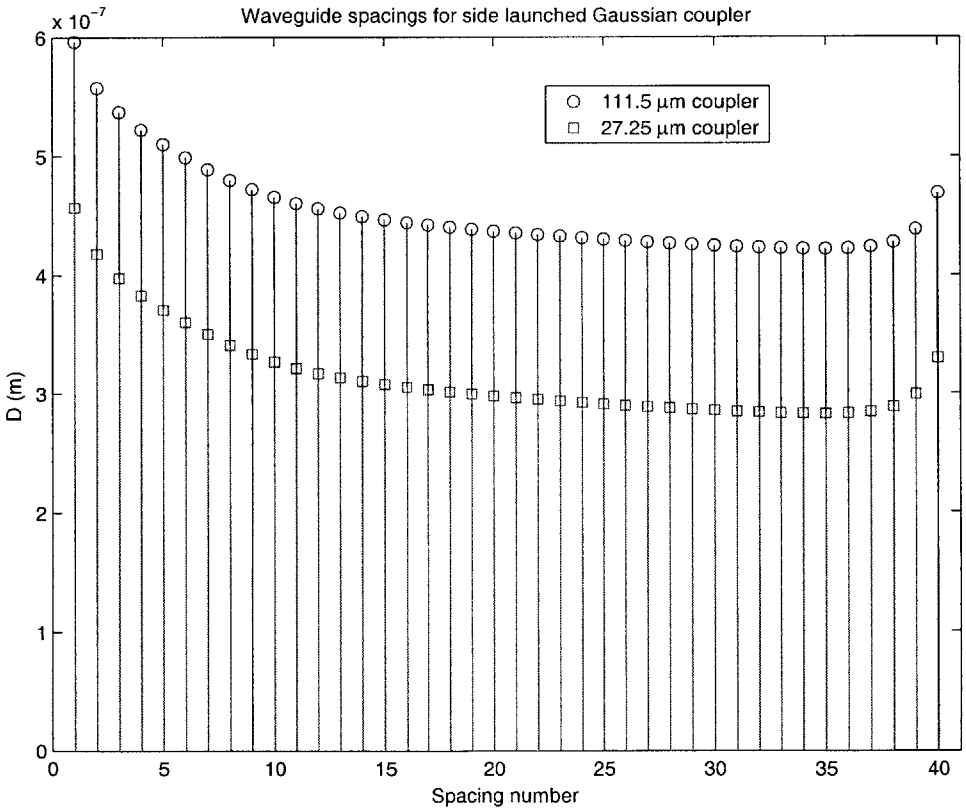

Figure 4-13: Waveguide spacings D required to create 41 waveguide side launched Gaussian coupler

one pair of waveguides to another. Therefore, even though the coupled waveguides work very well in theory, they may be hard to fabricate. We ignore this problem, however, and assume that it can be solved by using thinner, less confining waveguides or a lower, but still high, index contrast system. Otherwise, the coupling will differ from the ideal case.

4.3 Output coupler for new AWG

Unfortunately, the output coupler for this new AWG cannot be constructed like the input, using coupled modes. Consider a profile generated across coupled waveguides some distance L from the excitation of one of the waveguides. If the complex conjugate of the profile is taken, the process will be reversed. All of the light will return to the original waveguide in the same distance L. This behavior is a direct consequence of (4.12). Taking the derivative in the $-z$ direction,

$$\frac{d\mathbf{a}}{d(-z)} = -\frac{d\mathbf{a}}{dz} = -\mathbf{K}\mathbf{a} \tag{4.30}$$

Thus going in the $-z$ direction switches the sign of the $\pi/2$ phase change in $\mathbf{K}$. We can then conclude that the mode amplitudes at $z = -L$ are the complex conjugates of those at $z = L$. Since the profile at $z = -L$ converges to one waveguide at $z = 0$, we conclude that the conjugates of the profile at $z = L$ will also converge back to the original waveguide in distance L. This property is shown for the 41 waveguide side launched Gaussian coupler in Figure 4-14(a).

Unfortunately, it does not seem to be simple to add a phase to the mode amplitude profile to make it converge to a different waveguide. Figure 4-14(b) shows the result when a linear phase of slope $3\pi/40$ is added in addition to the complex conjugation. It is clear that the light focuses to the second to last waveguide more than the others, but it is far from 100%. Furthermore, the output appears at a smaller z than the

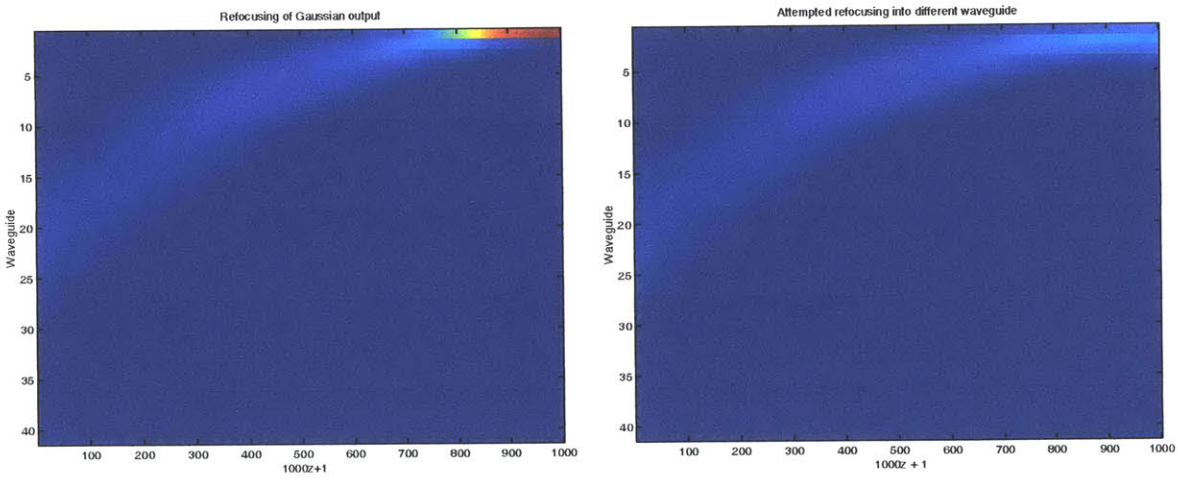

Figure 4-14: (a) Refocusing of Gaussian output into original waveguide, (b) Attempted refocusing into different waveguide

simple reversal. Thus, we cannot remove the signal focused here without affecting the operation of the coupler beyond that z. Even if we were to somehow engineer 100% power into different output waveguides as a function of phase, we would also need to ensure that this effect happened at the same z. For these reasons, we quickly abandoned the idea of using a waveguide coupler as the output of our AWG. It may be possible to do so, but the waveguide coupler physics do not suggest an immediate solution.

4.3.1 Free space (star) coupler

Since we decided not to use some sort of waveguide coupler for the AWG output, and since we wanted to avoid the disadvantages of MMIs, we instead chose to use a traditional free space, or star, coupler for this purpose. Although the star coupler has not yet been designed, we will discuss briefly the basic principles behind the device and outline its design.

Fourier transforms and the far field

The star coupler is basically a free propagation region, where the walls of the structure are far enough away that reflections do not significantly affect the device operation. Consider a field profile $u(x_0, 0)$ at the input of the star coupler, $z = 0$. The output of the star coupler is in the far field, so the field profile there is given by the Fraunhofer diffraction equation [10]:

$$u(x,z) = \frac{nj}{\lambda_0 z} e^{-jnk_0 z} e^{-jnk_0 x^2/2z} \int_{-\infty}^{\infty} u(x_0, 0) e^{jnk_0 x x_0/z} dx_0 \tag{4.31}$$

Notice that the integral in (4.31) defines a Fourier transform, so the far field is, within scaling factors and phase, the Fourier transform of the input. The most basic description of the star coupler's role in an AWG assumes that we have a Gaussian profile across the array waveguides. Since the Fourier transform of a Gaussian is a Gaussian, and since we put the array waveguide output along a spherical surface, we expect the Gaussian at that point to focus into a smaller Gaussian in the center output waveguide, just as if it had passed through a lens. When we put a linear phase across the array waveguides, Fourier transform theory tells us that the output will be shifted in space, i.e., into a different output waveguide. This is the reason we required a linear phase in the waveguide coupler.

Although we have been able to achieve a Gaussian profile in the mode amplitudes of the array waveguides, this is not the same as a Gaussian profile across the whole array. The individual mode shapes also matter. If the output of the array is merely a Gaussian envelope multiplied by the cosinusoidal mode amplitudes, the output will be a spatial convolution of the transforms of these two components. This causes sidelobes, or extra diffraction orders, to appear in the output. We wish to avoid sidelobes, because all power that goes into them is lost. To counteract this problem, we place the array guides as close to one another as possible at the entrance to the star coupler. Figure 4-15 shows the resulting field assuming 0.225 μm waveguides and no spacing between the waveguide cores. The figure was obtained by simply

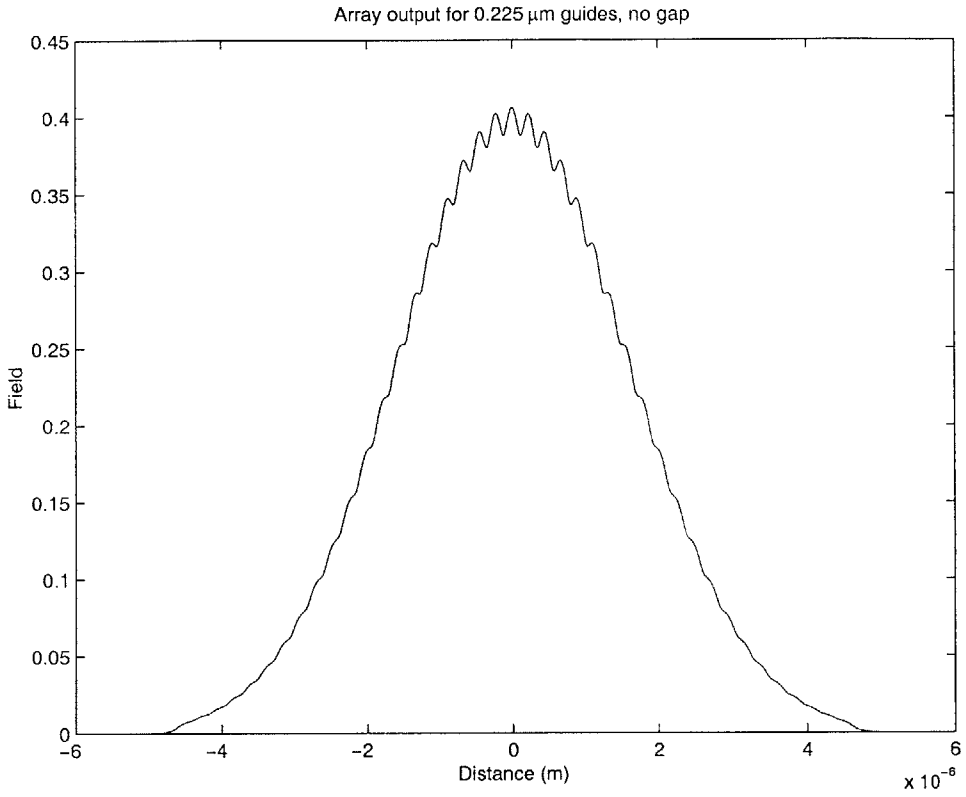

Figure 4-15: Input to star coupler for 0.225 μm array waveguides, no gap in between guides

summing the waveguide modes. In reality, no matter how sharply we approach the star coupler, there will be some coupling between the array waveguides, changing this profile. (Such coupling may actually be useful in its own right, however, as we mention in Chapter 5.) The profile is mostly Gaussian and should produce small sidelobes when Fourier transformed. Using thinner, less confining waveguides could improve this profile or allow the same profile to be created with more space between the guides. However, such a choice may increase loss in the array bends. Another option is to taper the waveguides as they approach the star coupler, reducing the coupling and simplifying the layout.

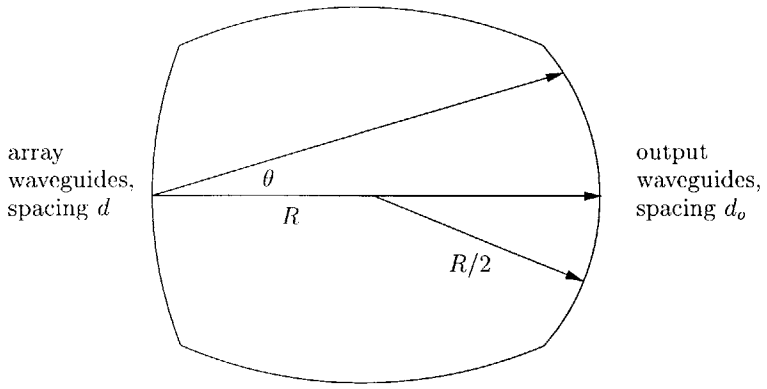

Figure 4-16: Rowland circle geometry for free space/star coupler

Detailed description of star coupler [18], [34]

A more detailed description of the star coupler in an AWG is given by Madsen and Zhao [18] and Smit [34]. This description is useful in the design of the coupler. We present some of the important results, modifying them if necessary for the coupled mode AWG.

Most star couplers used in AWGs are based on the geometry of the so-called Rowland circle, which is pictured in Figure 4-16. In this geometry, the array waveguides are spaced a distance d apart on a circle of radius R centered at the middle output waveguide. The output waveguides, on the other hand, are spaced distance d_o apart on a circle of radius $R/2$. It can be shown that the path length difference from two adjacent array waveguides to a particular output waveguide is given by $d \sin\theta$, where θ is the angle from the middle of the array to the output waveguide. In order to get constructive interference at the desired output guide, the phase arising from the length difference in the array must counteract the phase of the Rowland circle, as well as the phase from the input waveguide coupler:

$$-\frac{\pi}{2} - n_1 k_0 d \sin\theta = -n_{eff} k_0 \Delta L + 2\pi m \tag{4.32}$$

Here, $-\pi/2$ is the phase difference accumulated in the waveguide coupler[2], n_{eff} is the effective index of the array guides, and n_1 is the index of the core material which makes up the free space region.

The frequency response of the system can be understood better using a cascade of transfer matrices for each element. Madsen and Zhao [18] do this and show that the frequency response for output n can be expressed as the product of an envelope function and an integral. The envelope function is $U_g(nd_o)$, the far field of the array guide mode given by (4.31) and evaluated at output n. It causes nonuniformity between the outputs, something which is not easily predicted just by looking at the interference condition (4.32). The integral is a convolution of the output waveguide mode and the so-called array factor. This array factor contains information about both the Gaussian coefficients of the input coupler and the phases of (4.32). It determines the demultiplexing properties of the AWG. In addition, the overlap between the array factor and the output mode determines the loss due to mode mismatch and diffraction into sidelobes. We have already seen that creating a smooth Gaussian of the correct width across the array waveguides will help with these problems.

It turns out that the modes of the array and output waveguides can in fact be approximated by a Gaussian shape with an effective radius w_0 [18], [34]:

$$u(x) = \frac{1}{w_0\sqrt{\pi}} e^{-(x/w_0)^2} \tag{4.33}$$

The magnitude in the far field is given by

$$|U(x')| = \frac{n_1}{\lambda R} e^{-(n_1 \pi w_0 x'/\lambda R)^2} \tag{4.34}$$

[2]We now return to the correct sign for the phase between coupled waveguides and assume that the originally excited guide has the longest path length in the array.

Thought of as the far field of the array waveguides, this expression helps us find the envelope function for the frequency response or, alternatively, the loss uniformity of the output channels:

$$L_u \approx -10 \log_{10} e^{-2(n_1 \pi w_0 x'/\lambda R)^2} \approx 8.7 \left(\frac{n_1 \pi w_0 x'}{\lambda R} \right)^2 \tag{4.35}$$

Thought of as the far field of the output waveguides, (4.34) gives the required width of the Gaussian distribution produced by the input waveguide coupler.

Beginning the design

To begin the star coupler design, we first chose the array and output waveguide size to be 0.225 μm. As mentioned earlier, this is the largest size allowed by our usual 0.025 μm FDTD spacing which is still single mode. We then chose the output waveguide separation to avoid coupling. Conservatively, we selected $d_0 = 1.125$ μm. In reality, the separation will have to be one which is allowed by the interference condition (4.32). We then decided on a maximum nonuniformity of 3 dB and used (4.35) with $x' = 4d_0$ to determine a good radius for the coupler. We found $R = 16.4$ μm. The next step is to use coupled modes to create a new Gaussian input profile which looks like (4.34) for the chosen waveguide width and coupler radius. (This step is unnecessary in traditional star coupler AWGs because the input and output star couplers are symmetric.) Our current Gaussian profiles are unacceptable because we did not think carefully enough about the star coupler at the time of the earlier work. Finally, we must choose ΔL to achieve the phase match condition of (4.32).

The star coupler is difficult to simulate in FDTD simply because it is hard to properly set up the dielectric file. Some Matlab code has been written to automatically create the profile. It is incomplete, but a good start. Figure 4-17 shows the dielectric structure generated by this code. It is clear that the output side of the coupler has not been generated, though this is not necessary to begin testing the fo-

cusing properties. In addition, we used 0.2 μm waveguides spaced 0.225 μm apart at the coupler boundary instead of the desired structure of 0.225 μm waveguides spaced 0.225 μm apart. In the latter case, the waveguides are too closely spaced to make a meaningful figure. Even worse, the waveguides are hard to excite properly when spaced so closely, so some more cleverness might be needed to properly set up a star coupler simulation.

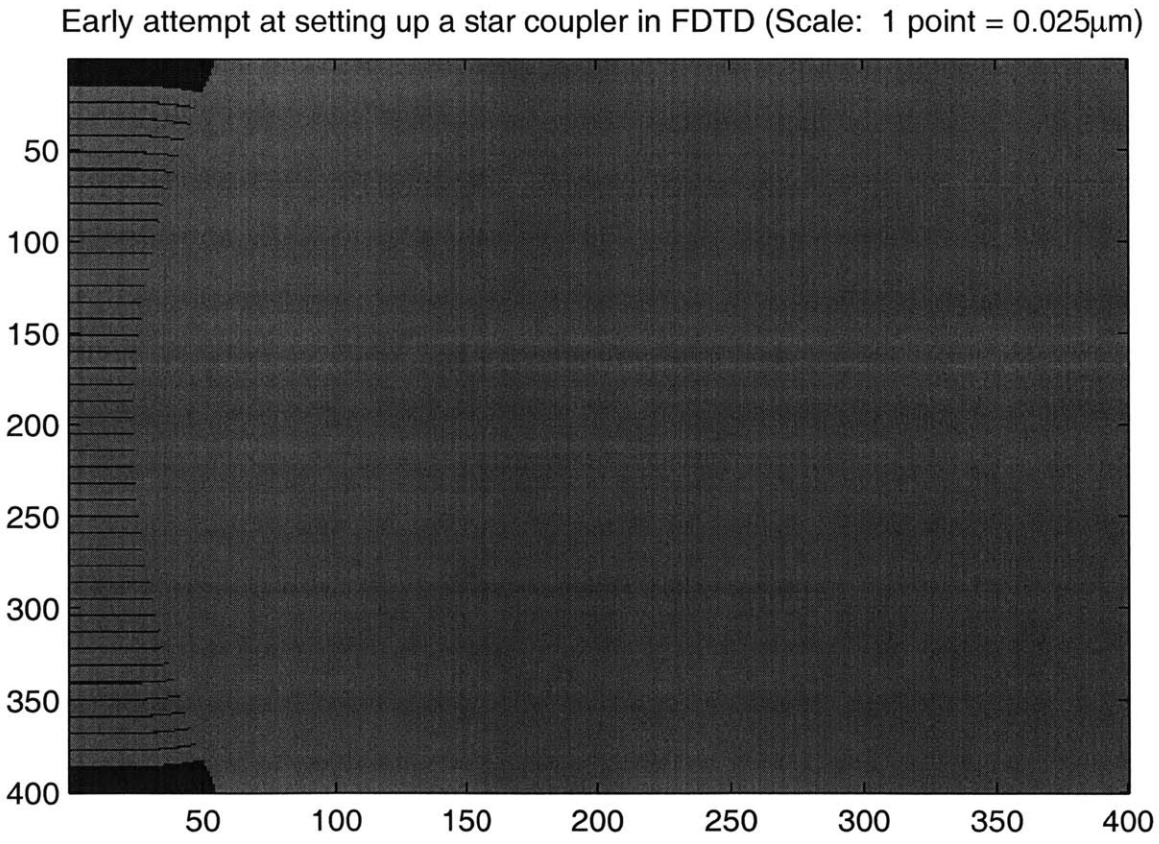

Figure 4-17: Early attempt at setting up a free space/star coupler in FDTD

Chapter 5

Conclusions and Future Work

In this thesis, we have described preliminary work in the design of an arrayed waveguide grating for high index contrast integrated optics. Unfortunately, the work is not complete; however, this is an accident of discovery more than error. Originally, the focus of the work was to be the design of a high index contrast AWG based on multimode interference couplers. After much effort, designs were produced for a 1×8 MMI splitter and an 8×8 MMI combiner. Then array waveguide lengths were selected to test the overall performance of the device. Unfortunately, the device performance, while qualitatively correct, did not meet the desired crosstalk requirements. Up until this point, most of the work, that described in Chapter 3, was fairly standard. Even though MMI-based AWGs are rare, they are still well known [16], [27], [28], [32], [24]. Here it seemed like the real work was about to begin, a project to identify and correct the problems with the MMI AWG in the particular setting of high index contrast, something which had not been tried before. Some progress was made identifying these problems and possible solutions before this part of the project was abandoned.

The focus of this thesis changed considerably when H. Haus proposed the idea for an input based on coupled mode theory. This type of input shows great promise for use in an AWG. By proper choice of the coupling coefficients, we can engineer any desired amplitude profile at a given distance in the coupler. This way, even though the output must be a star coupler, we can optimize the fields to produce the

smoothest Gaussian input to that device, reducing the sidelobes of the final response. In addition, the phase of the coupled waveguide output is well known, which was not the case for the error-riddled MMI couplers. Thus this new input structure is very exciting and merits further study. However, work on it has only just begun, so we are unsure of the final results.

This concluding chapter will discuss some of the future work which can be done on this project. It will begin by analyzing some of the possible solutions to the MMI AWG. Even though that project has been largely abandoned, it is still interesting, and this author would like to see it pursued. The concluding pages will then discuss the future of the coupled mode AWG, as well as some other work which could be done on the coupled mode input structure.

5.1 Future of the MMI AWG

One of the fundamental problems of the MMI AWG is the presence of large phase errors in the MMI couplers for the desired index contrast and size. The simplest solution, and one which could be tried easily, is to determine the actual output phases of the MMI couplers. The array waveguides could then be optimized around the actual phases of the MMI outputs, not the theoretical values.

A more radical solution to the phase error problem is to try to reduce or eliminate the phase errors. One solution is to implement the MMI couplers using an optimized low index contrast structure, as shown in [12]. This technique would theoretically cancel the two sources of phase error, but only within the approximations made in that paper. Of course, it would also increase the size of the MMI and also produce possible fabrication and mode matching problems at the interface to the (necessarily high index contrast) array guides. Another solution attempts to do away with phase errors in an even more fundamental manner. A so-called graded-index MMI, or GRIN MMI, eliminates phase errors by creating an index gradient in the slab region [17]. The index gradient is chosen so that the propagation constants have the quadratic dependence

$$\beta_0 - \beta_i = \frac{i(i+d)\pi}{(d+1)L_\pi} \tag{5.1}$$

It turns out that self-imaging will occur for even positive integer d. Inserting $d = 2$ recovers the relation (3.8) we had before for an MMI. Here, however, the index profile makes the relation exact, and no phase errors will occur. Such a structure seems tempting; however, the technical issues in creating this structure may be even worse than creating an optimized low contrast, step-index MMI.

A final step which could help the performance of the MMI AWG is to consider the windowing function interpretation of the array. It is impossible to create a larger number of waveguides in the array for a given number of output channels, but it is possible to shape the windowing function to reduce the sidelobes. Paiam and MacDonald [27] discuss their efforts to use a 1×8 nonuniform MMI splitter as the input device for their AWG. M. Bachmann, et. al., show that nonuniform power splitting is obtained by selecting the input position such that some of the self images overlap. In fact, the 1×8 uniform symmetric, shorter coupler we used in our AWG design is also a consequence of the same phenomenon [2]. Other researchers have found ways to make tunable MMIs by varying the index halfway along the MMI, but these designs are not yet ready for structures with eight outputs [14].

For the purpose of the MMI AWG, the splittings should be picked such that the transmitted power increases with d_j and then decreases again. This splitting gives a window shape which is closer to the popular Hamming window or the Gaussian of a free space AWG than the original even splitting. Paiam and MacDonald [27] found that use of a nonuniform input coupler helped reduce sidelobes by 15 dB. However, as a tradeoff, the main lobe was widened, causing different crosstalk problems. This problem was solved by cascading two multiplexers together for optimal performance. Such a solution would be the easiest to implement on the current high index contrast AWG problem. In fact, a nonuniform MMI was attempted in the last days of the MMI AWG work; however, the imaging was still too poor, suggesting that perhaps

this enhancement is not enough by itself. Still, it is the most promising lead for the MMI AWG.

5.2 Future of the coupled mode AWG

The coupled mode AWG still has much development to be done. The largest piece of the puzzle is the development of the output coupler. While some of the groundwork for the output coupler has been laid in Chapter 4, there is clearly much more work to be done on this device. One specific improvement beyond the standard design procedures could be the use of directional coupling in the array waveguides to create a rectangular envelope for the frequency response, thus making the loss more uniform across the output. This phenomenon was discovered by Okamoto [25] and explained by Song [38], [29], [37] using the Bessel function relations for coupled waveguides derived in Chapter 4. In addition, the input array of coupled waveguides has yet to be simulated in FDTD. Only a true electromagnetic simulation will show that this coupling technique is viable, especially over several wavelengths. Furthermore, it would be interesting to see how the desired profiles change when rounding error is introduced to the very finely spaced coupling coefficients we have derived. Finally, the input and output sections of the device need to be joined, with appropriate array guide delays, to test the overall AWG. Only at this stage will we know for sure how useful it is to be able to create an arbitrary profile in the array waveguides.

In order to achieve these goals, it might be useful to find ways to make the simulations more accurate. For this thesis, we have always ignored the three dimensional structure of the waveguides. We treated the problem as two dimensional without making any of the correct approximations, such as the effective index approximation. In the future, it may be wise to invoke this approximation. In addition, the continuation scheme to transfer fields from one simulation to another could be optimized. For instance, it would be better to transfer only the part of the field which matches the array waveguide mode from the input simulation to the output. Even more accuracy could be gained by simulating the array guides directly. Finally, it would be useful

to create a quick way to find mode overlaps over a range of frequencies, though this may be difficult because the desired mode profile changes with frequency. However, these accuracy improvements may be delayed until rough solutions are found.

5.3 More on waveguide couplers

The coupled mode AWG is not the only possible use for the large N waveguide couplers we have designed. As mentioned in Chapter 4, the uniform distribution coupler could be useful for many applications requiring uniform splitting. One of these is the MMI AWG considered in Chapter 3. An 8 waveguide coupler could replace the input MMI, though the output MMI was by far the worse problem. It is also possible to engineer an optimal nonuniform splitter for the input to an AWG with an MMI output. This splitting would not look Gaussian across the waveguides as they are laid out, but taken in order of d_j, it would look somewhat Gaussian.

There are also important theoretical questions surrounding the large N waveguide couplers. One problem is the choice of coupling coefficients. In this thesis we fairly easily found the coupling coefficients for N as large as 41, even in the antisymmetric side launched case. But the problem became much more difficult as N increased. It would be interesting to find a more sophisticated algorithm to find the coupling coefficients for a desired profile. A large part of such an algorithm should be more advanced convergence methods.

Finally, an interesting problem is the description of what actually happens in a large N waveguide coupler. For example, consider the paraxial wave equation in two dimensions [10]:

$$\frac{\partial^2 u}{\partial x^2} - 2jk\frac{\partial u}{\partial z} = 0 \tag{5.2}$$

It can be discretized in the x direction as

$$\frac{u_{n+1} - 2u_n + u_{n-1}}{(\Delta x)^2} - 2jk\frac{\partial u_n}{\partial z} = 0 \tag{5.3}$$

which becomes

$$\frac{\partial u_n}{\partial z} = \frac{-j}{2k(\Delta x)^2}(u_{n+1} - 2u_n + u_{n-1}) \tag{5.4}$$

This has some terms which are oddly similar to (4.18), suggesting that the infinite coupled waveguide problem can be somehow linked to paraxial free space propagation with an index profile (alternatively, a k profile) along the x direction. A finite number of waveguides would simply introduce reflecting boundary conditions into the free space problem. However, the term containing u_n is not found in the coupled waveguide problem, and we have not accounted for the mode profiles of the waveguides, so it is uncertain if the similarities between the two problems are meaningful. Nevertheless, it would be interesting to see if there is in fact some way of linking the theory of coupled waveguide optics and that of paraxial free space optics, especially since we use both in our coupled mode AWG design.

Appendix A

Sample FDTD Input File

The following is a sample input file for the FDTD code. It is constructed from the file for the 1×8 MMI coupler; however, the file has been somewhat edited to save space. For example, though nine sources may be declared, only two are listed. Needless to say, the file would not actually run correctly. All distances are in microns.

```
1 0                     ! indicates TE polarization
.025 .025 1.1           ! Discretization: dx, dy,
                          cfl (used to determine time step)
1360 480 16 16          ! Grid/PML Size (in grid points): x size,
                          y size, PML x size, PML y size
0 60000 1000 0          ! Time stepping: start, finish, interval
                          to write out field, take snapshot at end?
10 1.46 1               ! number of dielectric objects, background
                          refractive index, write out dielectric?
0 1                     ! object type (0 = rectangle),
                          number of repetitions
-1. 5.7 2. 0.6 3.5      ! corner x coordinate, corner y coordinate,
                          length 1, length 2, index of refraction
1 0  0 1                ! directions of length 1 and length 2
                          (here, +x and +y)
0 1                     ! More rectangles...we omit most of them
                          to save space.  Note that the format can
                          differ for other object types.
1. 1. 27.25 10. 3.5
1 0  0 1
2. 1. 2. 1. fldfile     ! Initial field (here, none): xmin, xmax,
                          ymin, ymax, filename
```

```
9                          ! number of sources (eight are dummies
                             for mode overlap)
2 1.55 500 3. 0. 0.        ! time type (2 = gaussian pulse), center
                             wavelength, pulsewidth in fs,
                             number of pulsewidths to shift, group
                             delay (times c), phase delay (times c)
.1 .1 5.5 6.5              ! Source cross section: xmin, xmax,
                             ymin, ymax
-2 0 2. 1. 1.              ! spatial type (-2 = from modesolver),
                             excitation direction (1 = forward),
                             effective index guess (ignored for -2
                             spatial type, needed otherwise),
                             amplitude if Gaussian spatial type,
                             width if Gaussian spatial type
0 1.55 500 3. 0. 0.        ! time type 0 = none, a dummy source for
                             mode overlap.  We omit the other seven.
33.9 33.9 1.125 2.125
-2 0 2. 1. 1.
0 5                        ! Monitored points (here, none): number,
                             interval to write out field
2. 1. 2. 1.                ! Energy box (here, none): xmin, xmax,
                             ymin, ymax
8                          ! number of mode overlaps
33.9 33.9 1.125 2.125 2 1  ! Mode overlap: xmin, xmax, ymin, ymax,
                             source to take frequency from,
                             write out finmag and finphs?
33.9 33.9 2.375 3.375 3 1  ! We omit the other six.
1.544 1.556 .00001         ! Spectrum: start wavelength, end
                             wavelength, spacing
9                          ! number of power fluxes
.2 .2 5.5 6.5              ! Power flux: xmin, xmax, ymin, ymax
33.9 33.9 1.125 2.125      ! We omit the other seven.
```

Bibliography

[1] M. Bachmann, P.A. Besse, and H. Melchior. General self-imaging properties in $N \times N$ multimode interference couplers including phase relations. *Applied Optics*, 33(18):3905–3911, June 1994.

[2] M. Bachmann, P.A. Besse, and H. Melchior. Overlapping-image multimode interference couplers with a reduced number of self-images for uniform and nonuniform power splitting. *Applied Optics*, 34(30):6898–6910, October 1995.

[3] J. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200, 1994.

[4] P. Bienstman. *Rigorous and efficient modelling of wavelength scale photonic components.* Ph.D. thesis, Ghent University, Department of Information Technology, May 2001.

[5] O. Bryngdahl. Image formation using self-imaging techniques. *Journal of the Optical Society of America*, 63(4):416–419, 1973.

[6] W.C. Chew. Electromagnetic theory on a lattice. *Journal of Applied Physics*, 75(10):4843–4850, May 1994.

[7] J.H. den Besten, M.P. Dessens, C.G.P. Herben, X.J.M. Leijtens, F.H. Groen, M.R. Leys, and M.K. Smit. Low-loss, compact, and polarization independent PHASAR demultiplexer fabricated by using a double-etch process. *IEEE Photonics Technology Letters*, 14(1):62–64, January 2002.

[8] C. Dragone. An $N \times N$ optical multiplexer using a planar arrangement of two star couplers. *IEEE Photonics Technology Letters*, 3(9):812–815, September 1991.

[9] P.L. Hagelstein. *Introduction to Numerical Simulation in Electrical Engineering.* Massachusetts Institute of Technology, course notes for subject 6.673, 2000.

[10] H.A. Haus. *Waves and Fields in Optoelectronics.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1984.

[11] H.A. Haus. Personal communication, September 2002.

[12] J.Z. Huang, R. Scarmozzino, and R.M. Osgood, Jr. A new design approach to large input/output-number multimode interference couplers and its application to low-crosstalk WDM routers. *IEEE Photonics Technology Letters*, 10(9):1292–1294, September 1998.

[13] L.C. Kimerling and K. Wada. Chip scale ultra-compact WDM demultiplexer with electrical conversion and amplification. MIT Materials Processing Center research proposal, December 2001.

[14] J. Leuthold and C.H. Joyner. Multimode interference couplers with tunable power splitting ratios. *Journal of Lightwave Technology*, 19(5):700–707, May 2001.

[15] H. Li, C. Lee, W. Lin, and S. Didde. 8-wavelength photonic integrated 2×2 WDM cross-connect switch using $2 \times N$ phased-array waveguide grating (PAWG) multi/demultiplexers. *Electronics Letters*, 33(7):592–593, 1997.

[16] L.O. Lierstuen and A. Sudbø. 8-channel wavelength division multiplexer based on multimode interference couplers. *IEEE Photonics Technology Letters*, 7(9):1034–1036, September 1995.

[17] Y. Lin, J. Liu, and W. Wang. Analysis of multimode interference couplers with lateral graded-index profiles. In *Conference on Lasers and Electro-Optics/Pacific Rim 2001*, volume 2, pages 244–245, Chiba, Japan, July 2001.

[18] C.K. Madsen and J.H. Zhao. *Optical Filter Design and Analysis: A Signal Processing Approach.* John Wiley & Sons, Inc., New York, New York, 1999.

[19] C. Manolatou. *Passive Components for Dense Optical Integration Based on High Index-Contrast.* Ph.D. thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, May 2001.

[20] C. Manolatou. Personal communication, November 2002.

[21] D. Marcuse. *Light Transmission Optics.* Van Nostrand Reinhold Company, New York, New York, 1972.

[22] D. Marcuse. *Theory of Dielectric Optical Waveguides.* Academic Press, New York, New York, 1974.

[23] MIT Microphotonics Center website. `http://web.mit.edu/mphotonics/www`.

[24] Y. Moreau, K. Kribich, P. Coudray, P. Etienne, and J. Galy. Code division multiple access with MMI mineral organic circuits. In *International Conference SPIE: Photonics West*, San Jose, California, January 2002.

[25] K. Okamoto. *Fundamentals of Optical Waveguides.* Academic Press, San Diego, California, 2000.

[26] A.V. Oppenheim, R.W. Schafer, and J.R. Buck. *Discrete-Time Signal Processing, Second Edition.* Prentice-Hall, Inc., Upper Saddle River, New Jersey, 1999.

[27] M.R. Paiam and R.I. MacDonald. Design of phased-array wavelength division multiplexers using multimode interference couplers. *Applied Optics*, 36(21):5097–5108, July 1997.

[28] M.R. Paiam and R.I. MacDonald. A 12-channel phased-array wavelength multiplexer with multimode interference couplers. *IEEE Photonics Technology Letters*, 10(2):241–243, February 1998.

[29] M.Y. Park and G.H. Song. A design formula for the optimized star coupler utilizing directional coupling in the input array. Submitted to *Optics Express*, 2003.

[30] M. Popović. *Air Trenches for Dense Silica Integrated Optics.* S.M. thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, February 2002.

[31] M. Popović. Personal communication, May 2003.

[32] J. Porque, P. Coudray, R. Charters, K. Kribich, P. Etienne, and Y. Moreau. WDM based on multimode interference-coupler built in an organic-inorganic material. *Optics Communications*, 183:45–49, September 2000.

[33] M. Smit. New focusing and dispersive planar component based on an optical phased array. *Electronics Letters*, 24(7):385–386, March 1988.

[34] M. Smit and C. van Dam. PHASAR-based WDM-devices: Principles, design, and applications. *IEEE Journal of Selected Topics in Quantum Electronics*, 2(2):236–250, June 1996.

[35] L.B. Soldano and E.C.M. Pennings. Optical multi-mode interference devices based on self-imaging: Principles and applications. *Journal of Lightwave Technology*, 13(4):615–627, April 1995.

[36] S. Somekh, E. Garmire, A. Yariv, H.L. Garvin, and R.G. Hunsperger. Channel optical waveguide directional couplers. *Applied Physics Letters*, 22(5):46–47, 1973.

[37] G.H. Song. *Principles of Photonics.* Volume 1, Chapter 6. In preparation for publication, 2003.

[38] G.H. Song and M.Y. Park. Bessel-function analysis of the optimized star coupler for uniform power splitting. Submitted to *Journal of Lightwave Technology*, 2003.

[39] Y. Tachikawa, Y. Inoue, M. Ishii, and T. Nozawa. Arrayed-waveguide grating multiplexer with loop-back optical paths and its applications. *Journal of Lightwave Technology*, 14(6):977–984, 1996.

[40] H. Takahashi, S. Suzuki, K. Kato, and I. Nishi. Arrayed-waveguide grating for wavelength division multi/demultiplexer with nanometre resolution. *Electronics Letters*, 26(2):87–88, January 1990.

[41] E.R. Thoen, L.A. Molter, and J.P. Donnelly. Exact modal analysis and optimization of $N \times N \times 1$ cascaded waveguide structures with multimode guiding sections. *IEEE Journal of Quantum Electronics*, 33(8):1299–1307, August 1997.

[42] R. Ulrich and G. Ankele. Self-imaging in homogeneous planar optical waveguides. *Applied Physics Letters*, 27(6):337–339, 1975.

[43] K. Wada and H.A. Haus. Photonic integrated circuits by high-index contrast optics. Massachusetts Institute of Technology, presentation for subject 6.191, October 2001.

[44] K.S. Yee. Numerical solution of the initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, AP-14:302–307, May 1966.

[45] M. Zirngibl, C. Joyner, and B. Glance. Digitally tunable channel dropping filter/equalizer based on waveguide grating router and optical amplifier integration. *IEEE Photonics Technology Letters*, 6(4):513–515, 1994.

[46] M. Zirngibl, C. Joyner, L. Stulz, U. Koren, M. Chien, M. Young, and B. Miller. Digitally tunable laser based on the integration of a waveguide grating multiplexer and an optical amplifier. *IEEE Photonics Technology Letters*, 6(4):516–518, 1994.