STARS

Retrospective Theses and Dissertations

Spring 2001

Design and applications of volume holographic optical elements

Steven C. Dunn

University of Central Florida

STARS Citation

Dunn, Steven C., "Design and applications of volume holographic optical elements" (2001). *Retrospective Theses and Dissertations*. 1163.
https://stars.library.ucf.edu/rtd/1163

DESIGN AND APPLICATIONS OF VOLUME HOLOGRAPHIC OPTICAL ELEMENTS

By
Steven C. Dunn

2001

UCF

DESIGN AND APPLICATIONS

OF VOLUME HOLOGRAPHIC

OPTICAL ELEMENTS

by

Steven C. Dunn
B.S. Drexel University, 1990
M.S. Drexel University, 1992

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the School of Electrical Engineering and Computer Science
in the College of Engineering
at the University of Central Florida
Orlando, Florida

Spring Term
2001

Major Professor: M.G. Moharam

ABSTRACT

Volume gratings were studied both theoretically and experimentally in order to design and analyze practical volume holographic optical elements. The diffraction of finite (Gaussian) beams by transmission gratings is investigated. A closed form solution for the off-Bragg profiles is obtained, and the distortion in the resulting profiles is analyzed. Criteria for when the finiteness of the beam has a minimal effect on the angular selectivity response are given. A model is developed to investigate the cross talk in multiplexed gratings Experimental verification in $LiNbO_3$ and PTR glass is demonstrated. The diffraction efficiency and output profiles of finite beams diffracted from reflection gratings on and off-Bragg are investigated. Using these beams decreases the diffraction efficiency, and the diffracted profiles are shown to be distorted, with higher distortion off-Bragg. The fidelity is shown to improve with increasing grating strength. The output profiles of diffraction by overlap gratings is investigated. Lateral shift and beam distortion are observed. When the diffraction efficiency reaches 83% the diffracted beam is shown to split into multiple sections. A method is given to obtain a diffracted beam with any desired profile. It is shown that most of the previously mentioned effects can be corrected, however the splitting of the beam at 83% cannot. Multiplexing in reflection gratings is also investigated. A narrow transmission peak is shown to occur midway between the Bragg wavelengths of the individual gratings. By

adding more gratings, additional peaks are obtained. Similar devices are designed using thin film stacks. These devices have many advantages, including simpler scaling and no sidelobes. The number of transmission peaks is found using multi-mirror Fabry-Perot interferometer theory, and expressions are given for their location and width. The effects of including the variation in average index when modeling gratings recorded in the presence of absorption is presented. These effects are seen to increase as the grating type goes from un-slanted transmission to un-slanted reflection. Finally, the characteristics of plane waves incident on over-modulated transmission gratings are investigated. It is shown that they will experience one of three types of off-Bragg behavior.

To my children, Christopher and Julia who

make every day a good one …

To their mother, Cassie whose unending love

and strength have carried me

throughout this endeavor – this is

but a small token of my love…

And to my parents, who gave me the ability.

ACKNOWLEDGMETNTS

I would like to express my sincere thanks to professor Jim Moharam for his guidance, support and patience over the past five years. His remarkable understanding of diffraction and electromagnetics has helped immeasurably in my pursuit of this degree. In addition to his technical help, the overall direction he has given has allowed me to grow considerably. For this I am deeply grateful.

I would also like to thank the faculty members who have agreed to serve on my dissertation committee. Their many questions, comments and discussions have led to significant improvements in this work.

Additional thanks go to my colleagues and friends at the school of optics, for countless discussions and ideas. Among them, special mention is due to Don Jacob, whose influence exists throughout this work. I truly hope you reach all of your goals... when you do, I'll be there for the ride!

Finally, special thanks go to my wife Cassie, without whose patience, support, and encouragement this work would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS

$\hat{A}_R$ Reference Beam Amplitude Distribution (x-y coordinate system)

A_R Reference Beam Amplitude Distribution (r-s coordinate system)

$\hat{A}_{R0}$ Recording Reference Beam Amplitude Distribution

$\hat{A}_S$ Signal Beam Amplitude Distribution in (x-y coordinate system)

A_S Signal Beam Amplitude Distribution in (r-s coordinate system)

$\hat{A}_{S0}$ Recording Signal Beam Amplitude Distribution in (x-y coordinate system)

c Speed of Light ($m \cdot s^{-1}$)

c_R Reference Beam obliquity factor

c_S Signal Beam obliquity factor

d Grating Thickness (m)

f_x Spatial Frequency x Component (mm^{-1})

f_z Spatial Frequency z Component (mm^{-1})

F Finesse factor of a Fabry-Perot interferometer

g Normalized geometry factor

$\vec{K}$ Grating vector

K_x The x component of $\vec{K}$

L	Grating Period (m^{-1})
L_x	X component of the Grating Period
P_d	Diffracted Power (W)
P_t	Transmitted Power (W)
R	Reference Beam Amplitude
R_0	Reference Beam Initial Amplitude
r_1	Reflection coefficient (field) of mirror no. 1
r_2	Reflection coefficient (field) of mirror no. 2
R_1	Reflection coefficient (intensity) of mirror no. 1
R_2	Reflection coefficient (intensity) of mirror no. 2
S	Signal Beam Amplitude
S_0	Signal Beam Initial Amplitude
m	Arbitrary constant
n	Index of refraction
n_e	Effective index of a thin film stack
n_1	Index modulation
Y	Admittance
Y_e	Equivalent Admittance of a thin film stack
α	Absorption Coefficient ($n \cdot m^{-1}$)
α_1	Amplitude of the absorption modulation
β	Average Propagation Constant (m^{-1})
δ	Arbitrary phase difference

ε_r	Relative Dielectric Permitivity
ε_{r0}	Average Relative Dielectric Permitivity
ε_{r1}	Amplitude of the Relative Dielectric Permitivity Modulation
∇^2	Laplacian operator
ϕ	Grating Slant Angle (radians)
ϕ_e	Effective phase of a thin film stack
γ	Normalized grating strength parameter
η	Diffraction Efficiency
η_m	Characteristic Impedance
κ	Coupling Coefficient
λ	Wavelength of light (m)
σ	Gaussian Beam Width (m)
σ_m	Conductivity of the medium
σ_0	Average Conductivity
σ_1	Amplitude of the Conductivity Modulation
ϑ	Dephasing parameter
μ	Permeability
ω	Angular frequency (radians)
v	Grating Strength Parameter
θ	Arbitrary angle (°)
θ_i	Angle of incidence (°)

θ_d Angle of diffraction (°)

θ_t Angle of transmission (°)

$\bar{\rho}$ Reference Beam Propagation Vector

ρ_x The x component of ρ

$\bar{\sigma}$ Signal Beam Propagation Vector

σ_x The x component of σ

τ Normalized dephasing parameter

ξ Dephasing Parameter

$\mathfrak{I}$ Fourier transform

$\mathfrak{I}^{-1}$ Inverse Fourier transform

CHAPTER 1

INTRODUCTION

Early optical systems were composed entirely of elements whose properties are based on the reflection and refraction of light. These included such elements as mirrors, lenses and prisms. Due to the inherent simplicity of these devices, people have understood for centuries how to manipulate light with them, as well as how to combine them into compound optical systems. Although this required considerable insight into the properties of light, it was not until the wave nature of light was recognized that many common effects involving interference and diffraction could be properly explained. Once the notion of light as a wave was finally accepted, tools and techniques were developed to use the wave properties of light in many novel ways.

As our understanding of these principles increased, and our ability to manufacture optics to better tolerances improved, an additional class of optical elements was developed which has had an enormous impact on the design of optical systems. These are the diffractive optical elements (DOEs). DOEs are optical components which use diffraction itself to manipulate the wavefront of the incident light. The addition of a DOE to an optical system can improve performance in many ways such as reducing

chromatization effects, reducing system weight, and even replacing one or more traditional components [1]. In addition, DOEs can be used to create a number of devices which can not be made with refractives such as beam shapers, filters, head up displays and novel optical interconnects.

Diffractive optical elements can be of many different types including diffraction gratings, surface relief structures and computer generated holograms. Similar effects may also be obtained from stacks of alternating high and low index layers. This work will however, be limited to one type, the volume holographic optical element (VHOE). These are simply diffractive elements produced by recording the interference pattern of two coherent beams in a thick material [2]. By a judicious selection of the object and reference beams, one can store a pattern that, upon reconstruction, will yield such things as a focusing wave, diverging wave or even multiple waves. These would be equivalent to a positive lens, negative lens, and beam splitter respectively. The selection of other forms for the object and reference beams can lead to even more novel structures.

By their very nature volume holograms require thick material. This thickness has two advantages. First it leads to higher diffraction efficiencies for small index modulations (more light coupled into the diffracted beam). Second, it adds a requirement that there be phase matching between the incident beam and the element in order for the diffracted beam to contain significant energy. This allows the devices to be used as excellent narrowband angular and wavelength filters.

The objective of this work is to investigate the analysis, design, and applications of VHOEs. All devices will be based on the simple diffraction grating, so after a brief

2

discussion of Bragg chapter two will give a review of the mathematics of grating theory. Of the many ways to describe grating behavior (modal analysis, rigorous electromagnetic theories...) the coupled wave theory as presented by Kogelnik [3] will be the primary method used in this study, due to its simplicity and applicability. Since this is a one-dimensional theory (assuming incidence by plane waves on infinite gratings) and many of the problems in VHOE design and analysis are two dimensional; it must either be expanded to two dimensions, or used in combination with a spectral decomposition of the input beam. Each of these methods will be used in this work (the choice depending on the problem at hand), therefore a description of each method is given in chapter 2.

Chapter 3 briefly reviews the materials of volume holography, describing their main characteristics, along with listing some of their strong and weak points. This is intended to give the reader an introduction into some of the choices from which he can select his material for volume holography.

Chapters 4 and 5 are devoted to diffraction of finite beams in transmission and reflection holograms respectively. They will attempt to cover in some detail the main characteristics of these devices. The chapters are organized around their uses as angular and wavelength filters; covering angular selectivity profiles (both theoretical and experimental), along with the effects that these devices have on the diffracted and transmitted beam profiles (specifically gaussian).

Chapters 6 and 7 are devoted to overlap (finite) gratings, beginning with a general solution for gaussian beams in these types of structures. Using this solution, diffracted and transmitted beam profiles are given for both Bragg and off-Bragg incidence. Then in

3

chapter 7 these results are used to develop a general methodology for the design of VHOEs with particular desired characteristics. Three particular examples will be given which include an edge emitter and a spot generator.

Chapter 8 is dedicated to multi-line filters, and begins by discussing multiplexed transmission and reflection gratings. The emphasis is on effects that result specifically from the multiplexing process itself. In some cases multiplexing closely spaced (in wavelength) reflection gratings can give a narrowband wavelength filter which operates as a transmission filter. This filter is discussed in detail, along thin-film alternatives which have similar properties, and which are somewhat easier to manufacture.

Chapter 9 considers two additional issues of volume holography, absorption during the recording process and overmodulation. The first of these is present in all types of holograms, and can significantly affect the angular selectivity profiles. The second (overmodulation) results in regions which exhibit non-traditional angular and wavelength selectivity profiles. Operating in these areas gives some rather unexpected results, which could lead to some novel devices.

The work finishes with a chapter devoted to conclusions and a discussion of possible future work

CHAPTER 2

BACKGROUND

In order to analyze and design volume holographic optical elements (VHOEs) an adequate model of their behavior is needed. Such a model allows the simulation of different conditions and an investigation into the workings of any structure. All elements in this work are based on volume diffraction gratings, thus solutions for these devices will be derived. We begin by considering the required phase relationships for light to diffract from a volume structure, then use coupled wave analysis to develop grating solutions. This is followed by a discussion of plane wave decomposition, which can be combined with coupled wave analysis to handle more complex problems. The chapter ends by covering two dimensional coupled wave analysis, which is used like plane wave decomposition, but adds insight to the solution, as well as into the design problem.

2.1   Derivation of the Bragg Condition

Throughout this work many references will be made to the incident beam being either on or off-Bragg. In this context Bragg refers to a particular angle of incidence

(wavelength) that maximizes the diffracted intensity at the incident wavelength (angle) for a given structure. These conditions will now be developed, based on a description by Syms [1]. The geometry of the problem is shown below in Figure 2.1.. We start with part a of the figure, where we have an array of scatterers separated by a distance L_x. An infinite plane wave (of wavelength λ in all regions) is incident on the scatters at an angle θ_i.
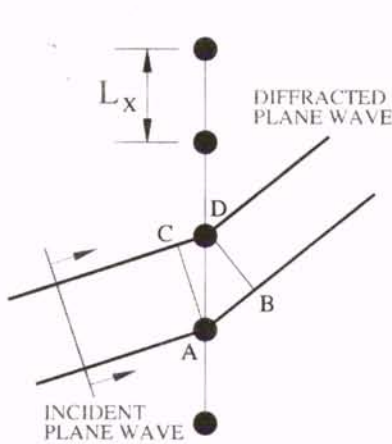
Figure 2.1a. Thin Grating   Figure 2.1b. Thick Grating
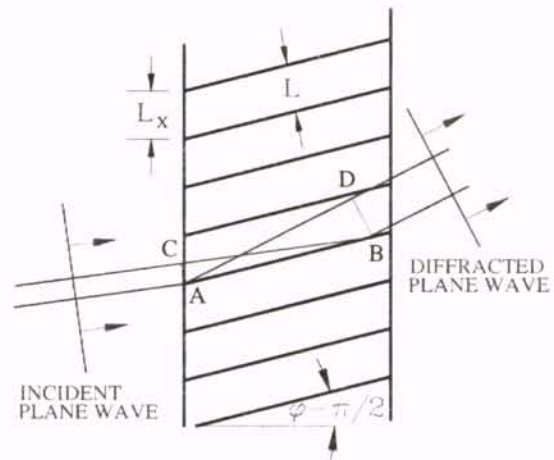
The effect of each scatterer is to distribute the incident light in all directions, resulting in a field that at any point will be the sum of the fields from each scatter. Since the incident light is monochromatic and spatially coherent, the periodicity of the structure will force additional interference effects. In particular, a plane wave output will be significant only if the fields from adjacent scatters add in phase. We see that this happens if the difference between paths AB and CD is an integral number of wavelengths. This path difference is given by $AB - CD = L_x(\sin\theta_d - \sin\theta_i)$. Forcing this to be an integer number of wavelengths and solving for the diffracted angle gives:

6

$$\sin\theta_d = \sin\theta_i + \frac{m\lambda}{L_x} \tag{2.1}$$

In this equation (which is the well known grating equation) m is the order number and refers to any integer. The equation states that the grating will distribute the light from an incident beam into a number of orders, each traveling in a direction determined by the properties of the grating along with the wavelength and angle of incidence. Since this relationship is based strictly on the periodicity of the grating, it is valid regardless of the thickness of the structure.

The above discussion arose from the fact that constructive interference will only occur when all individual waves add up in phase. If we consider part b of the figure where the grating has become thick, and treat each grating line as a partial mirror, then we get an additional requirement. Not only must the components from different fringes add up in phase, but those components scattered from different portions of the same fringe must also add constructively. Thus paths such as EF and HG must also be equal. This will occur if either $\theta_d = \theta_i$ or $\theta_d = 2(\phi - \pi/2) - \theta_i$ (where ϕ is the grating slant angle, and is measured perpendicular to the fringes). Inserting the first of these requirements into equation 2.1 (with $L_x \equiv L/\cos(\phi - \pi/2)$) leaves m=0. This describes the propagation of the transmitted portion of the incident wave. Inserting the second of these requirements into equation 2.1 and using some simple algebra along with a bit of trigonometry yields:

$$2nL\cos(\phi - \theta) = m\lambda \tag{2.2}$$

7

which is one form of the well known Bragg equation for volume gratings. From this equation we see that each order requires a different angle of incidence in order to propagate. Deviation from that angle will result in a significant intensity reduction for that order. This is an inherent angular selectivity, which makes these devices appropriate for filtering applications. The amount of reduction as a function of this deviation depends on the parameters of the structure, and equations for it will be derived later.

2.2   Modeling Grating Behavior - Coupled Wave Theory of Thick Holograms

In this section the properties of volume diffraction gratings are analyzed using the Coupled Wave Analysis presented by Kogelnik [2], due to its simplicity and physical intuitiveness. In spite of this simplicity, it gives quite accurate predictions of the diffraction process. It is based on an assumption of monochromatic light incident on an infinite grating at or near the Bragg angle. Only two significant light waves are assumed present, the incoming reference wave and an outgoing signal wave. For this to be true the grating must be thick enough that all but one order violates the Bragg condition strongly and thus can be ignored. The grating under study is shown below in Figure 2.2.
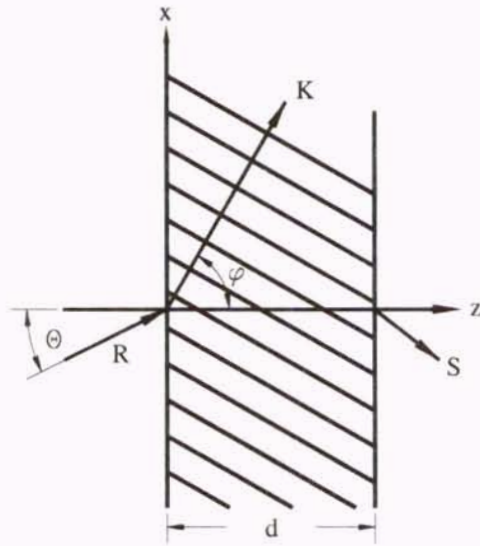
Figure 2.2. Grating under Study

As mentioned above the grating is assumed to extend to infinity in the x and y directions, and to have thickness d in the z direction. Its fringes are described by way of a K vector, whose direction is at an angle φ perpendicular to the fringe lines and whose magnitude is given by:

$$\left|\vec{K}\right| = \frac{2\pi}{L}$$

(2.3)

(where L is the period of the grating). A transverse electric (TE) plane wave is incident on the material at an angle θ, given by:

9

$$E = R_0 e^{-j(\vec{\rho} \cdot \vec{r})}$$

$$\text{where:} \tag{2.4}$$

$$\vec{\rho} = \begin{bmatrix} \rho_x \\ \rho_y \\ \rho_z \end{bmatrix} = \beta \begin{bmatrix} \sin(\theta) \\ 0 \\ \cos(\theta) \end{bmatrix}$$

For simplicity the average dielectric constant will be taken as the same inside and outside the medium (allowing us to neglect refraction). Using this model of a grating, expressions for the relative permittivity and the conductivity of the grating are written as:

$$\varepsilon_R = \varepsilon_{R0} + \varepsilon_{R1} \cos(\vec{K} \cdot \vec{r})$$
$$\sigma_m = \sigma_0 + \sigma_1 \cos(\vec{K} \cdot \vec{r}) \tag{2.5}$$

Since we assume the presence of only a reference wave and a signal wave in the medium the following form for the electric field is obtained:

$$E = R(z)e^{-j(\vec{\rho} \cdot \vec{r})} + S(z)e^{-j(\vec{\sigma} \cdot \vec{r})} \tag{2.6}$$

This is now inserted into the scalar wave equation:

$$\nabla^2 E + k^2 E = 0 \tag{2.7}$$

where k is obtained by inserting (2.5) into :

$$k^2 = \frac{\omega^2}{c^2}\varepsilon - j\omega\mu\,\sigma_m \qquad (2.8)$$

resulting in:

$$k^2 = \beta^2 - 2j\alpha\beta + 2\kappa\beta\left(e^{j(\vec{K}\cdot\vec{r})} + e^{-j(\vec{K}\cdot\vec{r})}\right) \qquad (2.9)$$

In this equation κ is a coupling coefficient which describes the coupling of light between the reference and signal waves, and is given by:

$$\kappa = \frac{1}{4}\left(\frac{2\pi}{\lambda}\frac{\varepsilon_1}{\sqrt{\varepsilon_0}} - \frac{j\mu c\sigma_1}{\sqrt{\varepsilon_0}}\right) \cong \left(\frac{\pi n_1}{\lambda} - j\frac{\alpha_1}{2}\right) \qquad (2.10)$$

The propagation vectors ρ and σ give the magnitude and direction of the reference and signal waves respectively. Since R always propagates, the magnitude of ρ is equal to β (the unmodulated propagation constant for the material), and the direction of ρ is given by the reference wave propagation direction. The magnitude and direction of σ are determined by the grating as follows. From the grating equation we have:

$$\sigma_x = \rho_x - K_x \qquad (2.11)$$

It is further assumed that:

$$\sigma_z = \rho_z - K_z \qquad (2.12)$$

11

Any corrections required from this assumption will be accounted for in the magnitude of R(z) and S(z). In addition, if the condition: $|\vec{\rho}| = |\vec{\sigma}| = \beta$ is met then we get the well known Bragg condition of thick holograms:

$$\cos(\phi - \theta) = \frac{K}{2\beta} \qquad (2.13)$$

Clearly this condition is limited to very weak modulation, where the propagation constant can be accurately taken as equal to its average value in the medium. This relationship is easily described by Figure 2.3 below. When the magnitude of σ is equal to β (as in part a of the figure), the Bragg condition is satisfied and the diffracted wave will have its maximum amplitude. In part b of the figure the reference wave vector was changed, σ ≠ β, and the Bragg condition is no longer satisfied.
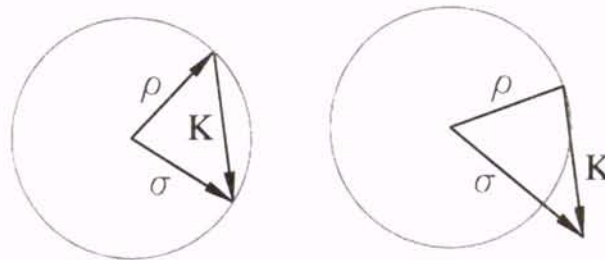
Figure 2.3. Graphical representation of the Bragg Condition

Equations (2.6) and (2.9) are now inserted into (2.7), which results in the following set of coupled wave equations:

$$\frac{d^2 R(z)}{dz^2} - j2\rho_z \frac{dR(z)}{dz} - j2\alpha\beta R(z) + 2\kappa\beta S(z) = 0$$

$$(2.14)$$

$$\frac{d^2 S(z)}{dz^2} - j2\sigma_z \frac{dS(z)}{dz} - j2\alpha\beta S(z) + (\beta^2 - \sigma^2)S(z) + 2\kappa\beta R(z) = 0$$

We now solve for the field amplitudes. In doing so the second derivatives of R and S are neglected, which should be valid if the coupling is slow on the order of a wavelength. Additionally any waves propagating well off Bragg are ignored. The resulting coupled wave equations are:

$$c_R \frac{dR(z)}{dz} + \alpha R(z) = -j\kappa S(z)$$

$$c_S \frac{dS(z)}{dz} + (\alpha + j\vartheta)S(z) = -j\kappa R(z)$$

where

$$(2.15)$$

$$\vartheta \equiv (\beta^2 - |\vec{\sigma}|^2)/2\beta \approx \Delta\theta \cdot K\sin(\phi - \theta_0) - \Delta\lambda \cdot K^2 \frac{\lambda}{4n\pi}$$

$$\text{and} \quad c_R = \frac{\rho_z}{\beta} = \cos\theta, \quad c_S = \sigma_z/\beta = \cos\theta - \frac{\kappa}{\beta}\cos\phi$$

In this expression c_R will always be positive, however the sign of c_S will depend on the hologram type. In transmission holograms (where $\phi > \pi/4$) c_S remains positive, however it becomes negative in reflection holograms ($\phi < \pi/4$). The characteristics of each hologram type vary in a smooth fashion as the slant angle (ϕ) is changed, however when

13

the sign of c_S changes (when φ crosses $\pi/4$), the hologram type changes abruptly and a correspondingly abrupt (and interesting) transition takes place in these characteristics.

The general solution to equation 2.15 is given below:

$$R(z) = r_1 \exp(\gamma_1 z) + r_2 \exp(\gamma_2 z)$$

$$S(z) = s_1 \exp(\gamma_1 z) + s_2 \exp(\gamma_2 z)$$

(2.16)

where r_1, r_2, s_1 and s_2 are constants that are determined by the boundary conditions. The values for γ_1 and γ_2 are found by plugging equation 2.16 into the wave equation. Doing this yields:

$$\gamma_{1,2} = -\frac{1}{2}\left(\frac{\alpha}{c_R} + \frac{\alpha}{c_S} + j\frac{\vartheta}{c_S}\right) \pm \frac{1}{2}\left[\left(\frac{\alpha}{c_R} - \frac{\alpha}{c_S} - j\frac{\vartheta}{c_S}\right)^2 - 4\frac{\kappa^2}{c_R\,c_S}\right]^{\frac{1}{2}}$$

(2.17)

By multiplying equation 2.15(a) by R* and 2.15(b) by S*, and adding the resulting equations to their complex conjugates the following energy balance expression is found:

$$\frac{d}{dz}\left(c_R\,RR^* + c_S\,SS^*\right) + 2\alpha\left(RR^* + SS^*\right) + j\left(\kappa - \kappa^*\right)\left(RS^* + R^*S\right) = 0$$

(2.18)

In the absence of absorption the last two terms reduce to 0 and we are left with $c_R\,RR^* + c_S\,SS^* = k$ where k is a constant. The value of this constant can be found by looking at the plane z=0. At this point S=0 so (assuming an incident uniform plane

14

wave) the constant becomes c_R. We can then define the diffraction efficiency (the fraction of incident wave power which is coupled into the diffracted wave) as:

$$\eta = \frac{|c_S|}{c_R} SS^*$$

(2.19)

The transmission and reflection cases are now considered separately (due to their differing boundary conditions) to solve equation 2.16.

Transmission Holograms

For transmission holograms the boundary conditions are $R(0)=1$, $S(0)=0$. Using these boundary conditions, expression for the diffracted field becomes:

$$S(\xi,v) = -j\left(\frac{c_R}{c_S}\right)^{\frac{1}{2}} \exp(-\alpha d/c_R) e^{\xi} \frac{\sin(v^2 - \xi^2)^{\frac{1}{2}}}{(1 - \xi^2/v^2)}$$

$$v = \kappa d \Big/ (c_R c_S)^{\frac{1}{2}}$$

(2.20)

$$\xi = \frac{1}{2}d\left(\frac{\alpha}{c_R} - \frac{\alpha}{c_S} - j\frac{\vartheta}{c_S}\right)$$

In these equations, v is a grating strength parameter, ξ a dephasing parameter (measuring the deviation from Bragg) and α is the readout absorption coefficient. Often this is small enough to be neglected. Doing so, and redefining ξ and v to be real yields the following equations for the diffracted field and the diffraction efficiency:

$$S = -j\left(\frac{c_R}{c_S}\right)^{\frac{1}{2}} \frac{\sin\left(v^2 + \xi^2\right)^{\frac{1}{2}}}{\left(1 + \xi^2/v^2\right)} e^{-j\xi}$$

$$v = \pi n_1 d \Big/ \lambda\left(c_R c_S\right)^{\frac{1}{2}} \qquad \xi = \vartheta d/2c_S \qquad (2.21)$$

$$\eta = \sin^2\left(v^2 + \xi^2\right)^{\frac{1}{2}} \Big/ \left(1 + \xi^2/v^2\right)$$

Note that at zero dephasing (incidence at the Bragg angle) the diffraction efficiency becomes: $\eta = \sin(v)^2$. Thus we can see that when v is a multiple of $\pi/2$ the "On Bragg" diffraction efficiency reaches 100%.

Figure 2.4 below shows the diffraction efficiency (normalized to its value on Bragg) plotted vs the dephasing parameter ξ for three values of the grating strength parameter.

16
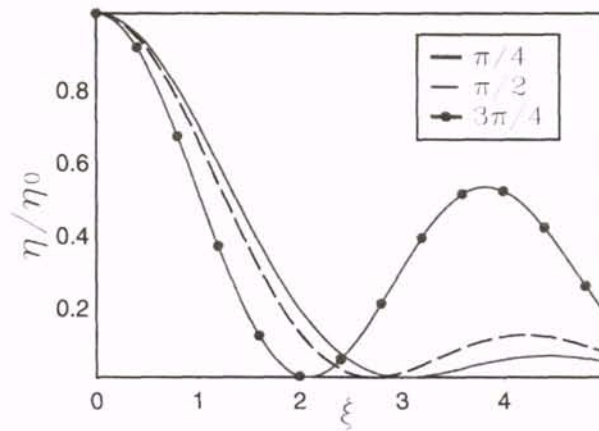
Figure 2.4. Normalized Diffraction Efficiency vs. Dephasing - Transmission Holograms

Reflection Holograms

The major difference between transmission and reflection holograms is that in reflection holograms the diffracted wave comes out on the same side as the incident wave ($c_S < 0$). The boundary conditions then become $R(0)=1$ and $S(d)=0$, where again d is the thickness of the hologram. Inserting these in the coupled wave equations and applying the same simplifications as before yields the following equation for the signal beam and diffraction efficiency at the output (the plane z=0) in the absence of absorption:

17

$$S(\xi,v)=\left(\frac{c_R}{c_S}\right)^{\frac{1}{2}}\bigg/\left\{j\xi/v+\left(1-\xi^2/v^2\right)^{\frac{1}{2}}\coth\left(v^2-\xi^2\right)^{\frac{1}{2}}\right\}$$

$$v=j\pi n_1 d\big/\lambda\,(c_R\,c_S)^{\frac{1}{2}}\quad \xi=-\vartheta d/2c_S \tag{2.22}$$

$$\eta=1\bigg/\left\{1+\left(1-\xi^2/v^2\right)\big/\sinh^2\left(v^2-\xi^2\right)^{\frac{1}{2}}\right\}$$

In these equations v has been redefined from that used in equation 2.21 to remain purely real. The figure below shows these equations plotted as a function of the dephasing parameter ξ for three values of the grating strength parameter v.
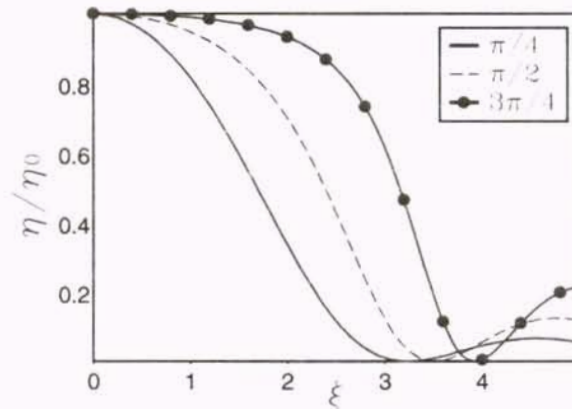
Figure 2.5. Normalized Diffraction Efficiency vs. Dephasing - Reflection Holograms

Since a principal use for these devices is an optical filter, an important parameter is the FWHM of the main lobe with deviations in angle or wavelength. Table 2.1 below shows these values for each hologram type with various efficiencies.

18

Bragg Efficiency	Transmission		Reflection	
	$\Delta\theta_{FWHM}$	$\Delta\lambda_{FWHM}/\lambda$	$\Delta\theta_{FWHM}$	$\Delta\lambda_{FWHM}/\lambda$
50 %	$0.866\dfrac{L}{d}$	$0.866\dfrac{L}{d\tan\theta}$	$1.12\dfrac{L}{d\tan\theta}$	$1.12\dfrac{L}{d}$
100 %	$0.8\dfrac{L}{d}$	$0.8\dfrac{L}{d\tan\theta}$	$1.97\dfrac{L}{d\tan\theta}$	$1.97\dfrac{L}{d}$

Table 2.1. Full width at half maximum for various grating types

Due to the presence of the tangent function in several of these expressions, we can see that for typical angles of incidence transmission gratings are better suited as angular filters, while reflection gratings are more suited to wavelength filters.

2.3   The Angular Spectrum of Plane Waves

As noted in the introduction, coupled wave analysis will be the main tool with which we will study VHOEs. This is, however a one-dimensional theory, which assumes plane wave incidence and allows the amplitude of the waves to vary in only one direction. All practical gratings and replay beams are of course finite, thus assuming plane waves is an approximation (although usually a reasonable one).

In order to get more accurate results, or to consider problems which are inherently two-dimensional (such as investigating beam profiles), a means of extending coupled wave analysis to two dimensions is needed. The simplest way to do this is with the angular spectrum of plane waves. This has been well described in the literature [3-6], so only a brief review will be given here. For simplicity this discussion will be restricted to waves traveling in the plane y=0. Generalization to three dimensions can be found in the references.

For this analysis the following two-dimensional transform pairs will be used:

$$f(x,z) = \int_{-\infty}^{\infty} F(f_x, f_z) e^{-j2\pi(f_x x + f_z z)} \, df_x \, df_z$$

and

$$F(f_x, f_z) = \int_{-\infty}^{\infty} f(x,z) e^{+j2\pi(f_x x + f_z z)} \, dx \, dz$$

(2.23)

where f_x and f_z are the spatial frequencies (in lines per millimeter) of the distribution along the x and z axes respectively. These correspond to the traditional frequency component in one-dimensional Fourier analysis.

Assume an arbitrary wave is traveling in the x-z plane as in part (a) of Figure 2.6.
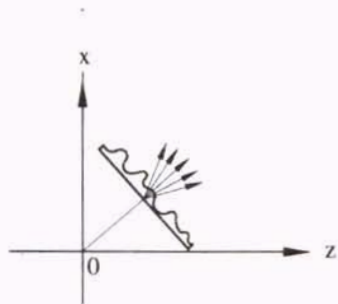
Figure 2.6a. The angular spectrum
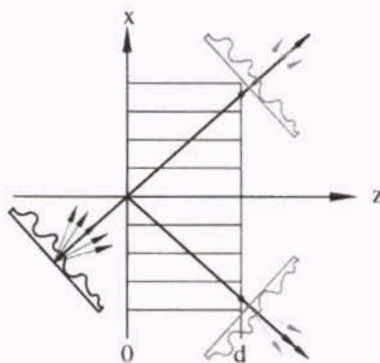
Figure 2.6b. Propagation of the angular spectrum through a grating

If the complex field of this wave at z=0 is denoted by E(x,0), and its Fourier transform by $\tilde{E}(f_X,0)$, then we can write E in terms of the inverse Fourier transform of $\tilde{E}$ as:

$$E(x;0) = \int_{-\infty}^{\infty} \tilde{E}(f_X;0) e^{-j2\pi f_X x} \, df_X \qquad (2.24)$$

Recalling that a plane wave traveling at an angle θ to the z axis can be written as:

$$A e^{-j\frac{2\pi}{\lambda}(\sin\theta \, x + \cos\theta \, z)} \qquad (2.25)$$

we can recognize equation (2.24) as the superposition of a number of plane waves, each propagating in a direction given by $\sin^{-1}(\lambda f_x)$, with a complex amplitude given by $\tilde{E}(f_X,0)$ evaluated at $f_x = \sin\theta/\lambda$. The ability to decompose any function into a number

21

of plane waves in this manner gives the function $\tilde{E}(f_X,0)$ the designation "the angular spectrum" of the field $E(x,0)$.

Since our input is now just a number of plane waves, the diffracted and transmitted profiles can be calculated by applying coupled wave analysis to each of them. This will result in a transmitted and a diffracted wave for each component, which can be superposed to obtain profiles for both the diffracted and transmitted waves. This is described graphically in part b of Figure 2.6 above. In addition, a numerical integration across these profiles will yield diffraction efficiencies for the overall beam.

Most of the analyses presented in this work will consider the incident beam to have a Gaussian shape, since this is approximately the profile of a typical laser beam. It therefore seems appropriate to determine the angular spectrum of this distribution. This is performed as follows. The field distribution at the plane z=0 of a Gaussian beam traveling in a direction θ_0 can be written as:

$$E(x,0) = E_0 \, e^{-\left(\dfrac{x^2 \cos^2 \theta_0}{\sigma^2} + j k_0 X \sin \theta_0 \right)}$$

(2.26)

Using the Fourier transform expressions of equation 2.23, the amplitudes of the corresponding plane waves can be written as:

$$\tilde{E}(k_x) = \int_{-\infty}^{\infty} E_0 \, e^{-\left(\frac{x^2 \cos^2 \theta_0}{\sigma^2} + j k_0 x \sin \theta_0\right)} e^{j k_x x} \, dx$$

(2.27)

where $\quad k_x = 2\pi f_x$

where this expression is evaluated at $f_x = \sin\theta/\lambda$. This can be simplified by the use of the following identity:

$$\int_{-\infty}^{\infty} e^{-\left(a x^2 + b x\right)} dx = \sqrt{\frac{\pi}{a}} \, e^{\frac{b^2}{4a}}$$

Inserting this into equation 2.27 results in the following plane wave amplitudes:

$$\tilde{E}(k_x) = \frac{2 E_0 \sqrt{\pi}}{\sigma_k} e^{-\left(\frac{k_x - k \sin \theta_0}{\sigma_k}\right)^2}$$

(2.28)

where $\sigma_k = \dfrac{2 \cos \theta_0}{\sigma}$

These amplitudes are then inserted into equations 2.21 or 2.22 (depending on the hologram type) to obtain the transmitted and diffracted profiles and diffraction efficiency.

In the above derivation the decomposition was performed in the plane z=0. It is permissible however to decompose the beam in a direction perpendicular to its propagation. In this case the amplitude distribution changes slightly to

23

$$\tilde{E}(k_r) = \sigma \sqrt{\pi} E_0 \, e^{-\left(\frac{\sigma k_r}{2}\right)^2}$$

$$(2.29)$$

$$\text{where} \quad k_r = 2\pi f_r$$

Again this is evaluated at $f_r = \sin\theta/\lambda$, which is the spatial frequency in the direction perpendicular to propagation (r). Since these angles are measured from the direction of propagation they actually represent differential angles to that of the central spectrum. We can therefore consider this as additional angular dephasing (which may be positive or negative) for the corresponding angular spectrum component. In a typical analysis a total dephasing amount (determined by the sum of the central and differential angles) for each component would be inserted into equation 2.21 or 2.22 to find the response.

It is possible however to use this concept in a slightly different way in order to get a quick feel for how a given beam profile will change the filtering characteristics, or whether the diffracted beam will be highly distorted by a particular grating. This is because this "additional" dephasing is intrinsic to the particular beam. Thus, by calculating the intrinsic angular width of a particular beam one can see if this will comprise a large portion of the angular acceptance region of a given grating (the main lobe of its angular selectivity curve). By the term intrinsic angular width I am referring to the angle corresponding to the highest spatial frequency in the beam (given by equation 2.29). Technically equation 2.29 has an infinite width since a Gaussian distribution never drops to zero, however if we use the customary (albeit arbitrary) $1/e^2$ beam width criteria then this spatial frequency component is $f_{r-\max} = \sqrt{2}/\pi\sigma$. Using

24

$f_r = \dfrac{\sin(\theta)}{\lambda}$ and assuming a small angle yields an intrinsic angular width of $\theta_{max} = \dfrac{\sqrt{2}\lambda}{\pi\sigma}$.

Inserting this into the dephasing equations (2.21 or 2.22) gives the resulting intrinsic angular dephasing. For unslanted transmission and reflection gratings these can be approximated as:

$$\xi_{T-max} \approx \frac{\sqrt{2}\lambda d}{\sigma L}$$

and $\hspace{6cm}$ (2.30)

$$\xi_{R-max} \approx \frac{\sqrt{2}\lambda d}{\sigma L}\tan(\theta_B)$$

As a quick example of this we see that a helium-neon laser beam with a Gaussian width of 0.5 mm has an intrinsic angular width of 0.03°. As mentioned above determining whether this will significantly alter the diffraction characteristics depends on the grating itself (through its angular selectivity curve), thus examples showing the effects of this and other beams will be left until chapters 4 and 5, where specific cases are considered.

2.4   Two Dimensional Coupled Wave Analysis

In the previous section the combination of plane wave decomposition and traditional coupled wave analysis was used as a means to describe the diffraction of a beam whose amplitude may vary in two directions. Another method to do this is to use two-dimensional coupled wave analysis. This method was first used in 1962 by Takagi

[7] in x-ray diffraction in order to account for changes in crystal parameters. It was not used in volume holography until 1977 when Kogelniks equations were rederived in two-dimensional form by Solymar [8]. It has since been extended to give solutions for gratings with non flat boundaries [9-11], non-uniform gratings and replay beams [12-14], and in a limited manner reflection gratings [9,15].

Although this derivation closely follows that of one dimensional coupled wave analysis, which was reviewed at the beginning of this chapter, its ability to include effects resulting from the profiles of the recording beams, along with the number of approximations which are required for a solution warrant a full re-derivation of the coupled wave equations.

The recording of the hologram is made by the interference of a reference and a subject beam, which can be written respectively as:

$$E_R = R_0 \, \hat{A}_{R0}(x,y,z) e^{-j\left(\vec{k}_r \cdot \vec{r}\right)} \quad \text{and} \quad E_S = S_0 \, \hat{A}_{S0}(x,y,z) e^{-j\left(\vec{k}_s \cdot \vec{r}\right)} \tag{2.31}$$

In these equations R_0 and S_0 are amplitude constants, $\hat{A}_{R0}$ and $\hat{A}_{S0}$ describe the variation of these amplitudes with position (i.e. the recording beam profiles), k_r and k_s are the propagation constants, and r is the position vector. Although there is no inherent difficulty in allowing k to be complex, the derivation here will be limited to real values for k (equivalent to ignoring absorption during recording). For the interested reader a derivation of the coupled wave equations when the propagation constants are complex is given in [8]. The holographic recording material is assumed linear with incident field

26

intensity, thus after recording and development, the dielectric constant of the material

will be proportional to the field intensity distribution:

$$\Delta\varepsilon_r = c|E_R + E_S|^2 = ct\left\{\begin{array}{l}R_0^2\,\hat{A}_{RO}^2(x,y,z) + S_0^2\,\hat{A}_{SO}^2(x,y,z) \\ + 2R_0\,S_0\,\hat{A}_{RO}(x,y,z)\hat{A}_{SO}(x,y,z)\cos\left[\left(\vec{k}_S - \vec{k}_R\right)\cdot\vec{r}\right]\end{array}\right\} \tag{2.32}$$

where equation 2.31 was inserted for E_R and E_S. This can be written in a simpler form by

defining a vector K as $\vec{K} = \left(\vec{k}_S - \vec{k}_R\right)$, along with the following two parameters:

$$\varepsilon_{r1} \equiv 2cR_0S_0 \qquad \varepsilon_{r0}(x,y,z) \equiv \frac{\varepsilon_{r1}}{2}\left[\hat{A}_{RO}^2(x,y,z)\frac{R_0}{S_0} + \hat{A}_{SO}^2(x,y,z)\frac{S_0}{R_0}\right] \tag{2.33}$$

With these expressions, eq. 2.32 for the permitivity of the material simplifies to:

$$\Delta\varepsilon_r = \varepsilon_{r0}(x,y,z) + \varepsilon_{r1}\,\hat{A}_{RO}(x,y,z)\hat{A}_{SO}(x,y,z)\cos\left[\vec{K}\cdot\vec{r}\right] = n^2(x,y,z) \tag{2.34}$$

A solution for the diffracted field is obtained by inserting this permitivity

variation into the wave equation: $\nabla^2(E) + k^2(E) = 0$. Since we are dealing with thick

media we can again assume that we are in a two wave region, thus during replay the field

in the grating will be of the form:

$$E = \hat{R}(x,y,z)\hat{A}_R(x,y,z)e^{-j(\vec{\rho}\cdot\vec{r})} + \hat{S}(x,y,z)\hat{A}_S(x,y,z)e^{-j(\vec{\sigma}\cdot\vec{r})} \tag{2.35}$$

27

which we can be written as $E = E_1 + E_2$. For simplicity we will take the Laplacian of each part separately and combine them later. Using straightforward (albeit tedious) algebra we get:

$$\nabla^2 E_1 = \left\{ \begin{array}{l} \nabla^2 \hat{R} \cdot \hat{A}_R + \nabla^2 \hat{A}_R \cdot \hat{R} + 2\left[\nabla \hat{A}_R \cdot \nabla \hat{R}\right] - 2j\rho_x \left[\frac{\partial \hat{R}}{\partial x} \cdot \hat{A}_R + \hat{R}\frac{\partial \hat{A}_R}{\partial x}\right] \\[4mm] - 2j\rho_z \left[\frac{\partial \hat{R}}{\partial z}\hat{A}_R + \hat{R}\frac{\partial \hat{A}_R}{\partial z}\right] - \beta^2 \hat{R}\hat{A}_R \end{array} \right\} e^{-j(\vec{\rho}\cdot\vec{r})} \quad (2.36)$$

In this equation the dependence of R' and A'_R on position is understood and has not been explicitly written. For the remainder of the chapter this will be done wherever it will not lead to confusion. Additionally, use was made in this equation of the fact that $\beta^2 = \rho_x^2 + \rho_z^2$ (propagation was confined to the x-z plane). Following a similar procedure as above, we can obtain the following expression for $\nabla^2 E_2$:

$$\nabla^2 E_2 = \left\{ \begin{array}{l} \nabla^2 \hat{S}\hat{A}_S + \nabla^2 \hat{A}_S \hat{S} + 2\left[\nabla \hat{A}_S \cdot \nabla \hat{S}\right] - 2j\sigma_x \left[\frac{\partial \hat{S}}{\partial x}\hat{A}_S + \hat{S}\frac{\partial \hat{A}_S}{\partial x}\right] \\[4mm] - 2j\sigma_z \left[\frac{\partial \hat{S}}{\partial z}\hat{A}_S + \hat{S}\frac{\partial \hat{A}_S}{\partial z}\right] - |\sigma|^2 \hat{S}\hat{A}_S \end{array} \right\} e^{-j(\vec{\sigma}\cdot\vec{r})} \quad (2.37)$$

Next, an expression is needed for $k^2 E$. This is obtained from eq. 2.34 as:

$$k^2 E = \left[\frac{2\pi}{\lambda} n\right]^2 E = \left\{ \frac{4\pi^2}{\lambda^2} \left[\varepsilon_{r0-orig} + \Delta\varepsilon_{r0}(x,y,z) + \varepsilon_{r1} \hat{A}_{R0} \hat{A}_{S0} \cos\left[\vec{K}\cdot\vec{r}\right]\right]\right\} E \quad (2.38)$$

By neglecting $\Delta\varepsilon_{r0}(x,y,z)$ (the effects of this will be discussed in chapter 9), and using the

fact that $\beta=2\pi n/\lambda$, this can be written (again after significant algebra) as:

$$k^2 E = \left\{ \begin{array}{l} \beta^2 \hat{A}_R \hat{R} e^{-j(\vec{\rho}\cdot\vec{r})} + \beta^2 \hat{A}_S \hat{S} e^{-j(\vec{\sigma}\cdot\vec{r})} + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_R\hat{R}e^{-j(\vec{\sigma}\cdot\vec{r})} \\ + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_S\hat{S}e^{-j(\vec{\rho}\cdot\vec{r})} + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_R\hat{R}e^{-j\left[(\vec{\rho}+\vec{K})\cdot\vec{r}\right]} \\ + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_S\hat{S}e^{-j\left[(\vec{\sigma}-\vec{K})\cdot\vec{r}\right]} \end{array} \right\} \quad (2.39)$$

By combining equations 2.36, 2.37 and 2.39 into the wave equation, the following

monstrosity is generated:

$$\left\{ \begin{array}{l} \nabla^2\hat{R}\,\hat{A}_R + \nabla^2\hat{A}_R\,\hat{R} + 2\left[\nabla\hat{A}_R \cdot \nabla\hat{R}\right] - 2j\rho_x\left[\dfrac{\partial\hat{R}}{\partial x}\hat{A}_R + \hat{R}\dfrac{\partial\hat{A}_R}{\partial x}\right] \\ -2j\rho_z\left[\dfrac{\partial\hat{R}}{\partial z}\hat{A}_R + \hat{R}\dfrac{\partial\hat{A}_R}{\partial z}\right] + (\beta^2-\beta^2)\hat{R}\,\hat{A}_R + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_S\hat{S} \end{array} \right\} e^{-j(\vec{\rho}\cdot\vec{r})}$$

$$+ \left\{ \begin{array}{l} \nabla^2\hat{S}\,\hat{A}_S + \nabla^2\hat{A}_S\,\hat{S} + 2\left[\nabla\hat{A}_S \cdot \nabla\hat{S}\right] - 2j\sigma_x\left[\dfrac{\partial\hat{S}}{\partial x}\hat{A}_S + \hat{S}\dfrac{\partial\hat{A}_S}{\partial x}\right] \\ -2j\sigma_z\left[\dfrac{\partial\hat{S}}{\partial z}\hat{A}_S + \hat{S}\dfrac{\partial\hat{A}_S}{\partial z}\right] + (\beta^2-|\sigma|^2)\hat{S}\,\hat{A}_S + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_R\hat{R} \end{array} \right\} e^{-j(\vec{\sigma}\cdot\vec{r})} \quad (2.40)$$

$$+ 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_R\hat{R}\,e^{-j\left[(\vec{\rho}+\vec{K})\vec{r}\right]} + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_S\hat{S}e^{-j\left[(\vec{\sigma}-\vec{K})\vec{r}\right]} \equiv 0$$

The complexity of this equation (in which the dependence of each term on all

coordinates is not shown) should be obvious. In order to proceed several simplifications

must be made to reduce its complexity. First we ignore ∇^2R' and ∇^2S', which as in the

29

one-dimensional case should be valid if the coupling is slow on the order of several wavelengths. Second, we will not allow beam profiles to spatially vary too quickly, again on the order of several wavelengths – thus allowing us to neglect $\nabla^2 A'_R$ and $\nabla^2 A'_S$. By analogy with one-dimensional analysis these assumptions are valid if the strength of the modulation is very small (i.e. $\varepsilon_r / \varepsilon_{r0} \ll 1$) [16], which should be valid for most of the materials considered in this work. Additionally we will ignore the last two terms in 2.40 since they are well off Bragg, and will thus not contain any significant power. Finally since we assumed that $\hat{A}_R$, $\hat{A}_S$, $\check{R}$ and $\hat{S}$ did not vary too quickly then we can also ignore $2(\nabla \hat{A}_R \cdot \nabla \hat{R})$ and $2(\nabla \hat{A}_S \cdot \nabla \hat{S})$ (see Solymar and Cooke [17]). This leaves:

$$
\begin{aligned}
&\left\{ \begin{bmatrix} -2j\left[\rho_x \dfrac{\partial \hat{R}}{\partial x} + \rho_z \dfrac{\partial \hat{R}}{\partial z} \hat{R} \right] \hat{A}_R - 2j\left[\rho_x \dfrac{\partial \hat{A}_R}{\partial x} + \rho_z \dfrac{\partial \hat{A}_R}{\partial z} \right] \hat{R} \\ + 2\kappa\beta \hat{A}_{R0} \hat{A}_{S0} \hat{A}_S \hat{S} \end{bmatrix} e^{-j(\vec{\rho}\cdot\vec{r})} \right. \\
&\left. + \left\{ \begin{bmatrix} -2j\left[\sigma_x \dfrac{\partial \hat{S}}{\partial x} + \sigma_z \dfrac{\partial \hat{S}}{\partial z} \right] \hat{A}_S - 2j\left[\sigma_x \dfrac{\partial \hat{A}_S}{\partial x} + \sigma_z \dfrac{\partial \hat{A}_S}{\partial z} \right] \hat{S} \\ + (\beta^2 - |\sigma|^2) \hat{S} \hat{A}_S + 2\kappa\beta \hat{A}_{R0} \hat{A}_{S0} \hat{A}_R \hat{R} \end{bmatrix} e^{-j(\vec{\sigma}\cdot\vec{r})} \right\} = 0
\end{aligned}
\tag{2.41}
$$

We now define P_1 and P_2 as: $P_1 \equiv \dfrac{1}{\beta}(\rho_x\, x + \rho_z\, z)$, $P_2 \equiv \dfrac{1}{\beta}(\sigma_x\, x + \sigma_z\, z)$, where P_1 is a unit vector in the direction of travel of the phase front of wave 1. P_2 is in the direction of travel of wave 2, however its magnitude varies slightly from unity as we deviate from Bragg. Using these definitions equation 2.41 becomes:

$$0 = \left\{ -2j\beta(\nabla P_1 \cdot \nabla \hat{R})\hat{A}_R - 2j\beta(\nabla P_1 \cdot \nabla \hat{A}_R)\hat{R} + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_S\hat{S} \right\} e^{-j(\vec{\rho}\cdot\vec{r})}$$

$$+ \left\{ \begin{array}{l} -2j\beta(\nabla P_2 \cdot \nabla S')\hat{A}_S - 2j\beta(\nabla P_2 \cdot \nabla \hat{A}_S)\hat{A}_S\hat{S} \\ + (\beta^2 - |\sigma|^2)\hat{S}\hat{A}_S + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_R\hat{R} \end{array} \right\} e^{-j(\vec{\sigma}\cdot\vec{r})} \qquad (2.42)$$

Per [18] any two-dimensional coupled wave analysis must obey the laws of geometrical optics. One requirement of this is that the power conservation theorem be met, thus: $2\nabla P_1 \cdot \nabla\hat{R} + a\nabla^2 P_1 = 0$. Furthermore, from our definition of P_1 and P_2 we have $\nabla^2 P = 0$.

Combining these with the previous condition gives: $2\nabla P_1 \cdot \nabla\hat{A}_R = 0$ & $2\nabla P_2 \cdot \nabla\hat{A}_S = 0$.

Inserting these requirements in 2.42 gives:

$$\nabla^2 E + k^2 E = \left\{ -2j\beta(\nabla P_1 \cdot \nabla\hat{R})\hat{A}_R + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_S\hat{S} \right\} e^{-j(\vec{\rho}\cdot\vec{r})}$$

$$+ \left\{ -2j\beta(\nabla P_2 \cdot \nabla\hat{S})\hat{A}_S + (\beta^2 - |\sigma|^2)\hat{S}\hat{A}_S + 2\kappa\beta\hat{A}_{R0}\hat{A}_{S0}\hat{A}_R\hat{R} \right\} e^{-j(\vec{\sigma}\cdot\vec{r})} = 0 \qquad (2.43)$$

For a solution at all places, each exponent must separately equate to 0. Doing this, and making use of Kogelnik's dephasing parameter ($\vartheta \equiv \dfrac{(\beta^2 - |\sigma|^2)}{2\beta}$) allows us to write eq. 2.41 as:

$$(\nabla P_1 \cdot \nabla\hat{R}) = -j\kappa\frac{\hat{A}_{R0}\hat{A}_{S0}\hat{A}_S}{\hat{A}_R}\hat{S}$$

$$(\nabla P_2 \cdot \nabla\hat{S})A_S' + j\vartheta\hat{S} = -j\kappa\frac{\hat{A}_{R0}\hat{A}_{S0}\hat{A}_R}{\hat{A}_S}\hat{R} \qquad (2.44)$$

Again the dependence of the variables on position in these equations is understood and has not been shown. Although the equations are now simpler, they are still not in a form where solutions are evident. In order to put them in such a form, a coordinate transformation must be made. The new coordinates used are r and s which are defined to be perpendicular to the directions of propagation of the reference and object beams (perpendicular to P_1 and P_2). The coordinate system is shown below in Figure 2.7.
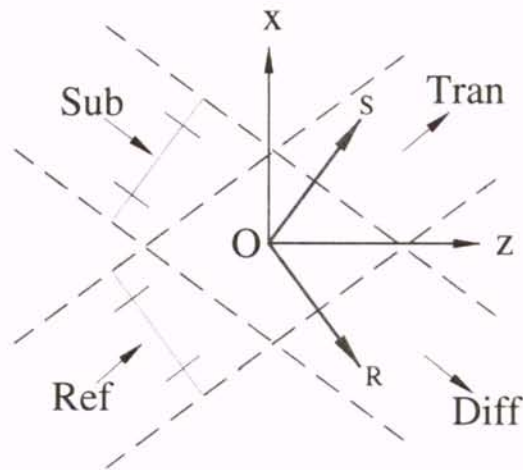
Figure 2.7. Transformed Coordinate System

and the coordinate transformations are given by:

$$r = \sin\left(\theta_t\right)z - \cos\left(\theta_t\right)x$$
$$s = \sin\left(|\theta_d|\right)z + \cos\left(|\theta_d|\right)x$$

(2.45)

As a result of this $\hat{R}(x,y,z)$ and $\hat{S}(x,y,z)$ are transformed into $R(s,r,y)$ and $S(s,r,y)$. additionally the quantities $\hat{A}_R(x,y,z)$ and $\hat{A}_S(x,y,z)$ are transformed into $A_R(r,y)$ and $A_S(s,y)$, while $\hat{A}_{R0}(x,y,z)$ and $\hat{A}_{S0}(x,y,z)$ are transformed into $A_{R0}(s,r,y)$ and $A_{S0}(s,r,y)$,. With this change in coordinates we can write the following:

$$(\nabla P_1 \cdot \nabla R) = \frac{\partial R}{\partial r}(\nabla P_1 \cdot \nabla r) + \frac{\partial R}{\partial s}(\nabla P_1 \cdot \nabla s)$$
$$(\nabla P_2 \cdot \nabla S) = \frac{\partial S}{\partial r}(\nabla P_2 \cdot \nabla r) + \frac{\partial S}{\partial s}(\nabla P_2 \cdot \nabla s)$$

(2.42)

Since the system is defined such that r and s are perpendicular to the directions of propagation, $(\nabla P_1 \cdot \nabla r) = (\nabla P_2 \cdot \nabla s) = 0$. Inserting these into equation 2.46 and plugging the result into 2.44 yields:

$$\frac{\partial R}{\partial u_s}(\nabla P_1 \cdot \nabla s) = -j\kappa \frac{A_{R0} A_{S0} A_S}{A_R} S$$
$$\frac{\partial S}{\partial u_r}(\nabla P_2 \cdot \nabla r)A_S + j\vartheta S = -j\kappa \frac{A_{R0} A_{S0} A_R}{A_S} R$$

(2.47)

Now from the definitions of P_1, P_2, r and s (see eq. 2.45) we can write:

$$\nabla P_1 \cdot \nabla s = \sin(\theta_t + |\theta_d|)$$
$$\nabla P_2 \cdot \nabla r = \frac{|\sigma|}{\beta}\sin(\theta_t + |\theta_d|)$$

(2.48)

33

Which, when inserted in equation 2.47 reduces the coupled wave equations to:

$$\frac{\partial R}{\partial s} = -j\kappa' \frac{A_{R0} A_{S0} A_S}{A_R} S$$

$$\frac{\partial S}{\partial r} A_S + j\vartheta' S \frac{\beta}{|\sigma|} = -j\kappa' \frac{A_{R0} A_{S0} A_R}{A_S} \frac{\beta}{|\sigma|} R \qquad (2.49)$$

where $\vartheta' = \vartheta / \sin(\theta_t + |\theta_d|)$ & $k' = k / \sin(\theta_t + |\theta_d|)$

As a final simplification we can approximate $\dfrac{\beta}{|\sigma|}$ as 1 everywhere except in the

dephasing term itself and our coupled wave equations become:

$$\frac{\partial R}{\partial s} = -j\kappa' \frac{A_{R0} A_{S0} A_S}{A_R} S$$

$$\frac{\partial S}{\partial r} A_S + j\vartheta' S = -j\kappa' \frac{A_{R0} A_{S0} A_R}{A_S} R \qquad (2.50)$$

The solution of which will give us the transmitted and diffracted wave amplitudes. Further development of this solution and describing the results depends on the particular values of the components. This will be given in the chapters where it is to be used.

CHAPTER 3

MATERIALS OF VOLUME HOLOGRAPHY

In order to design volume holographic optical elements, some understanding of the materials available is needed in order properly select one for a given application. Before listing these however, it seems prudent to give characteristics of an ideal material in order to see how existing ones compare. This material would be readily available, linear, capable of high diffraction efficiency with low scattering and have a low cost. Depending upon the required application, it would be either erasable (for read / write applications) or permanently fixable (for read only applications). As will be seen, none of the existing materials meets all of these requirements, and tradeoffs must be considered when selecting one. For more information on any of the materials, see Syms [1] or Solymar and Cooke [2]. They are listed in order of general popularity.

3.1   Photographic Emulsions

Photographic emulsion materials consist of a plate of glass on which a thin layer of gelatin containing silver halide is coated. They are stable, require additional

35

processing to obtain an image, have high sensitivity and can be used for both amplitude and phase gratings. Thicknesses are generally limited to approximately 15 μm, which is low, but considered acceptable for volume holography. Their principle disadvantages are processing difficulties and relatively large scattering. Additionally, they are becoming less available as many manufacturers have stopped making them.

The intrinsic response of these materials lies in the blue / UV end of the spectrum, however by including dyes, they can be sensitized to longer wavelengths. The recording process in these materials includes generating silver halide grains, which can be too large for recording high resolution patterns. To counter this, materials were developed with a smaller grains, but these had a huge drop in sensitivity. Thus a tradeoff exists between efficiency and resolution. Initial development of these materials produces an amplitude hologram, however they can be bleached into a transparent silver salt which results in a higher efficiency phase grating.

3.2   Dichromated Gelatin

After photographic emulsions the next most popular material is dichromated gelatin. This was first used for volume holography in the late 60's [3], and immediately yielded efficiencies above 90%. It consists of a gelatin layer doped with ammonium dichromate. Thicknesses can be as high as 100 μm, and both diffraction efficiency and optical quality are higher than photographic emulsion. Unfortunately its short useful lifetime has made it commercially unavailable, and it must be made in the laboratory by stripping the chemicals off of photographic plates and coating them on a plate of glass.

36

Unlike photographic emulsion where the light sensitive material is grainy, dichromated gelatin is homogeneous, giving low scattering and high spatial modulations. High index modulations (above 0.08) have been achieved by Chang and Leonard [4]. Again, the intrinsic sensitivity is in the UV / blue range, but can be shifted by the addition of dyes. The principle disadvantage of this material seems to be that it can experience shrinkage during development. This material has many excellent properties, however until it becomes commercially available its use will probably remain limited.

3.3   Photopolymers

Photopolymers typically consist of a film-forming polymer, a photoinitiation system, and one or more monomers. During exposure, light is absorbed by the initiator, causing the monomer to become polymerized. These polymers have a slight difference in refractive index, which gives the required index variation. The process occurs in real time and is considered self-developing. The exposure sensitivity is relatively high, and although scattering is lower than photographic emulsion, it is still considered a problem.

Many different photopolymers exist, and the results obtained with each are quite different. This makes it impossible to give minimum and maximum thicknesses, Δn's or even efficiencies. Historically most photopolymers have been experimental, thus it has been difficult to verify results or to select materials for future development. Nonetheless, the excellent properties of these materials has led to some becoming commercially available, and it is likely that more will become so in the future.

37

3.4   Photochromics

A photochromic material is one whose color is altered on exposure to light. Typically, exposure at one wavelength causes a change in one direction (activation), which is reversed by exposure at another wavelength (or by thermal relaxation). The thermally unstable state is usually darker, so the reversal is called bleaching. Reversibility implies reusability, which is advantageous, however the efficiency of these materials is typically low, and they have been eclipsed by other holographic materials.

Typically these are fine-grained materials, which allows high spatial resolutions (albeit with a low sensitivity). Holograms are made either by exposure of a bleached material or by bleaching of an activated material. In either case the grating is made in real time. The modulation is primarily of the absorption type. These materials can be made very thick giving increased selectivity and the ability to superimpose many gratings, however later exposures tend to degrade earlier ones significantly.

3.5   Photorefractives

Another important class of holographic material is photorefractive crystals. They were first used for holography by Chen, LaMacchia, and Fraser in 1968 [5]. These are real time materials, with many properties which make them quite complex. The most widely studied of them is lithium niobate ($LiNbO_3$), which will be emphasized here.

In any photorefractive crystal the hologram is recorded using the photorefractive effect. The first step is the photoexcitation of electrons, which then move through the crystal lattice. There are three contributions to this motion: the influence of electric fields, the photovoltaic effect, and diffusion. Depending on the particular crystal and experimental contributions, any one of these effects can dominate. Electrons are then trapped by nearby unexposed regions, and set up a space-charge field. This modulates the refractive index via the electro-optic effect, resulting in the recording of a volume phase hologram. In order for this to work, suitable traps are necessary, which are typically obtained by doping the $LiNbO_3$ with Fe impurities. The exact performance of any sample depends on the level of doping and on the ratio of Fe^{2+} and Fe^{3+} ions.

These crystals can be quite thick, which gives excellent selectivity and the ability to record many holograms. Using only angular multiplexing Staebler et al [6] recorded over 500 holograms in a 2mm thick sample, each with DE over 2.5% (although with poor uniformity across the recordings). Taking advantage of a scheduling method proposed by Mok et al [7] to improve this uniformity, An, Psaltis, and Burr recorded 10,000 holograms with notable uniformity in a single $LiNbO_3$ crystal [8]. The diffraction efficiency of each hologram was approximately 2×10^{-9}.

In general, additional exposure can re-excite electrons out of the traps and redistribute them, thus causing an erasure of the hologram. Thus some sort of fixing is required for nondestructive readback. In $LiNbO_3$ this is usually done by modest heating ($\approx 100°C$) after recording. This gives rise to a stable pattern which can last for months.

Due to the mechanism of real-time recording and to the crystallinity of the medium, modeling the output of photorefractives is quite difficult. Reasons for this

include the fact that the writing beams are themselves coupled during the recording, and the anisotropy of the crystals requiring a vectorial analysis.

3.6   Photo-Thermo Refractive (PTR) Glass

PTR glass is a new material for recording volume phase holograms. It is an aluminosilicate base glass doped with Ce_2O_3, Ag_2O F, Br. The index variation is obtained in the following manner. Upon exposure to ultraviolet light, an additional electron is removed from a Ce^{3+} ion, which is then absorbed by a Ag^+ ion. The glass is then heated at 450°C where the neutral silver atoms form metal colloid centers, which act as centers of nucleation. Further heat treatment at 520°C forms microcrystals where the colloid centers were. These microcrystals have a different index then the base glass, resulting in a spatial variation of the index of refraction [9].

The glass can be made quite thick, which allows high diffraction efficiencies and narrow angular and wavelength selectivities. This in turn gives the ability to multiplex many holograms on a single spot. The resolution and dynamic range are both high [10-11]. Because it is a glass, it has the additional benefit of being able to be shaped almost arbitrarily, even after exposure. It seems to have two main disadvantages – first it is only available experimentally, and second the recording must take place in the ultraviolet. Because of this requirement, if a pictoral hologram is recorded, and replay is attempted at a different wavelength, then additional distortion must be present. This arises because each spatial frequency making up the subject beam will record a separate hologram with

40

the reference beam. If replay is at the recording wavelength then a single angle (that of the recording subject beam) will Bragg match all of these gratings, however if replay is at some other wavelength, then only one of these gratings can be exactly Bragg matched for any incidence angle.

Although all of the materials listed above are considered volume holographic materials, the fabrication of VHOEs with strong selectivity requires very thick media. Thus all experimental portions of this work will be limited to the use of photorefractive crystals (LiNbO$_3$) and PTR glass. Additionally, this work will include considerable modeling of the diffraction characteristics resulting from a variety of effects. In making these models, it is assumed that the recording material has a linear response to the incident intensity, exhibits no scattering, and is capable of obtaining index modulations on the order of 10^{-4}. When comparing the predictions of these models to experimental data, any effects due to the material violating these assumptions will be noted.

CHAPTER 4

DIFFRACTION OF FINITE BEAMS BY

TRANSMISSION GRATINGS

In chapter 2 solutions were derived for the diffracted field and the diffraction efficiency of plane waves incident on transmission gratings. At this point we will consider the characteristics of these structures in more detail, concentrating on their filtering abilities and beam profiles. Theoretical predictions and experimental verification will be given and discrepancies between them will be discussed.

4.1   Angular Filtering Characteristics

The diffraction efficiency versus dephasing for a transmission grating was given in equation 2.21 (which is repeated below), and the resulting characteristics were summarized in table 1. This table indicated that transmission gratings are best suited for angular filters at reasonable angles of incidence (due to the presence of a tangent function in their wavelength characteristics).

$$S = -j\left(\frac{c_R}{c_S}\right)^{\frac{1}{2}} \exp(-\alpha d/c_R)e^{\xi}\sin(v^2 - \xi^2)^{\frac{1}{2}}\Big/\left(1 - \xi^2/v^2\right)$$

$$v = \pi n_1 d\Big/\lambda(c_R c_S)^{\frac{1}{2}} \qquad \xi = \vartheta d/2c_S \tag{4.1}$$

$$\eta = \sin^2(v^2 + \xi^2)^{\frac{1}{2}}\Big/\left(1 + \xi^2/v^2\right)$$

Because these filtering properties are potentially useful, an investigation into them for gratings recorded in photo-thermo refractive (PTR) glass was made. This involved writing several gratings in 2 mm thick samples of PTR glass, and recording their angular selectivities. Using this data, plots of the diffracted power versus angular deviation from Bragg were generated, a typical one of which is shown in Figure 4.1 below. Along with the experimental results, the graph includes theoretical predictions obtained from equation 4.1, which have been scaled to match the measured diffraction efficiency.
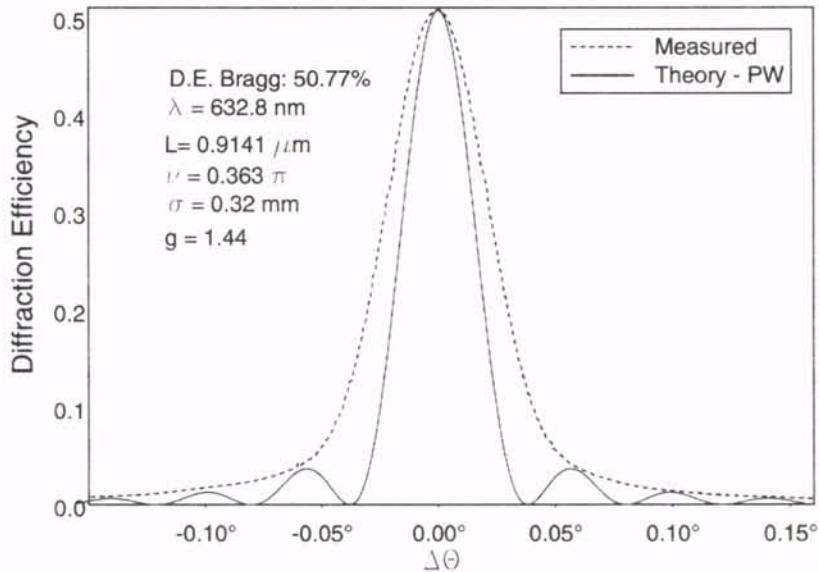
Figure 4.1. Diffraction Efficiency vs. Angular Dephasing – Theory and Experiment

When comparing the two plots, many similarities are apparent, however several obvious differences also exist. The first and most notable of these is the lack of sidelobes in the experimental data. The second (and perhaps more important in filtering applications) is the difference in width of each curve's main lobe. To get a quantitative idea of this difference, we again consider table 2.1, where the full width half power was calculated to be approximately 0.866 L/d, or about 0.023° in the material (0.0343° in air). In looking at the graph we see that the experimental value is roughly 0.052° in air.

The magnitude of these differences between the theoretical and experimental curves warranted additional investigation, and after some consideration the source of the discrepancy became apparent – the characteristics of the replay beam. Recall that equation 2.21 was derived for a uniform plane wave incident on an infinite grating, while the experimental curve was obtained using a typical helium neon laser beam (which is obviously finite, with an approximately Gaussian amplitude profile). By modeling the response of the grating to a similar input beam we should expect to improve on the agreement between the theoretical and experimental data.

As discussed in chapters one and two, the characteristics of the replay beam can be included by using either two-dimensional coupled wave analysis (allowing the amplitude of the wave to vary in two dimensions) or plane wave decomposition. Although either method is valid, the remainder of this section will use plane wave decomposition, with the two-dimensional analysis left until section 4.2 and chapter 6 (where overlap gratings are considered). For a Gaussian input beam the amplitudes of the plane wave components were shown in equation 2.28 to be:

$$\tilde{E}(k_x) = \frac{2E_0\sqrt{\pi}}{\sigma_k} e^{-\left(\frac{k_x - k\sin\theta_0}{\sigma_k}\right)^2}$$

(2.28 / 4.2)

$$\text{where } \sigma_k = \frac{2\cos\theta_0}{\sigma}$$

The values of σ and λ for the replay beam used in Figure 4.1 were inserted into equation 2.28 and coupled wave analysis was applied to the resulting plane waves. Finally, the diffracted amplitudes were summed to form an angular selectivity curve for the specific input beam used. This procedure was performed numerically, and the results are shown in Figure 4.2 below, along with that of the experimental data from Figure 4.1 and the plane wave case for comparison.
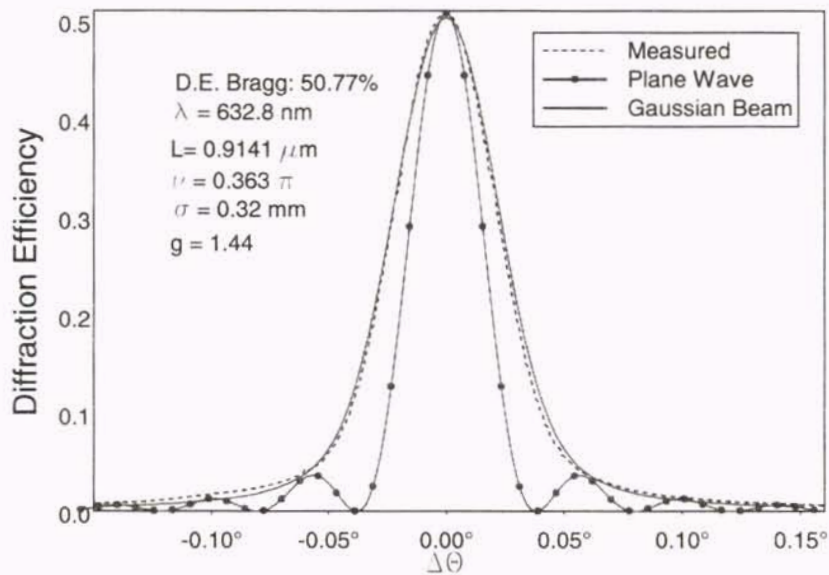
Figure 4.2. D. E. vs. Angular Dephasing – Plane Wave, Gaussian Beam Thy. & Exp.

In this plot we see considerably better agreement between the experimental data and theoretical predictions. First, the widths of the main lobes of both plots are almost exactly the same. Second, we see that the predicted curve no longer contains side lobes, which is in agreement which was experimentally obtained.

There is a simple physical explanation for why the Gaussian beam exhibits wider angular dephasing characteristics than those of a plane wave, and why it should not go to zero between the sidelobes. First, we recognize that the plane wave case is equivalent to keeping only the central angular spectrum component of the input beam. When all of the angular spectrum components are included, the diffracted power at each angle becomes the sum of all components. This power will therefore be higher than that of a single plane wave, and a wider main lobe will result. Additionally, although the power diffracted by the central angular component still drops to zero at points between the side lobes, the other spectral components continue to diffract power at these points, forcing the total diffracted power to be higher than zero (essentially "washing out" the sidelobes).

It is important to note that the changes in the angular selectivity just described are a result of the particular beam used in the experiment, and that the use of a different beam can cause substantially different results. To show this effect (and to verify that the match between theoretical and experimental curves was no mere coincidence) we see Figures 4.3 and 4.4 below, which show similar measurements for increasingly wider input beams.
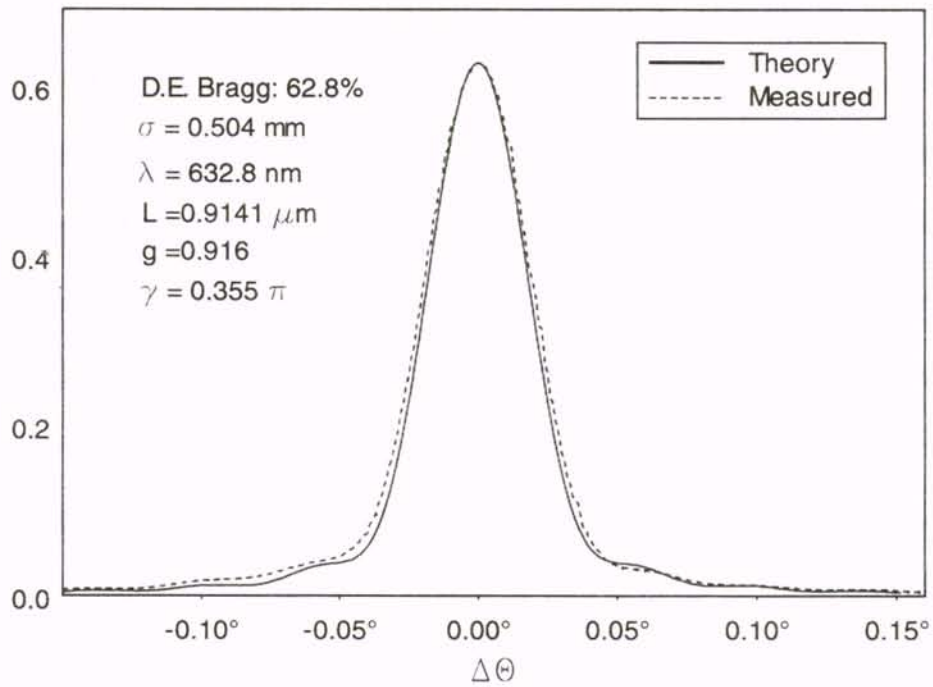
Figure 4.3. D.E. vs. Angular Dephasing – Gaussian Beam ($\sigma = 0.5$ mm) Thy. & Exp.
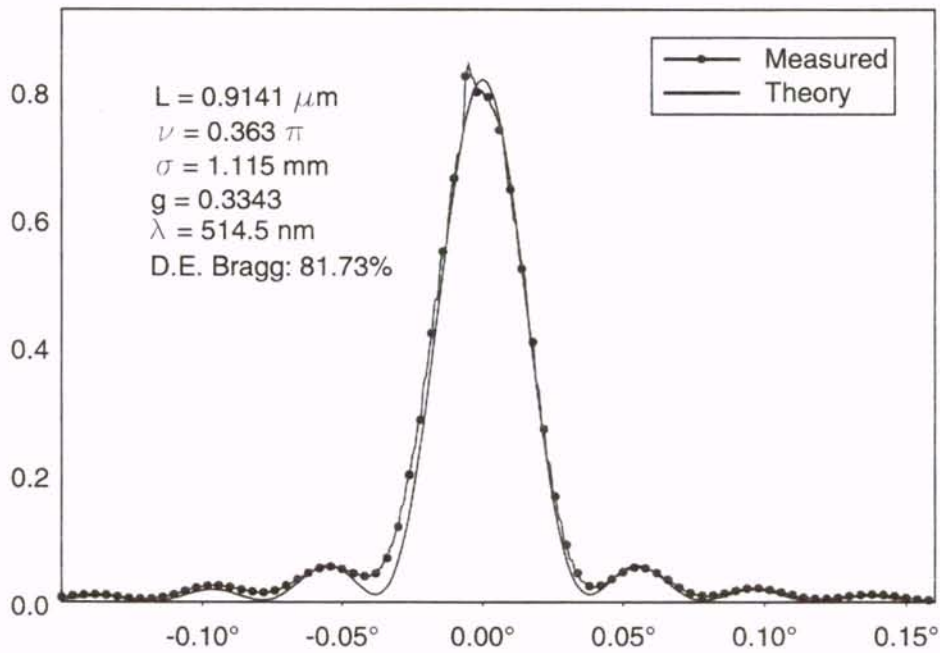
Figure 4.4. D.E. vs. Angular Dephasing – Gaussian Beam ($\sigma = 1.1$ mm) Thy. & Exp.

47

In each of these figures we see excellent agreement between the theoretical and experimental data. Additionally, we see that as the incident beam width increases, the characteristics start to resemble those of a plane wave (i.e. in Figure 4.3 we just begin to see the predicted sidelobes, while in Figure 4.4 the replay beam has become wide enough that they are clearly visible). This should have been anticipated, as Fourier theory tells us that this wider beam will have a narrower angular spectrum (which is closer to the plane wave case).

The discussion accompanying Figures 4.1 through 4.4 provides a nice qualitative analysis of the source of their particular shapes, as well as of the trends exhibited as the incident beam is widened. It is possible however to use the expressions given in chapter 2 for the intrinsic angular dephasing to get a quantitative feel for this behavior.

For the beam used in Figure 4.2 (λ=632.8 nm and σ=0.32 mm), θ_{max} is found from equation 2.30 to be 0.0009 radians or 0.05°. In Figures 4.3 and 4.4 these values decreased to 0.0006 radians (0.03°) and 0.0002 radians (0.012°) respectively. The grating used in all of these experiments however, had an angular FWHM of 0.0006 radians (0.04°) in air. One can get a feel for the amount non-uniform attenuation present by taking the ratio of these quantities. This ratio is 0.6 in Figure 4.2, 1.0 in Figure 4.3 and 3.0 in Figure 4.4.

These figures show that the profile of the replay beam must be considered when accurate predictions of a grating's replay characteristics are needed. It would be nice however, to know when this can be ignored in rough calculations. From Figure 4.3 (in which the effect of the beam profile is large) and Figure 4.4 (which shows "almost plane-wave" behavior), we can surmise that we can ignore the beam profile when the ratio is

48

approximately 3. This therefore becomes our criteria. By combining the equations for these values (table 2.1 and equation 2.30) we can derive the following criteria for ignoring the profile of the replay beam in angular selectivity calculations:

$$\sigma \geq \frac{1.6\lambda}{n} \frac{d}{L} \tag{4.3}$$

This value is useful for making quick assessments of the implications of the incident beam, however it is important to see what the actual attenuation of each component will be. This is shown for our three cases in Figure 4.5 below. Again, the behavior will look like that of a plane wave when the attenuation of each angular spectrum component is nearly the same.
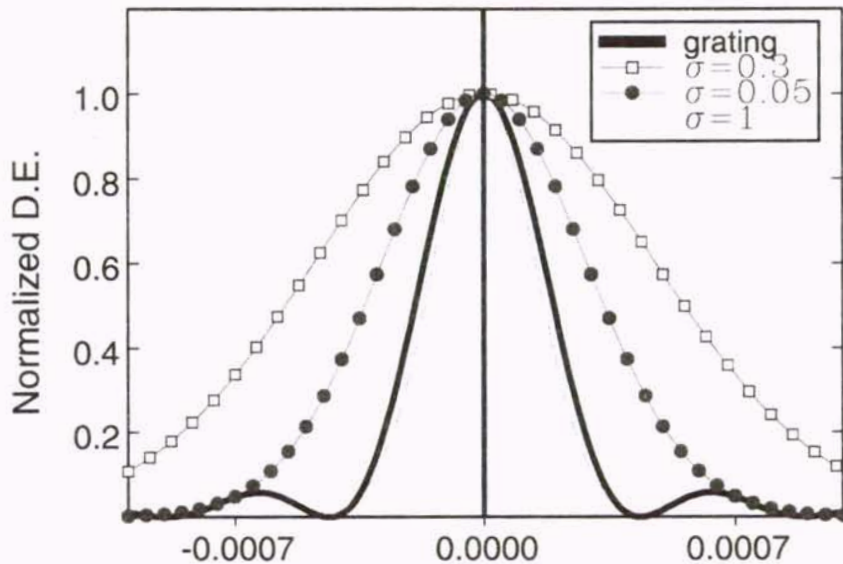
Figure 4.5. Grating angular selectivity and angular spectra with σ=0.3, 0.5 and 1 mm

49

4.2   Beam Profiles

In the last section, the profile of the replay beam was combined with coupled wave analysis in order to improve on the theoretical predictions of the angular selectivity. This was done by performing a plane wave decomposition, essentially forming a two-dimensional theory.   In addition to improved angular selectivity predictions, two-dimensional theory gives an additional benefit – the ability to readily calculate the profiles of the transmitted and diffracted beams, and to compare them to the input beam in order to determine the fidelity of reconstruction.

Both plane wave decomposition and two dimensional analysis have been used to determine these profiles for various input beams (including beams with uniform and gaussian amplitude distributions) [1-7], although all but one of the papers restricted consideration to exact Bragg incidence (with the single exception discussing general off-Bragg trends only). The following section will discuss the profiles of the diffracted and transmitted beams for an input beam with a Gaussian amplitude distribution (the approximate situation encountered in the laboratory). It will concentrate on off-Bragg incidence, although the Bragg case will be included for comparison purposes.

Before performing this analysis, it is noted that Moharam, Gaylord and Magnusson [1] used a two dimensional coupled wave theory to obtain the following expressions for the diffracted and transmitted portions of a Gaussian beam incident on a volume transmission grating at the Bragg angle:

$$S(\bar{s}) = -j\frac{1}{2}\gamma E_0 \int_{-1}^{1} \exp\{-[g(1-u)-\bar{s}]\} \cdot J_0\left[\gamma(1-u^2)^{\frac{1}{2}}\right] du$$

$$R(\bar{r}) = R_0(\bar{r}) - \frac{1}{2}\gamma E_0 \int_{-1}^{1} \exp\{-[g(1-u)-\bar{r}]\} \cdot \left(\frac{1+u}{1-u}\right)^{\frac{1}{2}} J_1\left[\gamma(1-u^2)^{\frac{1}{2}}\right] du \qquad (4.4)$$

where: $g = d\sin\theta/\sigma$ and $\gamma = \kappa d/\cos\theta_0$ $\bar{r} = r/\sigma$ $\bar{s} = s/\sigma$

In these expressions R represents the transmitted beam and S the diffracted one. The coordinate system used is shown below in Figure 4.6.
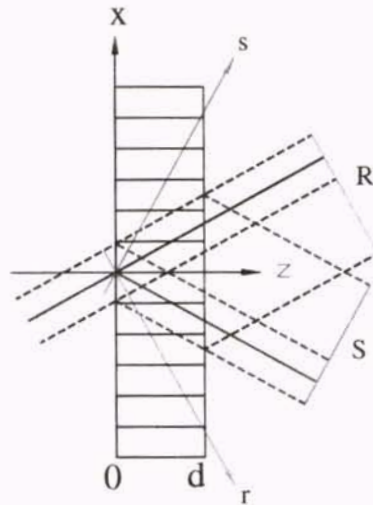
Figure 4.6. Geometry used for 2-dimensional solution for beam profiles

By integrating these solutions over the beam profile we can determine the on-Bragg diffraction efficiency for any value of g and grating strength. A plot of this information for several values of g and for grating strengths up to 4π was presented in [1]. Due to its interesting characteristics this plot has been repeated below.

51
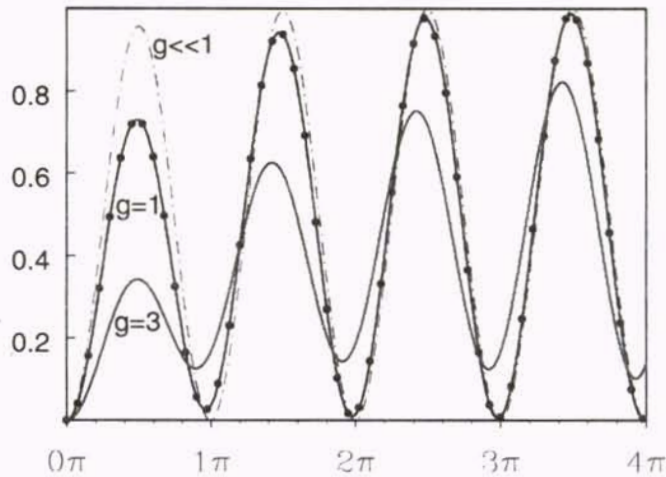
Figure 4.7. On-Bragg diffraction efficiency versus grating strength – various values of g

From this figure we see that the diffraction efficiency of finite beams on transmission gratings will not in general approach unity unless the grating has been well overmodulated. Additionally, this effect is larger as the value of g goes up (as the width of the replay beam is decreased.

The solution in equation 4.4 is very elegant since it is a function of only two normalized parameters: gamma – a normalized grating strength, and g – a geometry factor. These parameters allow a quick and encompassing investigation of many different input beam widths and Bragg angles to be made, which appears to give two dimensional coupled wave analysis an advantage over the angular spectrum of plane waves approach (where normalized parameters are not apparent). It therefore seems prudent to begin our off-Bragg analysis by deriving two-dimensional solutions for off-Bragg incidence in the manner of [1], with the hopes that the solution will closely resemble the Bragg case.

It turns out however that Benlarbi, et al [8] have shown that equation 4.4 can be derived directly from the general plane wave decomposition equations (eq. 2.24), by including the approximations inherent to two-dimensional coupled wave analysis, along with some semi-obscure mathematical identities [9]. For reasons that will become apparent in the next chapter, this method will be used to derive off-Bragg equations similar in form to equation 4.4. That is, starting with the off-Bragg plane wave decomposition equations and using similar identities and approximations to Benlarbi, a closed form expression for the off-Bragg solutions will be derived. In order to avoid duplication of effort this derivation will closely follow [8], differing only when we reach the off-Bragg portion.

The electric field of an incident two-dimensional wave can be written as:

$$E(r) = \int_{-\infty}^{\infty} \Theta(\beta \sin \theta) \exp[-j\beta(z \cos \theta + x \sin \theta)] d(\beta \sin \theta) \qquad (4.5)$$

Now, if we consider the input to be a "beam", with a central spectrum angle θ_0, then this expression can be normalized to that central angle, giving:

$$E(r) = e^{-j\bar{\rho}(\theta_0) \cdot \bar{r}} \int_{-\infty}^{\infty} \Theta(\tau) \exp\{-j[\tau x + \beta z (\cos \theta - \cos \theta_0)]\} d\tau \qquad (4.6)$$

where $\rho(\theta_0) \equiv \beta[\sin \theta_0 \, \hat{x} - \cos \theta_0 \, \hat{z}]$ is the wave vector of the central spectrum component, and Θ was re-defined in terms of the angular frequency parameter τ:

$$\tau = \beta(\sin\theta - \sin\theta_0) \qquad (4.7)$$

Since we have considered this to be a beam traveling in a given direction θ_0, then if the beam does not diverge too quickly, and if the z dependent term in the integrand can be approximated as $(-\tau_z \tan\theta_0)$ then the field can be written as a function of the direction perpendicular to its propagation, ie:

$$E(r) = a(\xi)\exp[-j\vec{\rho}(\theta_0)\cdot\vec{r}] \qquad (4.8)$$

Note that the accuracy of this approximation depends on how far the beam propagates from the boundary and on how quickly the distribution varies with ξ. Equating equations 4.6 and 4.8 at the boundary z=0 yields the following fourier transform pairs:

$$E(0,-x\cos\theta) = \int_{-\infty}^{\infty} \Theta(\tau)\exp(-j\tau x)\,d\tau$$
$$\Theta(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} a(-x\cos\theta_0)\exp(j\tau x)\,dx \qquad (4.9)$$

Note that the only difference in what we have done here and in the fourier decomposition of chapter 2 is that we have decomposed the beam in a direction perpendicular to its direction of propagation (usually we write the field in the x direction and decompose it directly). Standard grating theory tells us that the effect of a grating (with grating vector K_g) on each of these plane wave components will be to produce a

54

transmitted wave and n diffracted waves. This can be expressed (as a differential effect for an individual plane wave) as:

$$dE(\theta,r) = \Theta(\theta)\,d\theta \sum_n v_n(\theta,r)\exp\{-j[\vec{\rho}(\theta)-n\,\vec{K}_g]\cdot\vec{r}\} \qquad (4.10)$$

We can determine the effect of the entire input beam by integrating these individual results over the entire input spectrum. Before doing this, we note that the additional phase matching required by the thickness of the material considered in this work will cause each incident wave to yield only a transmitted beam and a single diffracted order. Thus equation 4.10 can be simplified to:

$$dE(\theta,r) = \Theta(\theta)\,d\theta\left[v_0(\theta,r)e^{-j[\vec{\rho}(\theta)\vec{r}]} + v_1(\theta,r)e^{-j\left[\left(\vec{\rho}(\theta)-\vec{K}_g\right)\cdot\vec{r}\right]} \right] \qquad (4.11)$$

The total filed can now be written as a sum of transmitted and diffracted beams:

$$E(r) = V_0(r)e^{-j\vec{\rho}(\theta_0)\vec{r}} + V_1(r)e^{-j\left(\vec{\rho}(\theta_0)-\vec{K}_g\right)\cdot r} \qquad (4.12)$$

where $V_0(r)$ and $V_1(r)$ are the amplitude distributions of the transmitted and diffracted beams, and are made up respectively of the transmitted and diffracted portions of all angular spectrum components. Thus:

$$V_{-1} = \int_{\sigma} \Theta(\tau) \exp\{-j[\tau x + \beta z(\cos\theta - \cos\theta_0)]\} S(\theta, z) \, d\tau$$

$$V_0 = \int_{\sigma} \Theta(\tau) \exp\{-j[\tau x + \beta z(\cos\theta - \cos\theta_0)]\} R(\theta, z) \, d\tau \tag{4.13}$$

In these expressions S and R are the diffracted and transmitted field amplitudes for each plane wave. Their values (which are obtained from Kogelnik's analysis) are given below, where S is repeated from equation 2.21 and R was obtained from [10]

$$S(\theta, d) = -j e^{-j\xi} v \sin\left(v^2 + \xi^2\right)^{\frac{1}{2}} \Big/ \left(v^2 + \xi^2\right)^{\frac{1}{2}}$$

$$R(\theta, d) = e^{-j\xi} \left[\cos\left(\sqrt{v^2 + \xi^2}\right) + j\xi \frac{\sin\left(\sqrt{v^2 + \xi^2}\right)}{\sqrt{v^2 + \xi^2}} \right] \tag{4.14}$$

In these equations ξ, ϑ and v are as defined in equation 2.21. Inserting equation 4.14 into equation 4.13 yields:

$$V_{-1} = \int_{\sigma} \Theta(\tau) \left[-j e^{-j\xi} v \frac{\sin\left(v^2 + \xi^2\right)^{\frac{1}{2}}}{\left(v^2 + \xi^2\right)^{\frac{1}{2}}} \right] e^{-\left[\tau x + \beta z\left(\cos\theta - \cos\theta_0\right)\right]} d\tau$$

$$V_0 = \int_{\sigma} \Theta(\tau) e^{-j\xi} \left[\cos\left(\sqrt{v^2 + \xi^2}\right) + j\xi \frac{\sin\sqrt{v^2 + \xi^2}}{\sqrt{v^2 + \xi^2}} \right] e^{-\left[\tau x + \beta z\left(\cos\theta - \cos\theta_0\right)\right]} d\tau \tag{4.15}$$

Before continuing with the analysis a more convenient form for the dephasing parameter (in terms of the angular frequency parameter τ) will be derived. In doing this,

56

Benlarbi et. al. restricted their derivation to the case where the central beam component was on Bragg. This allowed Kogelnik's dephasing parameter (ϑ) to be written as $\vartheta = 2\sigma \sin\theta_0$. This restriction is however unnecessary, and it is relatively straightforward to generalize this to off-Bragg incidence, in which case the previously defined angular frequency parameter becomes:

$$\tau \equiv \beta(\sin\theta - \sin\theta_0) = \beta[\sin\theta - \sin(\theta_B + \Delta\theta)] \tag{4.16}$$

The coupled wave dephasing parameter can now easily be written in terms of τ. Doing so yields:

$$\vartheta = 2\tau \sin\theta_B + K_g[\sin\theta_0 - \sin\theta_B] \tag{4.17}$$

which can be written concisely by defining $\tau_0 \equiv K_g[\sin\theta_0 - \sin\theta_B]$ as:

$$\frac{\vartheta}{2} = \sigma\sin\theta_B + \tau_0 \tag{4.18}$$

By substituting the values of ξ and v from equation 2.21, along with ϑ from eq. 4.18, the following two expressions can be written:

$$V_{-1} = \int_\sigma \Theta(\tau) \left[\begin{array}{c} -\dfrac{j\kappa \sin\left[\dfrac{x}{\cos\theta} \sqrt{\kappa^2 + (\tau\sin\theta_B + \tau_0)^2} \right]}{\sqrt{\kappa^2 + (\tau\sin\theta_B + \tau_0)^2}} \\[3em] \times e^{-j\left[\tau\left(x + d\left(\dfrac{\sin\theta_B + \tau_0/\tau}{\cos\theta} \right) \right) + \beta(\cos\theta - \cos\theta_0)z \right]} \end{array} \right] d\tau \qquad (4.19.a)$$

$$V_0 = \int_\sigma \Theta(\tau) \left[\begin{array}{c} \cos\left(\dfrac{d}{\cos\theta} \sqrt{\kappa^2 + (\tau\sin\theta_B + \tau_0)^2} \right) \\[1.5em] + j\dfrac{(\tau\sin\theta_B + \tau_0)\sin\left[\dfrac{d}{\cos\theta} \sqrt{\kappa^2 + (\tau\sin\theta_B + \tau_0)^2} \right]}{\sqrt{\kappa^2 + (\tau\sin\theta_B + \tau_0)^2}} \\[3em] \times e^{-j\left[\tau\left(x + d\left(\dfrac{\sin\theta_B + \tau_0/\tau}{\cos\theta} \right) \right) + \beta(\cos\theta - \cos\theta_0)z \right]} \end{array} \right] d\sigma \qquad (4.19.b)$$

It is important to stress that these equations have been derived using only plane wave decomposition. Since no additional approximations have been made they are as accurate as Kogelnik's equations. Unfortunately their form is somewhat complicated and unwieldy. In order to obtain expressions similar in complexity to those of Moharam et. al [1], further simplification is necessary.

To obtain the desired form two assumptions must be made. These are:

$$\cos\theta \cong \cos\theta_0$$
$$\text{and} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.20)$$
$$\tau x + \tau z \sin\theta_0 / \cos\theta + \beta z (\cos\theta - \cos\theta_0) \cong \tau x$$

58

The first of these assumptions is obvious, and the second is valid to the first order. These requirements will hold as long as the incident beam is not too narrow. Benlarbi has shown the effects of gradual violations of these conditions [3], and it turns out that beams from standard laboratory helium-neon lasers fall well within the region of validity.

Using these approximations, and defining $\bar{\tau}_0$ and P as $\bar{\tau}_0 \equiv \tau_0 / \sin \theta_B$ and $P \equiv d \sin \theta_B / \cos \theta_0$, yields the following pair of equations:

$$V_{-1} = e^{-j\frac{\tau_0 z}{\cos \theta_0}} \frac{j\kappa}{\sin \theta_B} \int_\sigma \Theta(\tau) \left[-\frac{\sin\left[P \sqrt{\left(\frac{\kappa}{\sin \theta_B} \right)^2 + (\tau + \bar{\tau}_0)^2} \right]^2}{\sqrt{\left(\frac{\kappa}{\sin \theta_B} \right)^2 + (\tau + \bar{\tau}_0)^2}} \right] e^{-j\tau x} \, d\tau \quad (4.21)$$

$$V_0 = e^{-j\frac{\tau_0 z}{\cos \theta_0}} \int_\sigma \Theta(\tau) \left[\cos\left(\frac{d}{\cos \theta} \sqrt{\kappa^2 + (\tau \sin \theta_B + \tau_0)^2} \right) + j (\tau \sin \theta_B + \tau_0) \right.$$
$$\left. \times \frac{\sin\left[\frac{d \sin \theta_B}{\cos \theta_0} \sqrt{\left(\frac{\kappa}{\sin \theta_B} \right)^2 + (\tau + \bar{\tau}_0)^2} \right]^2}{\sin \theta_B \sqrt{\left(\frac{\kappa}{\sin \theta_B} \right)^2 + (\tau + \bar{\tau}_0)^2 {}^2}} \right] d\tau \quad (4.22)$$

This can be further simplified by defining the following function:

$$G(\tau) \equiv \sin\left(P\sqrt{(\kappa/\sin \theta_B)^2 + \tau^2} \right) \bigg/ \sqrt{(\kappa/\sin \theta_B)^2 + \tau^2} \qquad (4.23)$$

59

Inserting this into equation 4.20 and 4.21 yields:

$$V_{-1} = -\frac{j\kappa\exp\left[-j\dfrac{\tau_0\,z}{\cos\theta_0}\right]}{\sin\theta_B}\int_\sigma \Theta(\tau)\,G(\tau+\bar{\tau}_0)e^{-j\tau x}\,d\tau$$

(4.24)

$$V_0 = \exp\left[-j\dfrac{\tau_0\,z}{\cos\theta_0}\right]\int_\sigma \Theta(\tau)\left[\dfrac{\partial G(\tau+\bar{\tau}_0)}{\partial P}+j(\tau+\bar{\tau}_0)G(\tau+\bar{\tau}_0)\right]e^{-j\tau x}\,d\sigma$$

Per the Fourier transform pair of equation 4.8 these are merely inverse transforms of products of our input spectrum and shifted versions of the G function defined above, i.e.:

$$V_{-1} = -\frac{j\kappa}{\sin\theta_B}e^{-j\frac{\tau_0\,z}{\cos\theta_0}}\,\Im^{-1}\left[\Theta(\tau)\,G(\tau+\bar{\tau}_0)\right]$$

(4.25)

$$V_0 = e^{-j\frac{\tau_0\,z}{\cos\theta_0}}\left\{\Im^{-1}\left[\Theta(\tau)\dfrac{\partial G(\tau+\bar{\tau}_0)}{\partial P}\right]+j\Im^{-1}\left[\Theta(\tau)(\tau+\bar{\tau}_0)G(\tau+\bar{\tau}_0)\right]\right\}$$

Equation 4.25 may be recast into a better form using the convolution theorem of Fourier transforms, along with help from the shifting theorem. The derivation of the transform pairs for Bragg incidence is outlined in [8], and developed more rigorously in [11]. The extension of these to off Bragg incidence (which is handled by the shifting theorem) is straightforward, and is given in appendix B. Applying the resulting transforms (equations B.9 through B.11), and inserting the expression for a Gaussian beam (eq. 4.1) as our input yields the following amplitude expressions:

$$V_{-1} = \frac{-j\gamma E_0 \kappa}{2} e^{-j\frac{\tau_0 z}{\cos\theta_0}} \int_{-1}^{1} J_0\left(\gamma\sqrt{1-u^2}\right) e^{-\left(g(1-u)-\frac{s}{\sigma}\right)^2} e^{-j\frac{\overline{\tau}_0 u\, d \sin\theta_B}{\cos\theta_0}}\, du \quad (4.26)$$

$$V_0 = e^{-j\frac{\tau_0 z}{\cos\theta_0}} \left[-\frac{1}{2}\gamma E_0 \int_0^1 \exp\left[-\left(g(1-u)-\frac{r}{\sigma}\right)^2\right]\left(\frac{1+u}{1-u}\right)^{\frac{1}{2}} J_1\left(\gamma\sqrt{1-u^2}\right) \times \exp\left(-j\frac{\overline{\tau}_0 u\, d \sin\theta_B}{\cos\theta_0}\right) du + \frac{1}{2\pi} E_0 \exp\left(-\left(\frac{r}{\sigma}\right)^2\right) \right] \quad (4.27)$$

where the normalized parameters γ and g are exactly as defined previously (eq. 4.4). This is the form that we have been searching for – a two dimensional solution for off-Bragg incidence in terms of a few normalized parameters. As we can see from these equations, when τ_0 goes to zero (when the input beam central component is on Bragg) the equations of [1] fall directly out, which provides a nice check.

In addition to showing the equality of plane wave decomposition and 2-D coupled wave analysis when the approximations are valid, this derivation allows us to use the normalized Bragg parameters for the off-Bragg case. This in turn will allow an encompassing number of cases to be quickly covered when investigating off-Bragg profiles. Finally, this long derivation shows us that even when using plane wave decomposition (whether for its accuracy or for another reason), the solution may still be in terms of only a few parameters, even though this may not be apparent.

Now that we have the desired output expressions, we can use them to show the diffracted and transmitted profiles away from Bragg. The first case covered (g=1) is shown in Figure 4.8. Plots are given for grating strengths from $\pi/8$ to $\pi/2$, along with dephasing parameters from 0 to 3. To get a feel for these paramerters we note that in a 2

mm thick grating grating with a period of one micron and an index of 1.5, g=1 corresponds to a helium neon laser with a Gaussian beam width of 0.42 mm (typical of a laboratory he-ne laser). Additionally the dephasing parameters of one, two and three correspond to angular deviations from Bragg of 0.009°, 0.018° and 0.027° respectively.
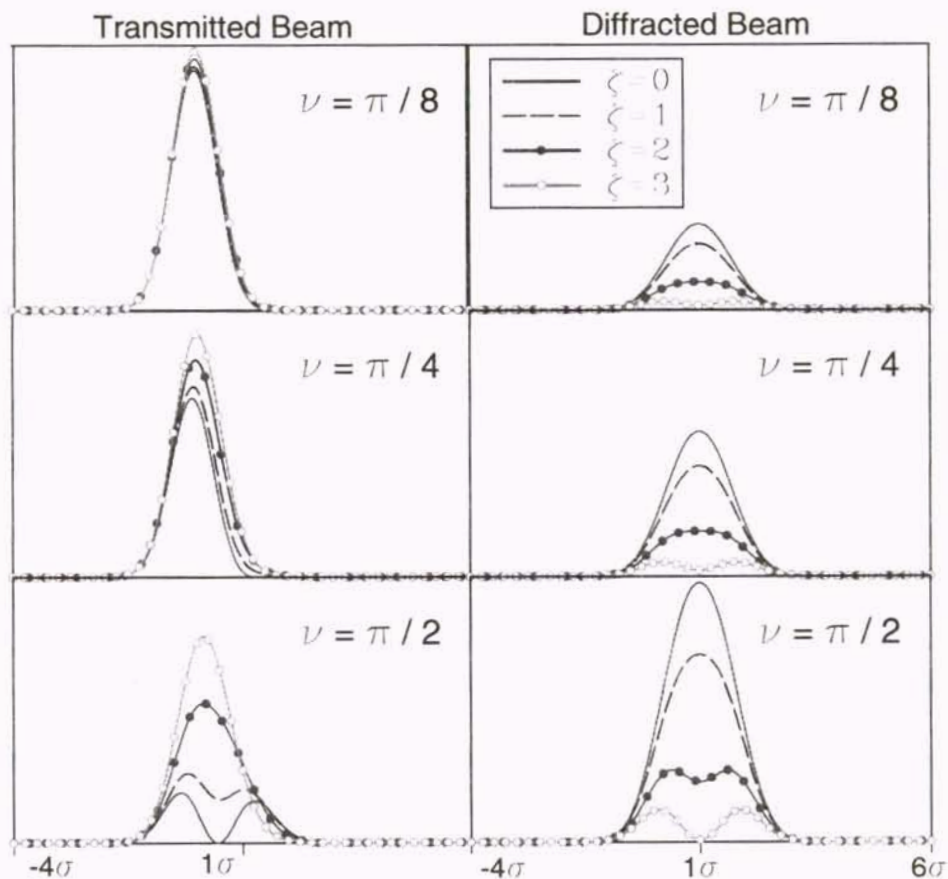
Figure 4.8. Diffracted and transmitted beam profiles, various strengths – g = 1.0

The discussion of the plots will start with the diffracted beam. We see in Figure 4.8 that when g=1, the diffracted profile on-Bragg remains reasonably Gaussian regardless of the modulation strength (the grating faithfully reproduced these replay

62

beams on-Bragg). As we deviate from Bragg however, not only is the diffracted power decreased (which is anticipated), but the diffracted beam also undergoes significant distortion. This distortion increases as we go further from Bragg, even to the point where a beam which looks quite good on-Bragg may break into more than one part. Furthermore we see an enhancement of this effect as the gratings gets stronger.

This off-Bragg behavior can be easily explained. Since distortion occurs when the angular spectrum components are non-uniformly attenuated, and since the uniformitiy of the grating's angular selectivity curve decreases as we deviate from Bragg (the slope of Figure 2.4 increases) then good profiles on-Bragg (where the uniformity is reasonably constant over the widest range) may become distorted as we go off-Bragg (where this uniformity decreases). The enhancement of these effects with stronger gratings occurs because this slope increases with increasing grating strength.

When considering the transmitted beam we see that the best beam profiles occur well away from Bragg, which should be obvious since little power is coupled from the incident beam under these conditions. Few overall generalizations regarding the transmitted beam can be made, except to note that the coupling of energy from the beam does not occur uniformly.

These plots seem to imply that on-Bragg incidence ensures that the diffracted beam will faithfully reproduce the recording beam. This was so in the last example due to the relatively narrow intrinsic angular width of the replay beam, however it is not necessarily true in the general case. To show this, Figure 4.9 below has been included. This is similar to Figure 4.8, except that the geometry factor (g) has been increased to 3. For identical gratings, an increase in g represents a decrease in the beam width, thus for

the grating described above (2 mm thick, n = 1.5, L = 1μm) this value of g represents a beam of width 0.14 mm.
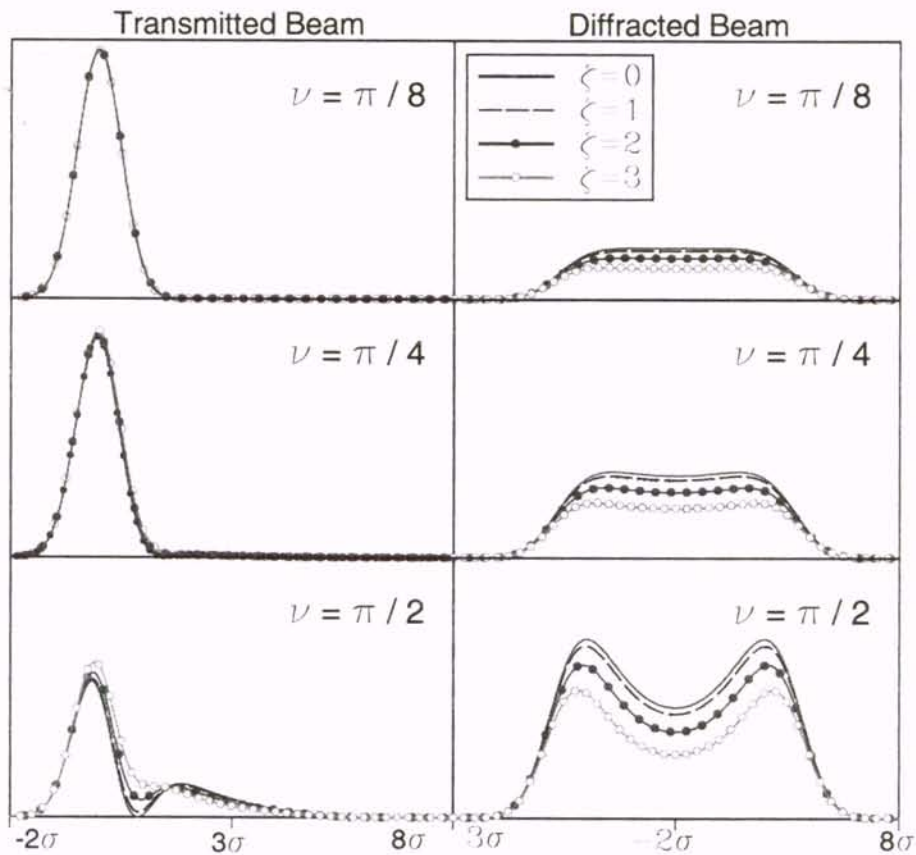
Figure 4.9.   Diffracted and transmitted beam profiles, various strengths – g = 3.0

Again concentrating on the diffracted beam, we see that Figure 4.9 tells a much different story than 4.8.   In this figure all of the diffracted beams are highly distorted. Furthermore, for each grating strength as we go off-Bragg we do see a decrease in power, but the additional distortion that was present throughout Figure 4.8 no longer exists.

Just as in Figure 4.8 these effects can be explained by considering the widths of the grating's angular selectivity curve and the input beam's angular spectrum. Since the angular spectrum of this input beam has increased then the attenuation uniformity across the spectrum has decreased, causing distortion (even on Bragg). Again these effects increase with increasing granting strength since the grating is acting as a low pass filter for angular spectrum components whose passband is narrowed as the strength is increased.

To summarize these plots we see that in Figure 4.8 the narrowness of the angular spectrum resulted in a somewhat uniform attenuation when the beam was centered on Bragg. As we deviated from this incidence however, the attenuation became much less uniform and the beam became significantly more distorted. In contrast to this is Figure 4.9, where the angular spectrum of the input beam was much wider. In this case the non-uniform attenuation was significant, even when the "beam" was on Bragg, thus all of the diffracted beams were highly distorted. Here the severity of the on-Bragg distortion made that which occurs off-Bragg less pronounced. For reasons described above all of these effects became worse as the grating got stronger.

Throughout this section a normalized geometry factor (g) has been used which contains information about both the grating (through the $\sin(\theta_B)$) and the incident beam. Since these are the parameters that were combined to give approximate criteria for ignoring the beam profile, it may be possible to simplify this criteria in terms of g. First the expression for $\sin(\theta_B)$ from the Bragg equation is inserted in the definition of g. This yields: $\sigma = d\lambda / (2nLg)$. From equation 2.30 this value of g gives an intrinsic angular width of:

$$\theta_{max} \approx 0.9 \, n \, g \frac{L}{d} \tag{4.28}$$

Using the expression for the FWHM of the angular selectivity curve from table 2.1 we can write our approximate criterial for near plane wave behavior ($\theta_{max} \leq 3 \, \Delta\theta_{FWHM}$) as:

$$3\left(0.9 \, n \, g \frac{L}{d}\right) \leq \left(0.85 \, n \frac{L}{d}\right) \quad \text{or} \quad g \leq 0.32 \tag{4.29}$$

Neither the gratings and replay beams used in generating Figure 4.8 nor those used to generate Figure 4.9 are close to meeting this criteria, thus we would anticipate seeing significant distortion in the beams. Clearly the figures indicate that this is the case, as almost all of the beam profiles presented have some amount of distortion. Since, however the parameters used for Figure 4.8 were much closer to meeting the required criteria (even if they did not meet them) then we would anticipate that the distortion would be lower in that figure. Again the data which is presented in Figure 4.8 agrees with this prediction. As mentioned above, when we look at the profiles of the beams incident away from Bragg in both figures there is additional distortion, since we are operating in a region with more non-uniformity in the attenuation..

At this point we have predictions for the on and off-Bragg beam profiles of beams transmitted through and diffracted by volume transmission gratings. It is important to check these predictions in some manner (as was done with the filtering characteristics of

66

section 4.1). To perform this checking a CCD camera was used to record the profiles of various beams diffracted from transmission gratings both on and off Bragg. Some of these measurements are shown below in Figures 4.10 and 4.11, alongside the theoretical predictions.

Figure 4.10 below shows the case of a helium neon laser beam with a gaussian width of 0.504 mm incident on a 2 mm thick grating with a period of about 0.914 microns. The medium used for the recording was photo-thermo-refractive (PTR) glass. Per equation 4.4 this represents a normalized geometry parameter of 0.91. The diffraction efficiency (which in the theoretical plot was obtained by integrating over the curve and in the experimental data was measured with a power meter) was slightly over 50%.
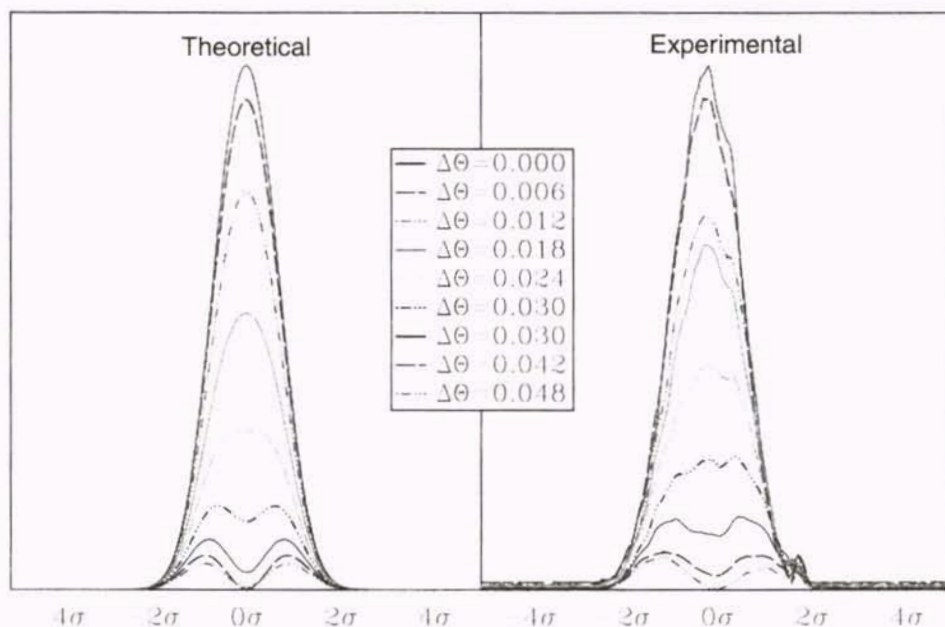
Figure 4.10. Theoretical and experimental beam profiles – g = 0.91

67

In this figure we see remarkable agreement between the theoretical and experimental plots for a wide range of incidence angles (deviations from Bragg ranged from 0 to almost 0.05°). There are however two features present in the experimental data which are not predicted – a lack of smoothness and some asymmetry to the profiles.

The first of these effects is due to the fact that the replay beam is not exactly gaussian, and contains some high frequency noise. Although this could have been reduced by adding a spatial filter, this would have complicated the system by requiring additional optics to place the waist of the beam back on the grating. The second effect is most likely due to a non-uniform grating (all theoretical predictions assumed a uniform grating). This non-uniformity arises from the grating being recorded with two gaussian beams, thus the grating strength will be strongest at the center of the overlap region and decrease as we go out from this. Even more non-uniformity will occur if the maxima of the recording beams do not coincide in the plane of recording. In this case the grating strength profile will have two maxima centered around a local minimum.

Since this was a relatively wide replay beam, it seems important to consider the more stringent case of a much narrower beam. This is shown below in Figure 4.11 where lenses were used to reduce the width of the beam to 0.25 mm, again incident on a 2 mm thick transmission grating with a slant of about 5° and a grating period of 1.02 μm. This represents a geometry factor of 2.38, approximately three times as high as above. In this case the diffraction efficiency was 30%.
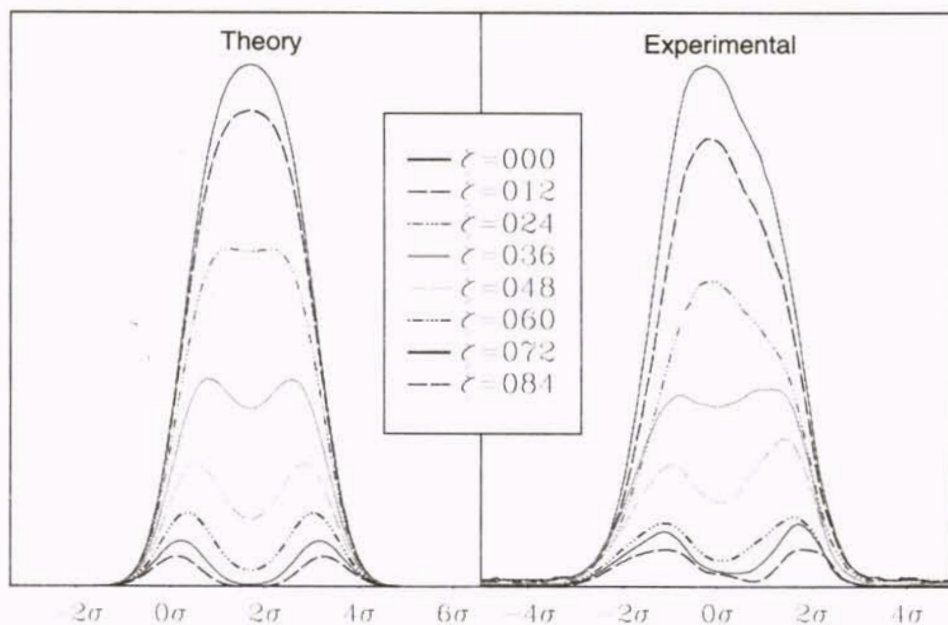
Figure 4.11.   Theoretical and experimental beam profiles – g = 2.38

Again we see good agreement between the theoretical and experimental profiles, both on and off-Bragg. Not only do the curves show agreement in the amplitudes of the beams at all cases, but the general shape of the beam (the flattening of the top, or the breaking of the beam into two portions) is also well predicted. Clearly the asymmetry discussed above is still present (and perhaps even worse), and the smoothness of the curves is still an issue.

These plots verify that when using typical laboratory helium neon laser beams we are well within the range of cases for which the theory is valid. They also serve to show how accurate the theory (including the additional off-Bragg parameter) can be, even with slightly non-uniform gratings.

4.3   Summary

High efficiency transmission gratings were recorded in photo-thermo-refractive glass and analyzed with various replay beams. Several differences were found between the measured angular selectivity and that predicted by coupled wave analysis. These were shown to result from the angular spectrum of the replay beam.

In order to investigate these finite beam effects further, a closed form solution for the off-Bragg diffracted and transmitted beam profiles of a Gaussian replay beam was derived. Using this solution it was shown that additional distortion (above that which happens on-Bragg) occurs when the devices are operated off-Bragg. This distortion (both on and off-Bragg) arises due to non-uniformity in the attenuation of the angular spectrum components. It was shown that when $\sigma > 1.6\lambda d/n L$ (or equivalently $g < 0.32$) the effects of the beam profile on the angular selectivity are minimal, and near plane-wave responses are found. All profile predictions were experimentally verified.

CHAPTER 5

DIFFRACTION OF FINITE BEAMS BY

REFLECTION GRATINGS

In the last chapter a discussion of the diffraction properties of transmission gratings was given, with the emphasis placed on the diffracted beam profiles and on their use as angularly selective filters. This chapter will perform a similar analysis for reflection gratings. Again the diffracted and transmitted beam profiles will be discussed, along with a discussion of their filtering characteristics (in terms of diffracted power versus dephasing). As in chapter four each section will begin with theoretical predictions and will finish with experimental verification.

5.1   Wavelength Filtering Characteristics

At the end of section 2.1 solutions were derived for the diffracted field and the diffraction efficiency of reflection gratings as a function of the grating strength and the amount of dephasing present. In this section the repercussions of those solutions will be

investigated in more detail. Naturally the starting point of this analysis is equation 2.22, which is repeated below for convenience.

$$S(\xi,v)=\left(\frac{c_R}{c_S}\right)^{\frac{1}{2}}\Bigg/\left\{j\xi/v+\left(1-\xi^2/v^2\right)^{\frac{1}{2}}\coth\left(v^2-\xi^2\right)^{\frac{1}{2}}\right\}$$

$$v=j\pi n_1 d\Big/\lambda\,(c_R\,c_S)^{\frac{1}{2}}\quad \xi=-\vartheta d/2c_S \qquad\qquad (2.22\,/\,5.1)$$

$$\eta=1\Bigg/\left\{1+\left(1-\xi^2/v^2\right)\Big/\sinh^2\left(v^2-\xi^2\right)^{\frac{1}{2}}\right\}$$

The filtering characteristics arising from these expressions were summarized in table 2.1, and the resulting selectivity curves were given in Figure 2.5. These plots were in terms of a general dephasing parameter, which could represent either angular or wavelength deviations from Bragg. The amount of each type of dephasing required to drop the diffracted power to half of its maximum value was also given in table 2.1. For reflection gratings this is:

$$\Delta\theta_{FWHM}=K_1\frac{L}{d\tan\theta}\qquad\frac{\Delta\lambda_{FWHM}}{\lambda_0}=K_1\frac{L}{d}\qquad\qquad (5.2)$$

where K_1 is a constant that varies roughly between 1 and 2. The presence of the tangent function in the denominator of the angular equation limits the usefulness these gratings as angular filters. This term is not present in the wavelength equation, thus the potential

exists for their use as wavelength filters. To get a feel for the degree of filtering possible with these devices, consider the example of a grating designed for normal incidence at 1.55 μm. The Bragg requirement sets the period at λ / 2n. Inserting this into equation 5.2 (assuming a value of K_1 around 1.5 and a 2 mm grating recorded in glass) yields:

$$\frac{\Delta\lambda}{\lambda_0} \cong 3.9 \times 10^{-4} \quad \text{or} \quad \Delta\lambda \cong 6 \, \text{angstroms} \tag{5.3}$$

The narrowness of this linewidth indicates strong potential for these devices as wavelength filters in DWDM systems.

These filtering characteristics were investigated experimentally in several gratings recorded in 5 mm thick samples of iron-doped lithium niobate. Since the theoretical curves were derived without regard to the type of dephasing, angular selectivity curves will be given. If reasonable agreement is found between the theoretical and experimental curves, we can assume that the model is adequate for the particular beams used on the gratings investigated, and simple scaling could be performed to see the wavelength filtering characteristics.

A typical example of an angular selectivity plot is given below in Figure 5.1. Along with the experimental data, the figure contains two theoretical plots – one for plane wave incidence (obtained from equation 5.1) and one which assumes incidence is with a Gaussian beam (obtained by combining equation 5.1 with plane wave decomposition).
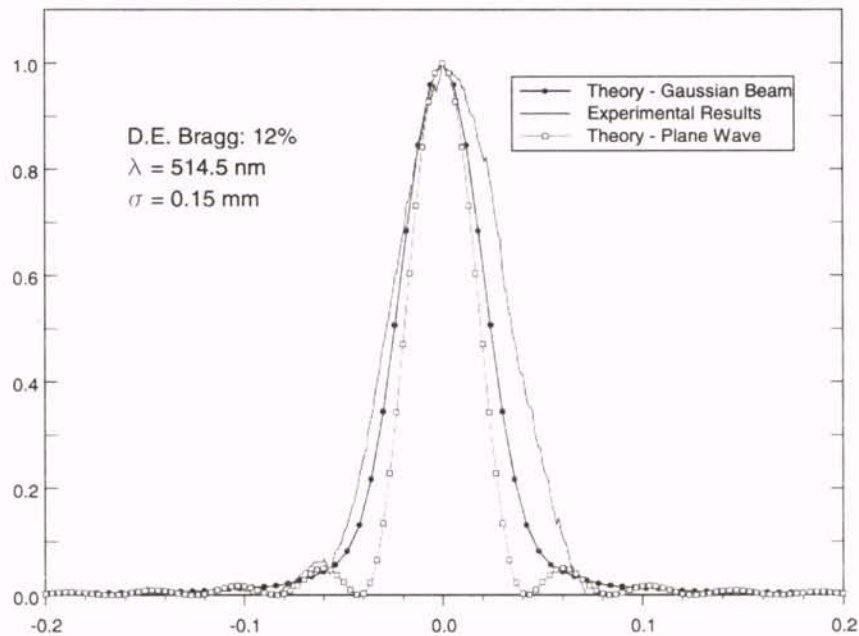
Figure 5.1. D.E. vs. Angular Dephasing: Theory and Experiment - reflection gratings

This figure has many features similar to its transmission grating counterpart (Figure 4.2). Just as in that figure, the predicted main reflection lobe for an incident plane wave is much narrower than the experimental results. Additionally, the side lobes that are present when the incident beam is a plane wave have again been washed out when the grating is replayed with a gaussian beam. As described in chapter 4 both of these effects are primarily due to the presence of an angular spectrum of input waves.

The third plot in the figure is a theoretical prediction which assumes the incident beam has a Gaussian amplitude distribution similar to the experimental beam. We see that using this profile has given some improvement in the predictions (we no longer expect sidelobes and the width of the main lobe has increased somewhat), however the match is still not as good as it was in the transmission gratings of chapter four. Most likely this is due to the grating being non-uniform.

74

This non-uniformity can happen in LiNbO$_3$ for a variety of reasons. Two examples (Gaussian recording beams and imperfect overlap) were discussed in chapter four. Along with these a taper (variation of grating strength through the thickness of the grating) will occur if the absorption of the material at the recording wavelength is relatively high. In LiNbO$_3$ this absorption depends strongly on the level of doping, however it can easily exceed 6 cm^{-1}, thus taper must be considered in thick gratings. The diffraction characteristics of tapered gratings are addressed in detail in chapter 9, thus a discussion of the effects will be left until then. Another cause of non-uniformity in LiNbO$_3$ is the fact that it is a real time material. In these materials the grating is present as the recording is taking place. This causes the recording beams themselves to be diffracted, leading to a non-linear chirp (variation of the period) throughout the grating thickness. Modeling this effect requires the grating to be sliced into a number of sections, the solution of each which becomes the input to the next until the grating is traversed. Since this effect is a function of the grating strength, not the recording time or power, it can not be alleviated by simply recording with a more powerful beam for less time (or a less powerful beam for more time). Finally some non-uniformity can arise due to erasure of the grating during replay (since the erasure will be highest where the intensity of the replay beam is highest). It is most likely that the mismatch between the theoretical predictions and the experimental data in these plots is due to a combination of all of these non-uniform grating effects.

An additional difference between Figure 5.1 above and the corresponding plot for a transmission grating is the scale of the x-axis. In Figure 4.2 most of the power was

confined in an angular region of about 0.1° (experimentally), while this has increased to about 0.15° in Figure 5.1. One might have expected a much larger increase in this parameter due to the presence of the tangent function described above. The reason that the parameter has not increased more dramatically is that Figure 4.1 was generated with an on-Bragg diffraction efficiency of approximately 51%, while that of Figure 5.1 was only 12%. Increasing this efficiency to a comparable level would have significantly increased this value, as it directly affects the value of K_1 in equation 5.2.

Since reasonable agreement was obtained between the experimental curves and theoretical predictions for angular dephasing, the use of the model in these materials has been validated, and the wavelength filtering characteristics (important for DWDM systems) can be obtained by simple scaling of this dephasing parameter.

5.2   Diffracted and Transmitted Beam Profiles

The previous section analyzed the angular and wavelength filtering properties of reflection holograms. This analysis included the profile of the replay beam for an improved match between the theoretical and experimental data. This was done using Fourier decomposition, since 2-D coupled wave analysis solutions are quite limited for reflection gratings [1-2]. Just as with the transmission gratings, we can obtain the diffracted and transmitted beam profiles from this analysis. After a brief discussion of normalized parameters, the remainder of this chapter will consider these profiles in detail.

It was shown in chapter 4 that when the additional requirements of 2-D CWA were added to Fourier decomposition, the two methods were mathematically equal. This allowed the normalized parameters of 2-D CWA to be retained for analyses performed with plane wave decomposition, which in turn allowed many combinations of recording beams and gratings to be considered with only a few plots. It would be nice to have similar normalized parameters for reflection gratings, however the lack of a two-dimensional solution suggests that these may not be available. At the same time, since normalized parameters existed when Fourier decomposition was used on transmission gratings (even if they weren't apparent), then it is possible that they exist for reflection gratings also.

It is obvious from Kogelnik's solutions for reflection gratings [3] (eq. 5.1), that the grating strength parameter will remain normalized for reflection gratings. Since his solutions were not intended for a two-dimensional analysis however there is nothing to suggest that the geometry factor would also remain normalized. Nonetheless a numerical investigation was performed to see if this would hold, and in all cases keeping this factor constant did result in the same output. It thus appears that this factor does remain normalized, and we will therefore use these same parameters (defined in equation 4.4) to show general cases for the reflected beam profiles.

The coordinate system used for this analysis is shown in Figure 5.2 below. All diffracted beam plots are given along the direction s, while any transmitted profiles are given along the direction r. This system will be used for both on and off-Bragg incidence.

$$r = -z \sin\Theta + x \cos\Theta$$
$$s = z \sin\Theta + x \cos\Theta$$
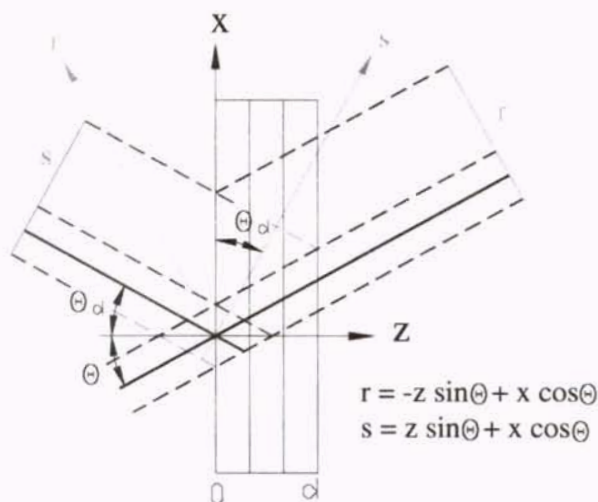
Figure 5.2. Geometry used for 2-dimensional reflection grating beam profiles

Our investigation of the diffracted beam profiles will begin with the exact Bragg incidence case (in transmission gratings Bragg incidence had been studied previously [4-6]). Our first results (which are given below in Figure 5.3) are for a geometry factor of 1. If we assume a 2 mm thick grating recorded in material of index 1.5, this represents an incident gaussian beam with a width of 0.14 mm at 633 nm. The figure contains diffracted profiles for gratings with strengths ranging from 0 to π.

A few things are instantly apparent when looking at this plot. First, as the grating strength is increased the profile of the diffracted beam becomes significantly better. This is in complete contrast to transmission gratings, where the opposite effect has been reported [4]. This effect can be easily explained by considering the properties of these gratings under plane wave illumination, along with the angular spectrum of the input beam. Figure 2.5, which gives the diffraction efficiency of a reflection grating vs. dephasing, shows that as the strength of the grating is increased, the grating became less

sensitive to angular or spectral dephasing. Thus, by increasing the strength of the grating, the attenuation of the angular spectrum components becomes more uniform, resulting in a better diffracted beam profile.
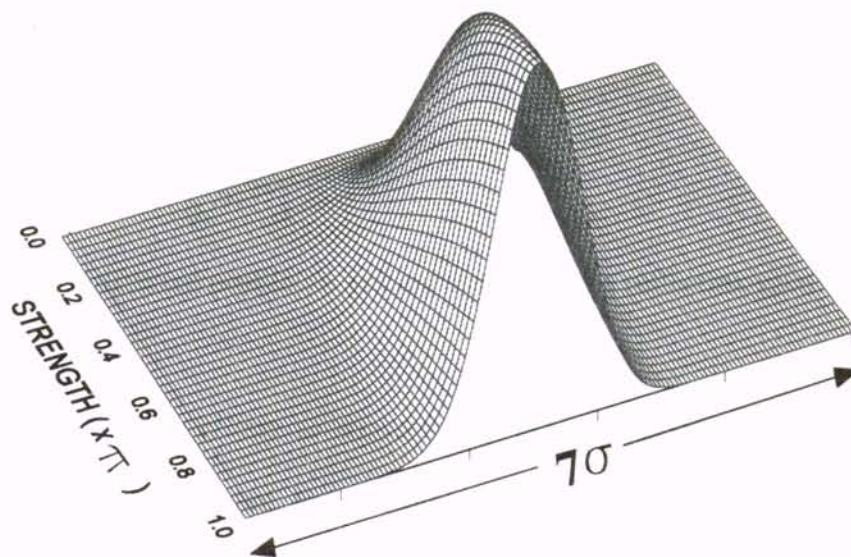
Figure 5.3. Diffracted profile vs. grating strength for a reflection grating ($g = 1$)

As will be shown below these effects will be more pronounced when the incident beam is narrowed. In Figure 5.4 the diffracted profiles of the same grating are investigated when the geometry factor has been increased to 3. This is equivalent to decreasing the width of the incident beam by a factor of three.
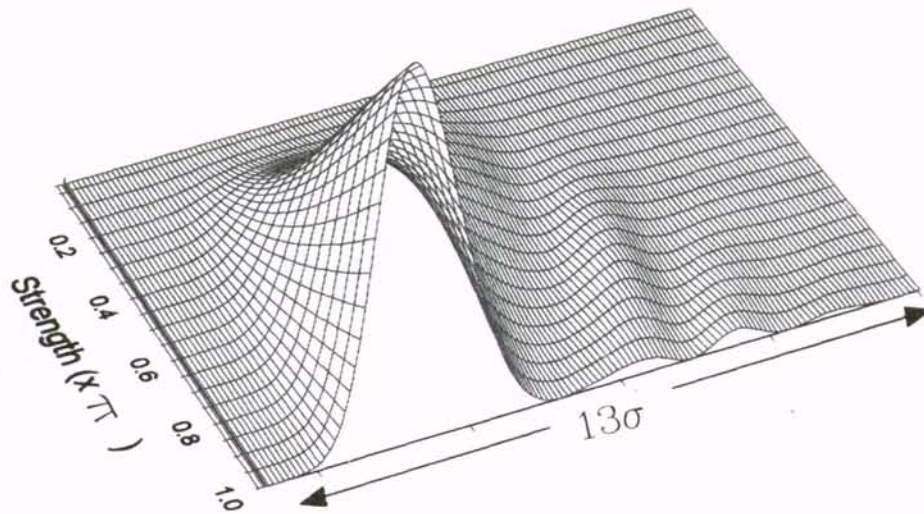
Figure 5.4. Diffracted profile vs. grating strength for a reflection grating ($g = 3$)

Again we see similar effects as above (when the grating strength increased, the fidelity with which the diffracted beam matched the incident one also increased), however with this narrow an incident beam there is considerable distortion, even with a grating strength as high as π. Obviously these are both because the narrow beam has a much wider angular spectrum, and the attenuation of these components is highly un-even. Additionally we see that the diffracted beam has spread out considerably (it is approximately 4 times as much as in the previous case).

One final thing that must be considered is the diffraction efficiency of finite beam when the incident beam is on-Bragg. This is shown below in Figure 5.5 for three different incident beams (a plane wave and Gaussian beams of two different widths). From this plot it is clear that the effect of a wider angular spectrum is to decrease the overall Bragg diffraction efficiency. Additionally we see that when the geometry factor

is below 0.1 the finiteness of the beam has almost no effect on the overall diffraction efficiency.
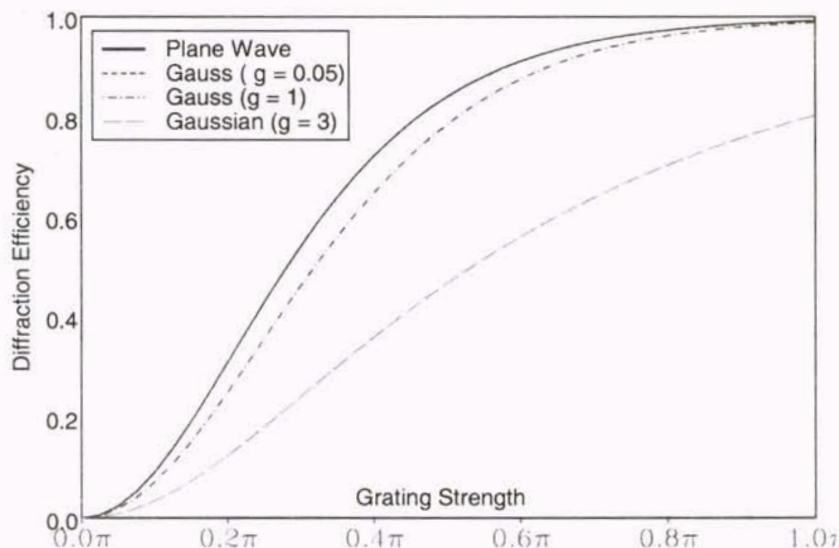
Figure 5.5. Diffraction efficiency for various input beams on reflection gratings

Now that we have looked at the angular selectivity curves as well as the important on-Bragg effects (beam profile and diffraction efficiency), we are ready to investigate the final effect - beam profiles when the incidence is off-Bragg. In all cases covered the dephasing parameters are calculated from equation 5.1, and were obtained by varying the angle of incidence (selected arbitrarily). Plots are included for grating strengths of $\pi/4$, $\pi/2$ and $3\pi/4$, with dephasing parameters ranging from 0 to 3. Again geometry factors of 1 and 3 are shown. These dephasing parameters correspond to angular deviations from Bragg of $0.000°$, $0.027°$, $0.054°$ and $0.082°$ respectively.

The case of g=1 is given below in Figure 5.6. As described above, if the grating is 2 mm thick in a material of index 1.5, this represents an incident gaussian beam with a width of 0.14 mm at 633 nm.
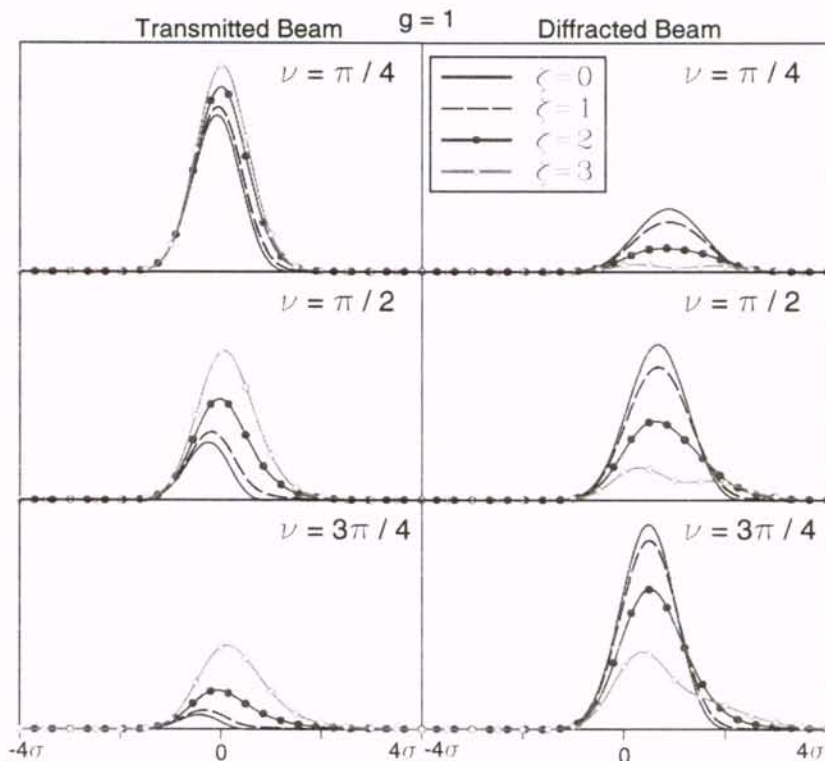
Figure 5.6.   Diffracted and transmitted beam profiles, reflection gratings – g = 1.0

As in the analysis of transmission gratings the discussion starts with the diffracted beam. Again we see that as we deviate from Bragg the decrease in diffracted power is accompanied by additional distortion. Not unexpectedly this distortion increases as we go further from Bragg, again to the point where a beam which looks good on-Bragg becomes unrecognizable. From these plots it is hard to make quantitative statements, but

82

it appears that the off-Bragg profiles do improve (albeit slightly) with increased strength (i.e. for a given amount of dephasing a stronger grating gives a slightly better profile).

When viewing the transmitted beams we see effects similar to transmission gratings, with the best profiles away from Bragg (where little power is coupled from them). One generalization seen by Figure 5.6 is that the coupling of energy from the beam is not uniform, it occurs principally from the side of positive r (see Figure 5.2).

The diffracted and transmitted profiles which occur with a narrower incident beam are shown in Figure 5.7 below. This is similar in content to Figure 5.6, however the incident beam width has been decreased by 2/3 (which results in a g value of 3).
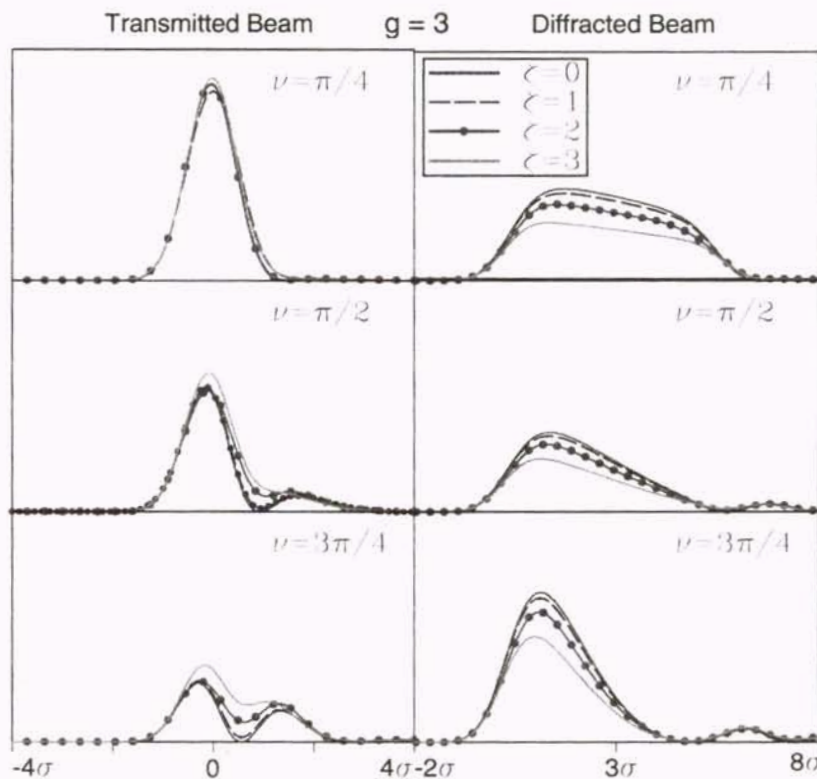
Figure 5.7. Diffracted and Transmitted beam profiles, reflection gratings – g = 3.0

We see with this narrow an incident beam that all of the Bragg profiles are significantly distorted, even when the grating strength reaches $3\pi/4$. For the two lower strength values ($\pi/4$ and $\pi/2$) the diffracted beam appears almost flat-topped due to the near truncation of the highest spatial frequency components. It is not until the strength has increased to $3\pi/4$ that any resemblance (albeit small) of a Gaussian beam is present.

In each of these plots there is minimal additional distortion as our incidence deviates from Bragg. Essentially for every strength the profiles have the same shape as their on-Bragg counterparts, however their amplitude is decreased. This is similar to the effects seen in Figures 4.9 where the non-uniform attenuation on-Bragg causes enough distortion that the additional off-Bragg distortion is barely noticible.

Again we can summarize these figures in the following manner. If the characteristics of the filter are such that on-Bragg significant non-uniformity exists in the attenuation, then any additional off-Bragg effects (where the slope of Figure 2.5 is highest) will be less pronounced. Essentially the grating acts as a low pass filter for angular spectrum components whose bandpass increases as the grating becomes stronger.

We now have theoretical predictions for the on and off-Bragg beam profiles diffracted by and transmitted through reflection gratings. Again we would like to check these predictions in real materials, and to do so several gratings were recorded in 5 mm thick samples of iron-doped LiNbO3. The diffracted beam profiles were then recorded with a CCD camera, and the results are given below in Figures 5.8 and 5.9.

Figure 5.8 covers the case when the grating was replayed with the beam from an argon laser ($\lambda = 514.5$ nm) with a width of about 1.7 mm. Per equation 4.4 this

represents a normalized geometry parameter of 0.13. The diffraction efficiency in these plots was 2.1%. This value was obtained experimentally by direct measurement and theoretically by performing a numerical integration of the curve.
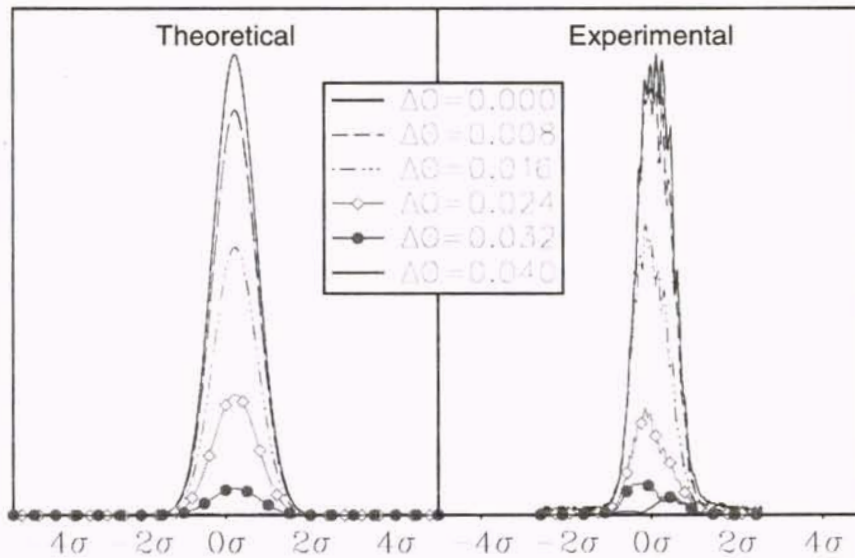
Figure 5.8. Theoretical and experimental beam profiles – g = 0.13

In this figure, aside from noise due to undersampling we see reasonable agreement between the theoretical and experimental plots for a wide range of angles of incidence (deviations from Bragg ranged from 0 to almost 0.04°). There is again a large amount of asymmetry to the experimental profiles. As discussed above this is most likely due to the grating being non-uniform (either from the amplitude profiles of the recording beams, imperfect overlap, or from taper and chirp effects which arise from the real-time nature of the material).

This replay beam was however somewhat wide, and it seems prudent to also consider the more stringent case of a much narrower beam. This situation is shown in Figure 5.9 below which was performed with an argon laser beam shrunk to $\sigma = 0.2$ mm, again incident on a 5 mm thick reflection grating in iron-doped lithium niobate. The geometry factor corresponding to this situation is approximately 1.13 (roughly 8 times that of the previous case). Again the diffraction efficiency is 2.1%.
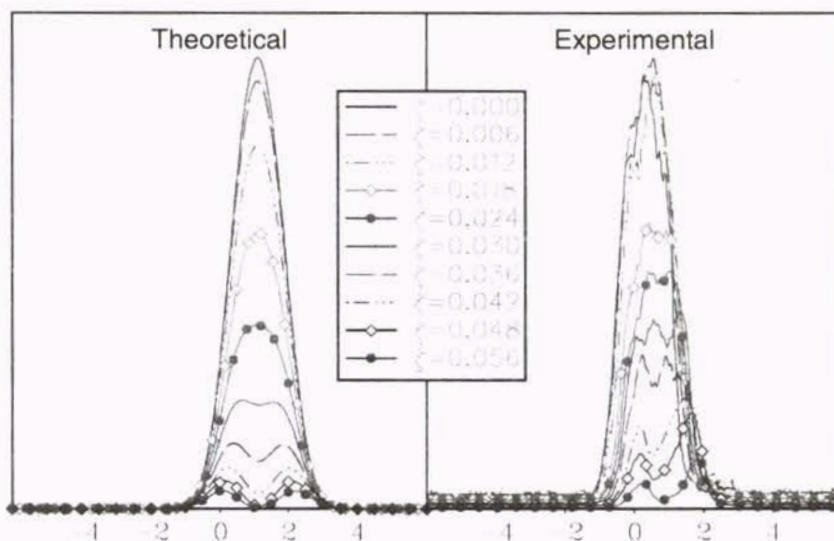
Figure 5.9 Theoretical and experimental beam profiles – g = 1.13

Although the match isn't as good as was obtained with the wider beam, even in this stringent case we still see decent agreement between the theoretical and experimental profiles, both on and off-Bragg. Along with a reasonable conformity in the amplitudes of the beams for each dephasing amount the general shape of the beam (the flattening of the top, or the breaking of the beam into two portions) is also well predicted. As in previous

cases, these plots give a good verification that we are well within the range of cases for which the theory is valid.

5.3   Summary

The diffraction efficiency and resulting profiles of finite beams diffracted from uniform reflection gratings on and off-Bragg were investigated. With regard to the diffraction efficiency the effect of replaying with a finite beam will be an overall decrease in the level. For a typical He-Ne laser beam this decrease will be about 5% when the grating strength is near $\pi/2$.

The profiles diffracted from these reflection gratings will generally be distorted versions of the replay beams, with the distortion increasing as incidence deviates from Bragg. Unlike in transmission gratings however, the fidelity of the on-Bragg profiles improve as the grating strength is increased. This happens because stronger reflection gratings have wider regions of approximately uniform attenuation. Again, all profile predictions were verified experimentally.

CHAPTER 6

OVERLAP GRATINGS

The previous two chapters dealt with investigating the properties of volume transmission and reflection gratings. In those chapters the gratings were assumed infinite with plane parallel front and back boundaries. The type of hologram present (reflection or transmission) was determined by the angle that the grating lines made with those surfaces. In this chapter another type of grating known as crossed beam [1] or overlap type will be studied. These generally arise due to finiteness of either the recording beams or the holographic material. Examples of overlap gratings arising from each of these situations are shown below in Figure 6.1. Regardless of the cause of the overlap, the end result will be a grating with non-uniform or non-planar boundaries. The solution to all of these problems will be found using the two-dimensional coupled wave analysis described in section 2.3. This is chosen for its ease in handling these non-planar boundaries.

This chapter will investigate the diffraction characteristics of these gratings both on and off-Bragg by looking at three specific examples of subject beams: a Gaussian beam, a truncated uniform beam, and an array of Gaussian beams. These beams have been selected to show effects that are general to overlap holograms, as well as some that are critical to systems currently under design.
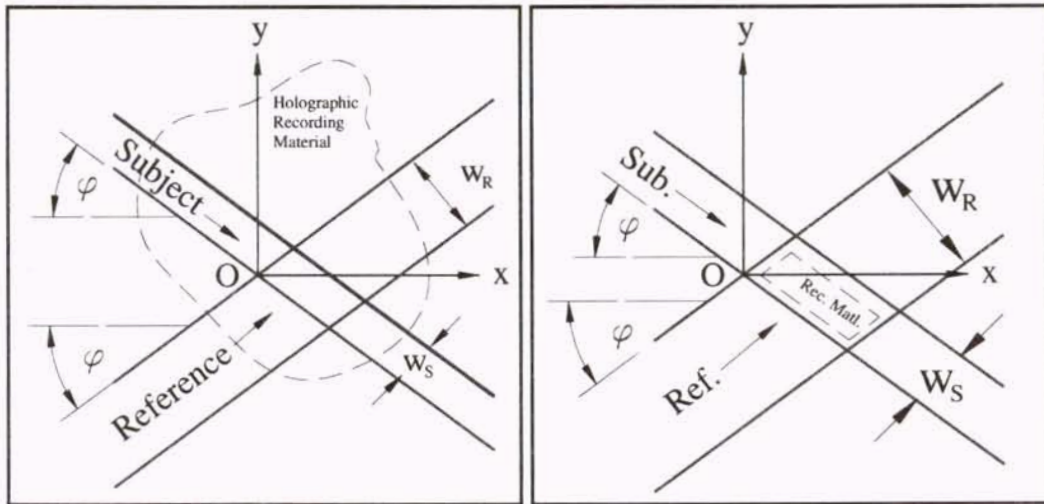
Figure 6.1. Types of Overlap Gratings

6.1   Coupled Wave Equation Solution for Overlap Holograms

In section 2.3 equations for two-dimensional volume holograms were derived. These expressions were given in equation 2.50, and are repeated below.

$$
\frac{\partial R}{\partial s} = -j\kappa' \frac{A_{R0}\,A_{S0}\,A_S}{A_R} S
$$

$$
\frac{\partial S}{\partial r} A_S + j\vartheta' S = -j\kappa' \frac{A_{R0}\,A_{S0}\,A_R}{A_S} R
$$

$$(2.50 / 6.1)$$

Now that we have these equations for overlap gratings we need to solve them to get the beam profiles and diffracted power. In order to obtain a solution using known methods, their form must be modified slightly, and an additional approximation must be

89

made. The required approximation is that $A_{r0}(r,s) \approx A_{r0}(r)$, and $A_{s0}(r,s) \approx A_{s0}(s)$. This is equivalent to forcing the beams to be read back at an angle close to that at which they were recorded. It must be noted that this does not nullify its use in angular dephasing problems (since the magnitude of the deviation is small enough in these instances ($<< 1°$) that the approximation remains valid). It does however restrict the validity of the solution to cases where the readback beam is the same wavelength as the recording one (since changing the wavelength requires a large change in incidence angle to satisfy the Bragg condition). With these approximations equation 6.1 reduces to:

$$\frac{\partial R}{\partial s} = -j\kappa' \frac{A_{R0}(r)A_{S0}(s)A_{S}(s)}{A_{R}(r)} S$$

$$\frac{\partial S}{\partial r}A_{S} + j\vartheta'S = -j\kappa' \frac{A_{R0}(r)A_{S0}(s)A_{R}(r)}{A_{S}(s)} R$$

6.2

Now, by making the following two substitutions into each of the equations:

$$R = \overline{R}\frac{A_{R0}(r)}{A_{R}(r)}e^{-j\vartheta'r} \quad \text{and} \quad S = \overline{S}\frac{A_{S0}(s)}{A_{S}(s)}e^{-j\vartheta'r}$$

6.3

We get:

$$\frac{\partial \overline{R}}{\partial s} = -j\kappa' A_{S0}(s)^2 \overline{S}$$

$$\frac{\partial \overline{S}}{\partial r} = -j\kappa' A_{R0}(r)^2 \overline{R}$$

6.4

These equations can now be combined into a single second order-linear hyperbolic differential equation. This can be done in terms of either the reference or the signal beam, with the resulting equation being one of :

$$\frac{\partial \overline{S}}{\partial r \partial s} = -\kappa^{2'} A_{R0}(r)^2 A_{S0}(s)^2 \overline{S}$$

$$\frac{\partial \overline{R}}{\partial r \partial s} = -\kappa^{2'} A_{R0}(r)^2 A_{S0}(s)^2 \overline{R}$$

6.5

Obviously the first of these is used when the signal beam profile is desired, while the second is used to determine that of the reference beam. From Figure 6.1 we see that the boundary conditions are that the replay beam has an amplitude of R_0 at the boundary s=0, and that the amplitude of the signal beam is 0 at r=0. These can be written as:

$$R(r,s) = R_0\big|_{s=0} \quad \text{and} \quad S(r,s) = 0\big|_{r=0} \tag{6.6}$$

Writing these in terms of our new variables (equation 6.3) and incorporating equation 6.4 results in the following boundary conditions:

$$\frac{\partial \overline{S}}{\partial r} = -j\kappa' A_{r0}(r,y) A_r(r,y) e^{+jK_1(r)} R_0 \quad \text{along the line } s = 0$$

$$\frac{\partial \overline{R}}{\partial s} = 0 \qquad\qquad\qquad\qquad\qquad \text{along the line } r = 0$$

(6.7)

At this point we can write our final expressions for the beam amplitudes, which we will do starting with the signal beam. The solution will be obtained using Reimenn's technique for second order linear hyperbolic partial differential equations. A detailed description of this method (from Courant & Hilbert [2]) can be found in appendix 1, where two overall forms of the solution are derived. Essentially the method gives a general form for the solution, into which problem specific parameters and boundary conditions are inserted. In calculating the signal beam we will use form 1 of the solution (equation A1.14.a), which is:

$$\overline{S}(r,s) = \overline{S}(-\infty,s) - \int_{-\infty}^{s} \overline{S} \frac{\partial G}{\partial s'} ds' + \int_{-\infty}^{r} G \frac{\partial \overline{S}}{\partial r'} dr' \tag{6.8}$$

In this equation G is the Riemann function, which represents any function that meets a number of requirements. These requirements are listed in equation A1.10 of appendix 1. It is easy to show that one such function that meets these requirements is:

$$J_0\left(2\kappa' \sqrt{v(s,y)(u(r,y) - u(r',y))}\right), \text{ where}$$

$$u(r,y) = \int_{-\infty}^{r} A_{r0}^2(r',y) dr' \quad \text{and} \quad v(u_s,y) = \int_{-\infty}^{s} A_{s0}^2(s',y) ds'$$

Inserting this for the Riemann function in 6.8, along with the boundary conditions from equations 6.6 and 6.7 gives the following solution for the signal beam at the output face:

92

$$\overline{S}(r,s) = -j\kappa' R_0 \int_{-\infty}^{r} A_{r0}(r,y)A_r(r,y)e^{jK_1(r)}J_0\left[2\kappa'\sqrt{v(s,y)(u(r,y)-u(r',y))}\right]dr' \quad (6.9)$$

In order to determine the reference beam solution we need to use form two of the solution (equation A1.14.b), which is:

$$\overline{R}(P) = \overline{R}(r,-\infty) + \int_{-\infty}^{s} G\frac{\partial \overline{R}}{\partial s'}ds' - \int_{-\infty}^{r} \overline{R}\frac{\partial \overline{G}}{\partial r'}dr' \quad (6.10)$$

Again inserting our boundary conditions, along with the Riemann function from above yields the following solution:

$$\overline{R}(r,s) = \frac{A_r(r,y)e^{jK_1(r)}}{A_{r0}(r,y)} - j\kappa' R_0 \int_{-\infty}^{r} \frac{A_r(r,y)R_0}{A_{r0}(r,y)}\left(\frac{\kappa'\sqrt{v(s,y)}A_{r0}^2(r,y)}{\sqrt{u(r,y)-u(r',y)}}\right) \quad (6.11)$$
$$\times J_1\left[2\kappa'\sqrt{v(s,y)(u(r,y)-u(r',y))}\right]dr'$$

These equations are now transformed back into R and S, which yields:

$$R(r,s) = R_0 - R_0\frac{A_{r0}(r,y)e^{j(\delta\cdot r)}}{A_r(r,y)}\kappa'\int_{-\infty}^{r}\frac{A_r(r,y)A_{r0}(r,y)e^{-j(\delta\cdot r)}\sqrt{v(s,y)}}{\sqrt{u(r,y)-u(r',y)}}$$
$$\times J_1\left[2\kappa'\sqrt{v(s,y)(u(r,y)-u(r',y))}\right]dr'$$
$$\quad (6.12)$$

$$\overline{S}(r,s) = \frac{-j\kappa' R_0 A_{s0}(s,y)e^{-j(\delta\cdot s)}}{A_s(s,y)}\int_{-\infty}^{r} A_{r0}(r,y)A_r(r,y)e^{-j(\delta\cdot r)} \times J_0\left[2\kappa'\sqrt{v(s,y)(u(r,y)-u(r',y))}\right]dr'$$

93

6.2   Resulting Beam Profiles

Now that we have general solutions for the diffracted and transmitted beams of overlap gratings, we can use them to show the resulting profiles and diffraction efficiencies for various inputs. In order to present results, specific recording and replay beams must be selected. Since the profile of a typical laser beam is approximately Gaussian, this distribution will be used for the reference beam in all cases. Three different subject beams have been chosen: a Gaussian beam of width σ_{s0}, an array of five Gaussian beams (each of width σ_{s0}), and a truncated uniform beam.

Each of these profiles has been selected for a different reason. The Gaussian beam is used because it is the situation that naturally occurs when the grating is made with standard laser beams. The uniform beam is valuable because its simplicity allows small profile distortions to be seen most easily. Finally, the array of Gaussian beams has been selected because this is the manner in which typical holographic data storage systems are operated (i.e. the signal beam consists of a number of spatially separated bits formed into a page of data. These pages are then angularly and spatially multiplexed to avoid crosstalk. This investigation will therefore look for effects that could be important in practical implementations of these systems.

Although expression 6.12 was developed for replay with an arbitrary beam, this analysis will be limited to replay with Gaussian beams of width σ_{r0} (ie: reading back with the original recording beam). Additionally, although the theory can provide three dimensional profiles (by slicing the beam into a number of two-dimensional sections and

solving each separately), generating these profiles is numerically intensive, and little additional information is gained from the off-zero sections; therefore only the section at y=0 will be shown. In situations where the profile of the beam's central portion is not typical of the entire beam or when the entire profile can not be ascertained from its central portion the full three-dimensional profile will be given.

Case 1: Gaussian Subject Beam

As discussed above, the first example will examine the diffraction properties when the grating was recorded with two Gaussian beams. For this example the width of the reference beam is taken to be σ_r and that of the signal beam to be σ_s, where $\sigma_s = 10 \sigma_r$. These recording and replay geometries are shown below in Figure 6.2.
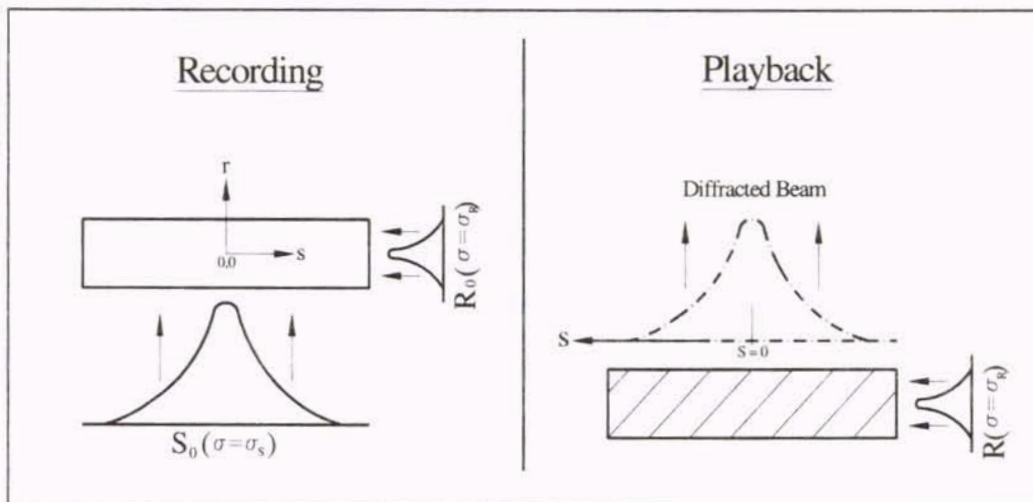
Figure 6.2. Recording and Reply: Case I - Gaussian Subject Beam

The diffracted beams which result from this grating are shown below in Figure 6.3. This figure covers overall Bragg efficiencies of 50%, 75% and 85%. For each of these, plots are given for dephasing amounts of 0 (on-Bragg) 1, 2 and 3, along with an appropriately scaled Gaussian beam. This additional beam allows the diffracted profile to be compared to the recording beam for checking the fidelity of reconstruction.
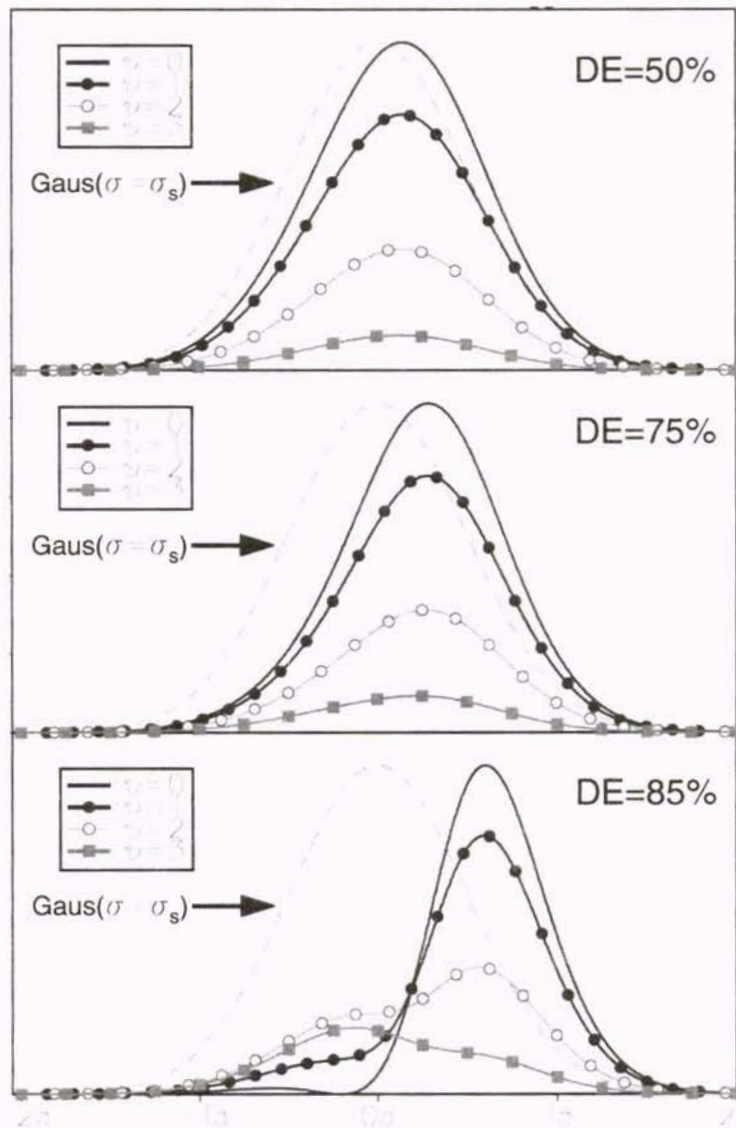
Figure 6.3. Diffracted beam profiles – single gaussian subject beam

For each efficiency the discussion will begin with the on-Bragg profiles. We essentially see two effects in all of these profiles, a shift in the location of the beam toward positive values of s (see Figure 6.2), and an asymmetry (the right side falls off quicker than the left). In the first plot (50% D.E.) this asymmetry is minimal, and the location shift is less than $\sigma/5$, however each of these effects becomes more pronounced as the overall efficiency increases. By the time this efficiency reaches 85%, the shift amount has become nearly σ.

The reason for each of these effects can be found by looking back at Figure 6.2 (the playback geometry). As the replay beam starts diffracting, its amplitude is decreased and beam depletion effects occur. This causes the right side of the diffracted beam (where the amplitude of the replay beam is highest) to have more power than the left. Essentially we are seeing the effects of violating the first Born approximation (which is typically made in calculating hologram outputs). Obviously these issues are more pronounced as the grating strength increases.

In addition to the distortion and shift discussed above, a third occurrence has shown up when the diffraction efficiency reached 85%. In looking at the on-Bragg characteristics we see that the amplitude drops to zero for a bit just beyond the center of the graph – essentially the beam has split into two. It will be shown that this effect is fundamental to overlap holograms, and results in an absolute limit to the diffraction efficiency with which a subject beam can be faithfully reproduced.

Some discussion of the off-Bragg profiles is now in order. As in previous chapters it is difficult to make sweeping generalizations. For all efficiencies however, the

off-Bragg profiles are more distorted than their on-Bragg counterparts. This distortion is not extreme when the overall diffraction efficiency is below 75% (the diffracted profiles resemble those found on-Bragg, just with less power). At a diffraction efficiency of 85% however, the resemblance of the on and off-Bragg plots quickly falls off with dephasing, and much of the structure of the diffracted beam is lost. These are interesting effects, since these profiles are caused more by the diffraction process itself than by beam depletion (less of the replay beam has been diffracted away in this case). Additionally we see that the off-Bragg beams are no longer split into multiple sections. Aside from these simple statements very few generalizations on the off-Bragg behavior can be made.

Case 2: Array of Five Gaussian Subject Beams

The second overlap grating structure to be investigated is that which results from the recording of a Gaussian reference beam of width σ_R and a subject beam which consists of an array of five Gaussian Beams, each of width σ_R. The separation of the maxima in this example is 4 σ_R. As noted above this is an important geometry to consider because it represents the type of subject beam that is typically used in holographic data storage systems. The recording situation is shown below in Figure 6.4, alongside the replay geometry.
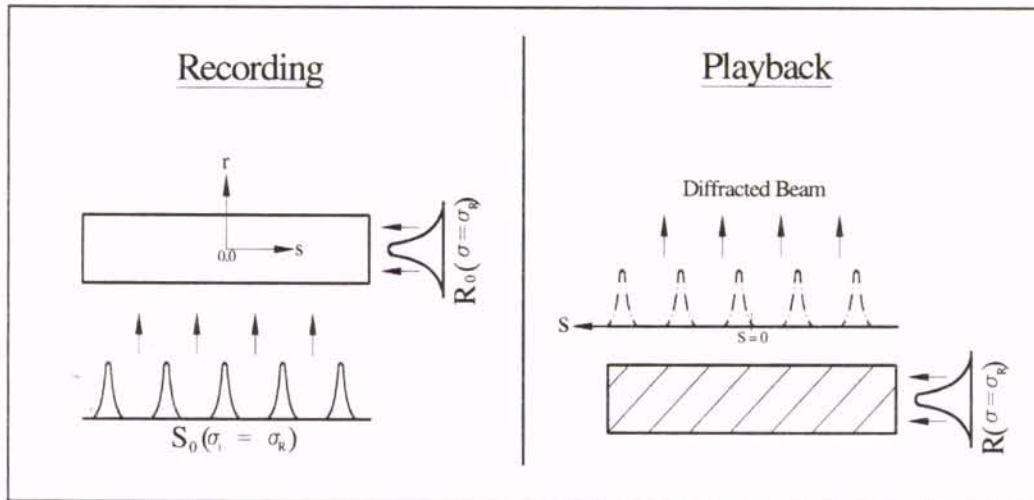
Figure 6.4. Recording and Reply: Case 2 – Five Gaussian Subject Beams

As in the previous example the analysis will consider the diffracted beam profiles when the overall diffraction efficiency is 50%, 75% and 85%. By overall efficiency we are referring to the percentage of the incident beam that has been diffracted into all of the output beams (regardless of whether each has the same amplitude). Again the dephasing parameters considered (found from equation 2.49) run from 0 (Bragg) through 3. These results are shown below in Figure 6.5. For the sake of clarity the recording beams have been omitted from these graphs. These profiles will, however be given later.

Again the on-Bragg results will be discussed first. We see in Figure 6.5 that when the overall diffraction efficiency is 50%, each of the individual beams remains approximately Gaussian. This does not imply that beam depletion effects are not present – they are. In this case they are manifested as a decrease in beam amplitude as we go from positive s (where the amplitude of the replay beam is highest) to negative.

99
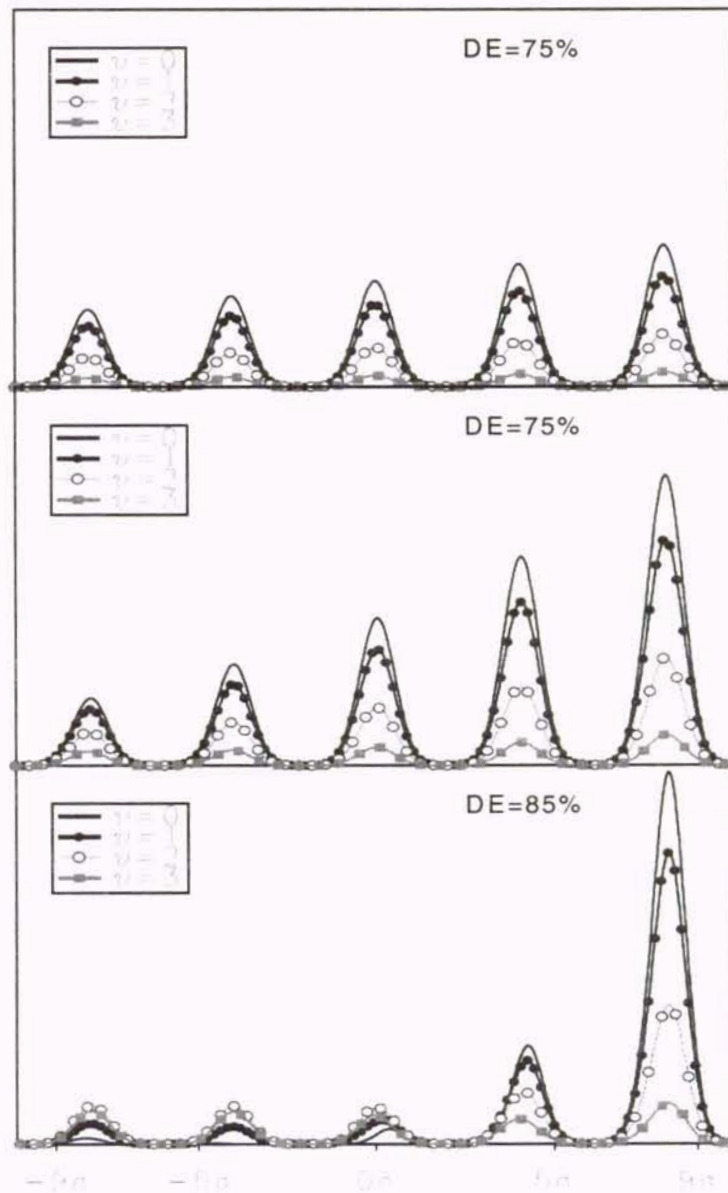
Figure 6.5. Diffracted beam profiles – array of five gaussian subject beams

This result becomes considerably larger when the overall diffraction efficiency reaches 75%. At that point the amplitude of the first diffracted beam is more than 4 times that of the fifth. Even more non-uniformity occurs when the diffraction efficiency

100

reaches 85%, and in that case it is combined with severe distortion of the third through fifth beams. In addition to this distortion we see that the amplitude of the fourth beam has dropped to almost zero. As will be shown in the final section of this chapter, this is equivalent to the break-up of the diffracted beam in Figure 6.3.

Just as was the case when the signal wave was a wide Gaussian beam, the latter bits are shifted along the s axis in the positive direction. This is difficult to see in Figure 6.5 due to the number of plots present in each graph, so Figure 6.6 has been included below. This figure shows a blow up of the on-Bragg diffracted profiles for overall diffraction efficiencies of 75% and 85%, along with the recording subject beams. In this figure the shift of the diffracted beam is clearly visible.
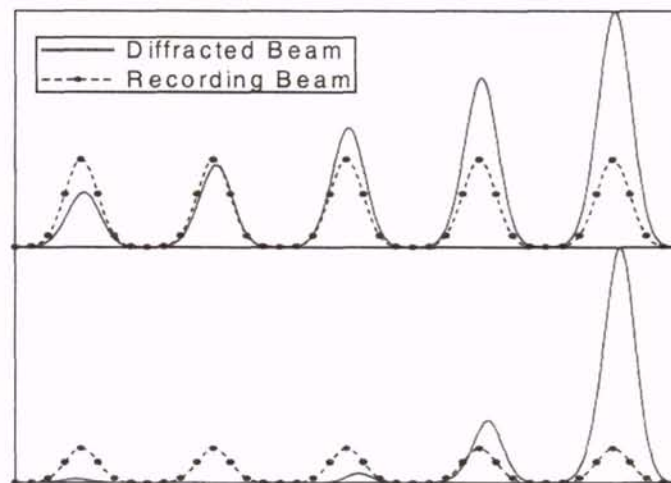
Figure 6.6. Detail of 75% and 85% five beam array diffracted profiles

Each of these two effects is of critical importance in holographic data storage systems. The shifting is important because the detection of bits in these systems is done

with an array of photocells. In this case if a bit is shifted it could cause a false reading in either its own cell or in an adjacent cell. The shifting seen in Figure 6.6 does not appear large enough for this to happen because a large spacing (4σ) was used between bits.

The other issue (large amplitude variation) causes additional concerns because each detector has minimum and maximum threshold levels. If a large number of bits are present then the last bit may not have enough power to meet this minimum threshold. Increasing the replay beam power to account for this could cause the first beam to exceed its maximum threshold. One solution to these issues is to operate at a low efficiency (where the first Born approximation holds). Another will be investigated in chapter 7.

When we look at the off-Bragg profiles of Figure 6.5 we see that the effects are similar to those that occur in the single beam case. In the 50% and 75% total diffraction efficiency cases the off-Bragg profiles look like reduced amplitude versions of their on-Bragg counterparts. Additionally, beam depletion effects are still present in these off-Bragg plots (up to about $\xi=2$). Finally, in the off-Bragg case at 85% diffraction efficiency the fourth bit has returned. As above, generalizations of this behavior are difficult to make.

Case 3: Truncated Uniform Subject Beam

The final case to be investigated is the recording of a Gaussian reference beam with a truncated uniform subject beam. The recording and replay situations are shown below in Figure 6.7.
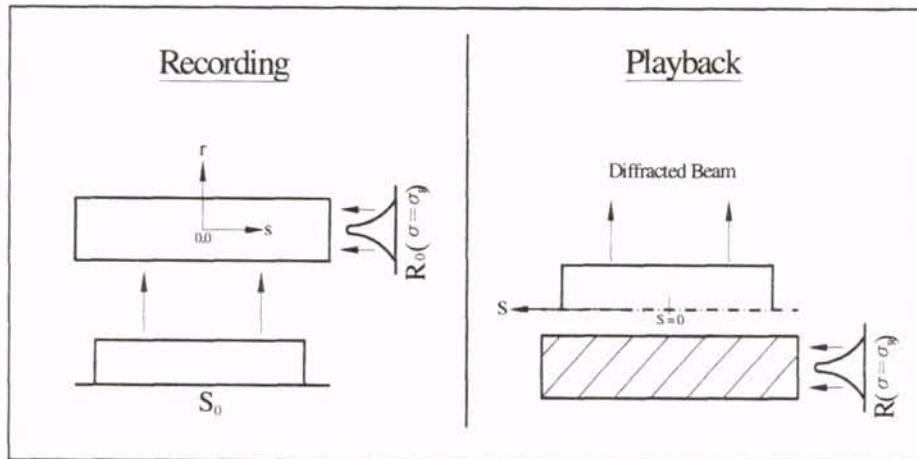
Figure 6.7. Recording and Reply: Case 3 – truncated uniform subject beam

The diffracted beam profiles of this grating are given below in Figure 6.8. As before overall diffraction efficiencies of 50%, 75% and 85% are investigated. Again for each of these efficiencies plots are included for dephasing amounts of 0, 1, 2 and 3.

In looking at the plots we see that just as in the two previous cases, when the total diffraction efficiency is 50% the resulting beam is quite uniform, and an excellent reproduction of the recording beam is obtained (the fidelity of reproduction is high). Naturally, for reasons described above the right side of the output (that which interacts with the replay beam first) is higher, but this effect appears insignificant. Off Bragg this effect is less pronounced since less depletion occurs in this case.
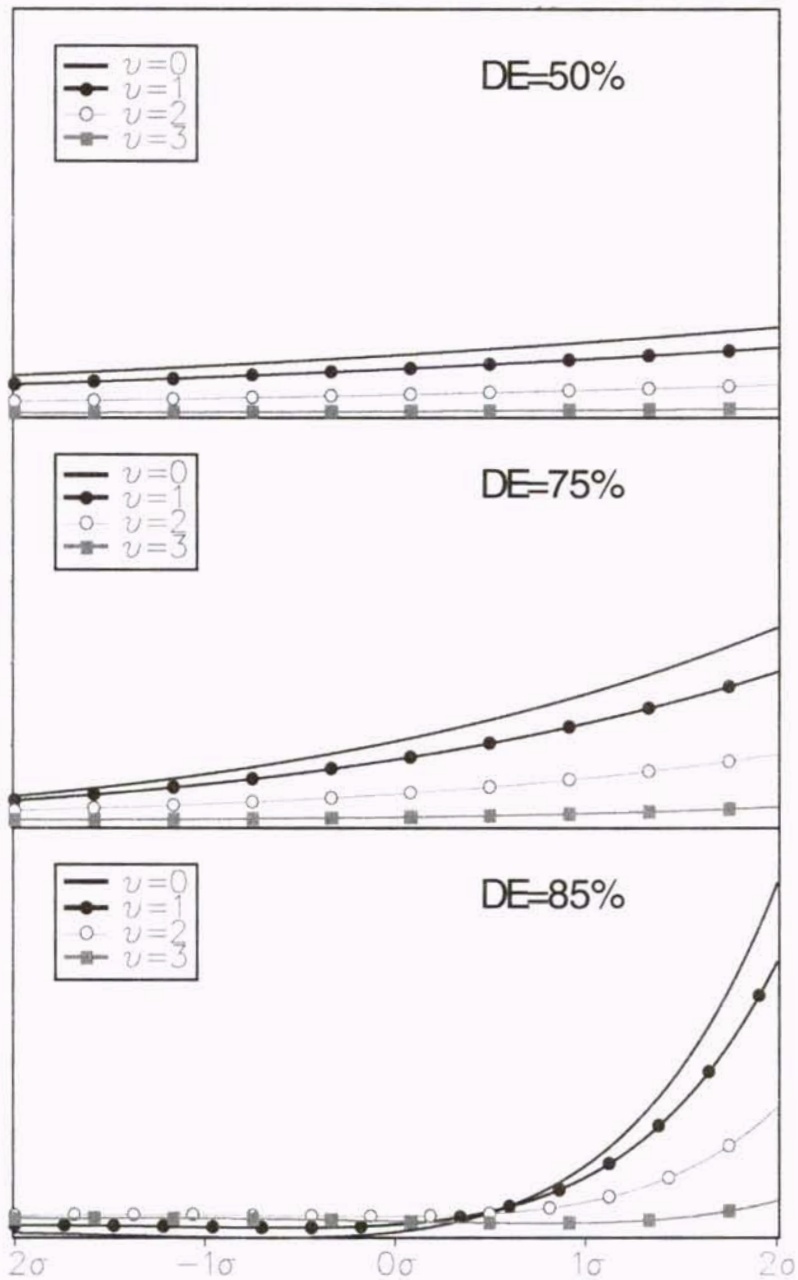
Figure 6.8. Diffracted beam profiles – truncated uniform subject beam

Again we see in plots 2 and 3 that increasing the total diffraction efficiency results

in additional beam degradation. At 75% diffraction efficiency those effects become

somewhat dramatic, with the right side of the beam approximately three times that of the left. Again this effect is lessened as the beam incidence deviates from Bragg, but this improvement in beam quality comes at the cost of lower efficiency, as well as with any other problems associated with operating off-Bragg.

Finally, these problems become extreme when the diffraction efficiency reaches 85%. By that time most of the power is contained in the first 25% of the original beam width. Again the beam drops to zero (at least on-Bragg), before coming back up in amplitude again. By the time the efficiency reaches this level, the reproduction fidelity is low enough to limit its use in any application for which this is an important issue.

6.3   Diffraction Efficiency Calculations

In each of these three cases the graphs were distinguished by their overall diffraction efficiency. Although the calculation of this efficiency is done in the same manner as in earlier sections (integrate the diffracted profile and divide by the known incident power), some insight is gained by considering this in detail. In the three-dimensional case Moharam et al [3] have shown these powers to be:

$$P_d = \eta_m^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_S^2(s,y) |S(s,\infty,y)|^2 \, ds \, dy$$
$$P_t = \eta_m^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_R^2(r,y) |S(\infty,r,y)|^2 \, dr \, dy$$

(6.32)

105

where η_m is the average characteristic impedance of the medium. The diffraction efficiency then becomes:

$$DE_{3D} = P_d / (P_d + P_t) \qquad (6.33)$$

When the problem is limited to the two-dimensional case (i.e. when we consider only the plane y=0) Solymar and Jordan [4], and Keenen [5] have shown that the expression for the diffraction efficiency can be reduced to:

$$DE_{2D} = 1 - J_0^2 \left[2\kappa' \left(u(\infty) \, v(\infty) \right)^{1/2} \right] - J_1^2 \left[2\kappa' \left(u(\infty) \, v(\infty) \right)^{1/2} \right] \qquad (6.34)$$

where u and v are as defined in section 6.1. In discussing this equation however, Moharam et al [3] pointed out that since $u(\infty)$ and $v(\infty)$ are proportional to the total powers of the recording reference and signal beams respectively, *the diffraction efficiency is independent of the beam profile and depends only on the product of the power in each of the two writing beams.* Therefore equation 6.34 represents a universal result for the diffraction efficiency of all two-dimensional crossed-beam gratings with a given recording power. On its own this is quite surprising, however this fact actually has important additional repercussions.

The first implication of equation 6.34 comes from the form of the expression itself. Although an initial look at the equation probably does not show this, a plot of this

equation versus $G \equiv 2\kappa'\left(u(\infty)v(\infty)\right)^{1/2}$ shows the function to be a non-decreasing function. This plot is given as Figure 6. 9 below.
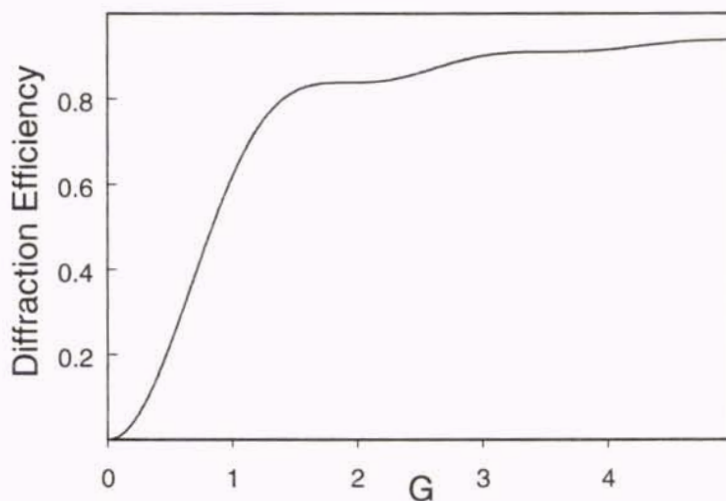
Figure 6.9.   Diffraction efficiency versus grating strength – overlap gratings

Since G is defined by the total amount of power in the recording beams, increasing the power in either of these beams will therefore increase the diffraction efficiency (or at least keep it constant). We now consider a grating recorded with an arbitrary reference beam and a uniform subject beam of width w. We can integrate over these profiles to find G, and then determine the diffraction efficiency from equation 6.34. If we now keep the reference beam the same, but increase the width of the subject beam, G will increase accordingly, as will the diffraction efficiency.   In this situation the increase in diffraction efficiency represents the amount of light diffracted into the additional width. From Figure 6.9 we see that after an initial ramp up period (g < 0.4) this increase will be approximately linear with G (i.e. with position). Within a certain

107

area however ($G \cong 1.8 \ldots 2.1$) widening the subject beam results in no additional diffraction efficiency, thus no light is diffracted into this area.

This is the origin of the beam break-up as the efficiency approached 85% in the previous plots (the value of G was between 1.8 and 2.1 in those areas). This effect can be seen much more clearly in Figure 6.10 below, which shows the derivative of equation 6.34 versus G. It therefore represents the additional light that is diffracted as the replay beam is widened.
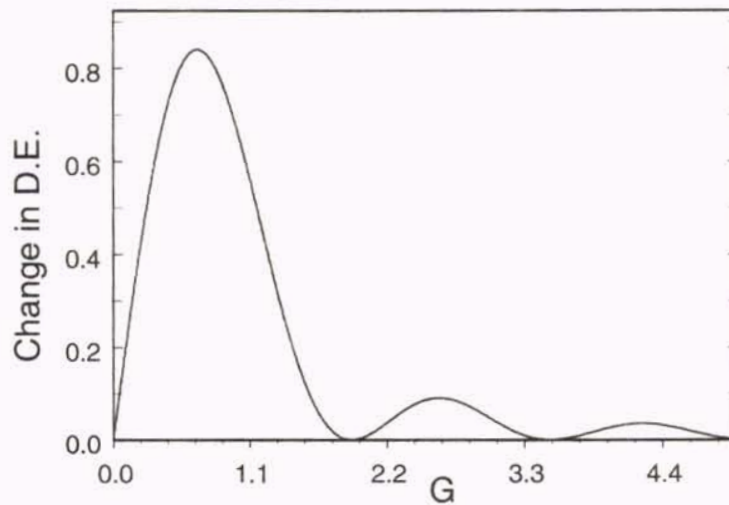
Figure 6.10. Change in diffraction efficiency with grating strength – overlap gratings

Once we have increased the length of our uniform subject beam such that G exceeds 2.1 we again get light diffracted. By looking back at Figure 6.9 we see that the diffraction efficiency where this starts to happen is about 83%. Clearly this effect (which is mathematically due to a zero in the J_1 Bessel function) is the cause of the break-up of the beams in all of the 85% graphs given earlier. Since there is no way around this limit (theoretically one would need an infinite amount of energy in an infinitely narrow area to

"jump" over this region, and even if this recording beam could be generated, the material would need to be linear across this infinite range), it can be said that 83% represents a fundamental limit to the diffraction efficiency with which an overlap hologram can give a faithful reproduction of the recording beams. Additionally from either Figure 6.9 or 6.10 we can see that a similar effect happens at $G \cong 3.5$ and $G \cong 5$. Continuing these plots out toward infinity would show an infinite number these discontinuities.

The fact that the relationship between the diffraction efficiency and the recording beam power is so simple has an additional benefit. Since we increase the incident beam power as we widen the recording beam we naturally get an increase in diffraction efficiency (up to the limit of 83%). By varying the amplitude while widening the subject beam it is possible to choose the shape of the diffracted beam. This effect may be used for a correction of some of the distortion discussed above (i.e. that which results from beam depletion during replay), or to perform somewhat arbitrary beam shaping. Each of these issues will be discussed in considerable detail in chapter 7.

6.4   Summary

The profiles of the diffracted and transmitted beams of overlap gratings were investigated using known solutions. When the grating is replayed with the exact recording reference beam, the resulting diffracted beams will be shifted and distorted versions of the recording subject beam. These effects arise because the first Born

approximation is violated (substantial beam depletion occurs). At overall efficiencies of up to 50% they are minimal, however they become pronounced at 75%. When the overall efficiency exceeds 83% the diffracted beam splits into more than one section. This distortion and shift can be problematic in data storage systems.

CHAPTER 7

DESIGN OF OVERLAP GRATINGS FOR BEAM

CORRECTION AND BEAM SHAPING

In chapter six, profiles were given for the beams diffracted by and transmitted through overlap holograms. In all cases the fidelity of the reconstructed beams was less than ideal, with the distortion increasing as the efficiency went up. At the end of the chapter a simple expression was given for the diffraction efficiency of these structures (when replayed on Bragg with the original reference beam). This expression indicated that the overall D.E. could be found simply from the total power in the recording beams. In this chapter we will attempt to use this relationship to perform beam shaping by varying the distribution of the recording waves. A general method will be given to find the required subject beam such that the wave diffracted by the resulting hologram will have some desired form. Once the required profile is found, it is obtained by placing an amplitude mask in front of a base subject beam (such as a Gaussian beam). After describing the method, an example of its use will be given. In this example a grating will be designed that diffracts a uniform wave from a Gaussian replay beam.

Several additional examples will then be given to show the improvement that can be obtained when the hologram is created with these scaled subject beams. The examples will be the same cases used in chapter 6 to describe the properties of overlap gratings (i.e. a Gaussian beam, an array of 5 Gaussian beams and a truncated uniform beam).

7.1   Basic method to obtain a desired beam profile

In this section a description of how to perform beam shaping with volume holography will be given. The geometry used throughout the discussion is shown below. The figure shows an overlap hologram being recorded by the interference of a Gaussian reference beam and an arbitrary subject beam (shown with a Gaussian amplitude).
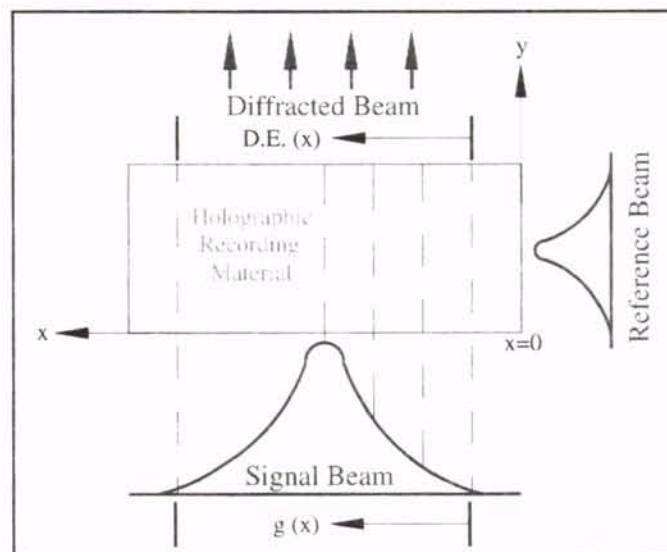
Figure 7.1. Geometry for beam shaping method

112

The basic idea is to determine the necessary grating strength profile (κ vs. position) that will cause the diffracted beam to have some desired form (e.g. uniform, Gaussian...). In general finding this profile is not a trivial task, however with some insight equation 6.34 can be used to determine it. As discussed in section 6.3 this equation gives the Bragg diffraction efficiency of an overlap grating as a function of the total recording power (regardless of the actual profiles). Naturally the value it gives is the overall efficiency.

Suppose, however that the recording signal beam is truncated at some point x. We can insert the power in this portion of the beam into equation 6.34 (along with that of the reference beam) to determine the diffraction efficiency of this modified hologram. By repeating this process for various values of x, the following function for the diffraction efficiency vs. position (i.e. D.E.(x)) can be obtained (see fig. 7.1):

$$DE(x) = 1 - J_0^2 \lfloor 2\kappa'(u(\infty)v(x))^{1/2} \rfloor - J_1^2 \lfloor 2\kappa'(u(\infty)v(x))^{1/2} \rfloor$$

(7.1)

where:

$$v(x) : \int_0^x A_s^2(\xi) d\xi$$

and $u(\infty)$ and κ' are as defined in chapter 6.

The amount of power diffracted into the output beam as we travel along the x-axis can then be determined from this expression for any given recording beam. This is essentially the diffracted amplitude profile. Obtaining these profiles in this manner is

113

straightforward enough, however it still must be performed numerically. Thus it holds little advantage over the use of equation 6.12.

This discussion assumed that we were given a recording signal beam, and from this the resulting diffracted profile was calculated. If however, it was desired that the diffracted beam have a particular profile, then we could use similar logic to calculate the D.E.(x) that would give this profile. These values could then be inserted into equation 7.1 to obtain the necessary form for $G(x) \equiv 2\kappa'(u(\infty)v(x))^{1/2}$.

Since we know the total power in the reference beam, all that is needed to find $v(x)$ from this equation is the value of κ'. This is obtained by choosing a value for the transmission of the amplitude mask's first section. It turns out however that this selection is arbitrary, as it merely represents a scaling factor for the entire mask (any changes resulting from a different value are absorbed in κ', which in turn is set by changing the recording time).

We now have values for $v(x)$, the derivative of which will yield the necessary recording profile. In most cases these last two steps will be done numerically, which will require the desired output beam to be split into a number of discrete sections, each of which is solved individually. The number of sections used will determine how well the profile matches its desired value, as well as how smooth the output is.

The entire process can be summarized into the following simple algorithm:

114

Algorithm for Performing Beam shaping

1. Select a value for the overall diffraction efficiency (De_{max})

2. Choose the desired diffracted beam profile.

3. Select the number of discrete sections (N).

   a. Using N and (2), calculate the desired $D.E._x$ (the desired diffraction efficiency at the end of each discrete section)

4. Using the equation: $DE = 1 - J_0(2\gamma) - J_1(2\gamma)$, solve for the γ_x's.

5. Calculate the power in each discrete section of the *base* subject beam. These will be designated as F_k.

6. Using the γ_k's from (4) and the F_k's from (5), calculate the amplitudes of the transmission mask (the A_k's) from the equation:

$$A_k = \sqrt{\left[2\left(\frac{\gamma_k}{c}\right)^2 - \sum_{G=1}^{k-1} F_G A_G\right]\frac{1}{F_k}}$$

7. Test the resulting profile to determine if it is smooth enough or whether the number of sections N must be increased.

Despite the number of individual steps, this process is actually quite simple. In order to show this, an example will be given. In this example the profile required to convert a Gaussian replay beam into a uniform diffracted beam will be found. For this example a base subject beam with a Gaussian profile will be used. Each step will be described below, as well as shown graphically in Figure 7.2 for additional clarity.

115

The first step in the process is to select an overall diffraction efficiency. For this example a value of 50% will be used. It is desired that the diffracted profile be uniform (step 2), and in step 3 the number of discrete sections is arbitrarily chosen to be 10. This allows us to calculate the desired cumulative diffraction efficiency of these 10 sections (i.e. DE_x). These values are given in Figure 7.2. Inserting them into equation 7.1 yields the cumulative grating strength of each section (i.e. $\gamma(x)$)(step 4).

Step 1:	Total Diffraction Efficiency: 50%									

Step 2: Desired Diffracted Profile: Uniform

Step 3:	N = 10	10	9	8	7	6	5	4	3	2	1
Step 3:	D.E.(x)	0.5	0.45	0.4	0.35	0.3	0.25	0.2	0.15	0.1	0.05
Step 4:	$\gamma(x)$	.840	.778	.718	.658	.598	.537	.473	.403	.325	.226

Step 5: Basic Recording Profile: Gaussian $(F_k\text{'s})$:

Step 6:	$(A_k\text{'s})$:	1.401	.401	.153	0.08	.057	.055	.072	.127	.313	1
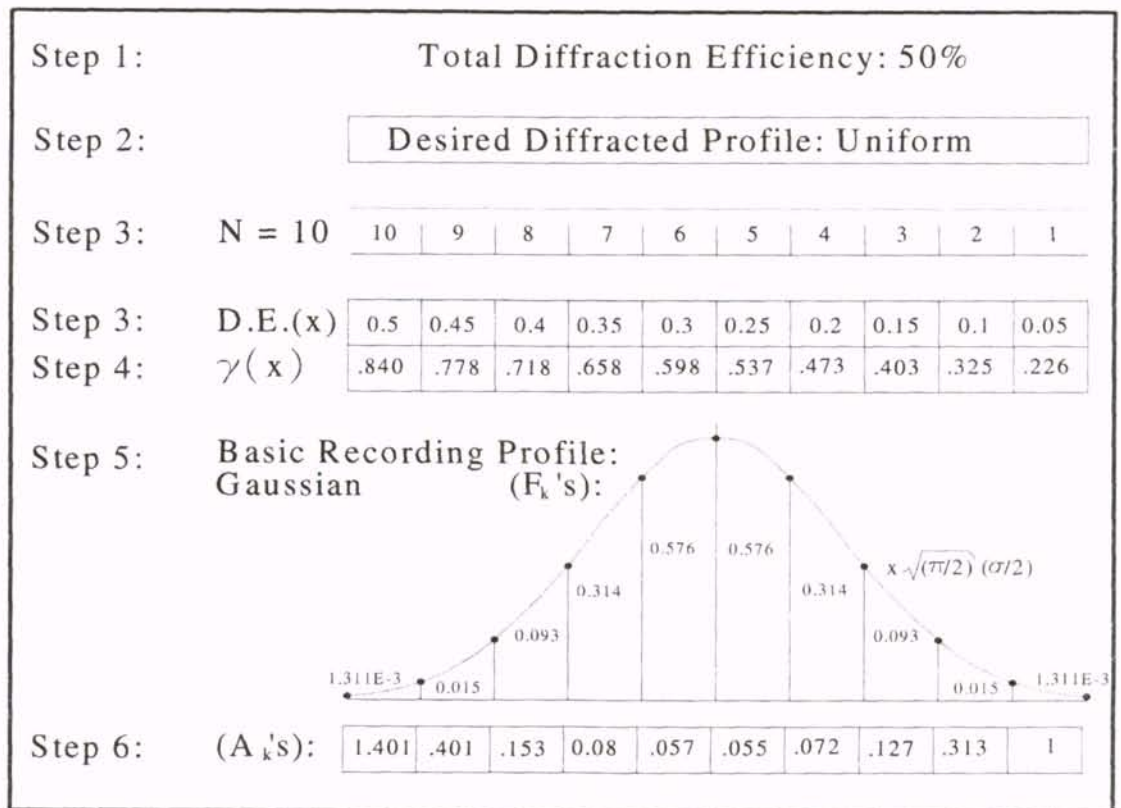
Figure 7.2. Graphical description of the beam shaping method

Step five calculates the power in each section of the base recording beam. For this example the beam is Gaussian and the corresponding powers are given in the figure.

116

The combination of steps 4 and 5 are then translated into scaling factors for the recording beam in step 6. This requires knowledge of the total power in the reference beam, which for a Gaussian beam is $A_{R0}^2 \sigma \sqrt{2\pi}$. The resulting scaling factors are shown in the figure. Two final steps are then performed. The first is to divide all of the A_k's by their largest value (to keep the transmission levels of the amplitude mask below unity), and the second is to apply a curve smoothing function to avoid discontinuities at the end of each section.

The design is now complete, and the resulting profile is inserted into equation 6.12 to find the output of the grating. The results are shown in Figure 7.3 below. Along with the output profile, the transmittance of the required amplitude mask is given.
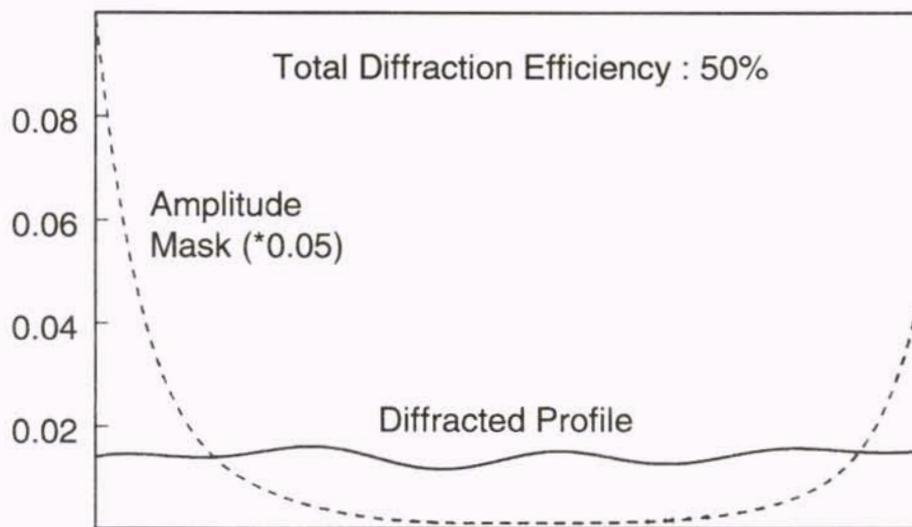
Figure 7.3. Results for a Gaussian to uniform beam converter

We see in Figure 7.3 that an approximately uniform diffracted beam has been generated. In this example there is an amplitude variation of about 15%. If this is too

high for a given application then it can be lowered by increasing the number of discrete sections used. The method therefore appears to work well for overall beam shaping.

When looking at the transmission characteristics of the amplitude mask, however a serious problem becomes evident. The transmittance must vary up and down between 0.04 and 1.0 across a relatively small area, which is difficult to fabricate. Additionally this analysis assumed that the phase difference in propagating through the mask would be constant for all sections, which is probably not the case with this large a difference in the section transmittances. The complexity of the mask in this example was a result of the large difference between the base subject beam (Gaussian) and the desired output beam (uniform). This was done to show the overall generality of the method. In most applications, however the base beam will be similar to the desired output, and better results will be obtained with simpler masks. The remainder of this chapter will give examples of the improvement that can be obtained by using the method.

7.2   Using the method to correct for distortion in the diffracted profiles

In order to present results of using a modified subject beam we again must select specific cases. The three cases to be shown are those of chapter 6 (i.e. a single Gaussian, multiple Gaussians and finite rectangular subject beams). In each case the diffracted profile will be given for an unapodized grating (from chapter 6) as well as for an apodized grating (using the method). In addition, a plot of the transmission characteristics of the required amplitude mask will be included.

118

Case 1: Gaussian Subject Beam

As in chapter 6, the first case to be considered will be the Gaussian subject beam. Again overall diffraction efficiencies of 50%, 75% and 85% will be covered, although only the Bragg diffracted profiles will be given. The results when the overall diffraction efficiency is 50% are shown below in Figure 7.4.
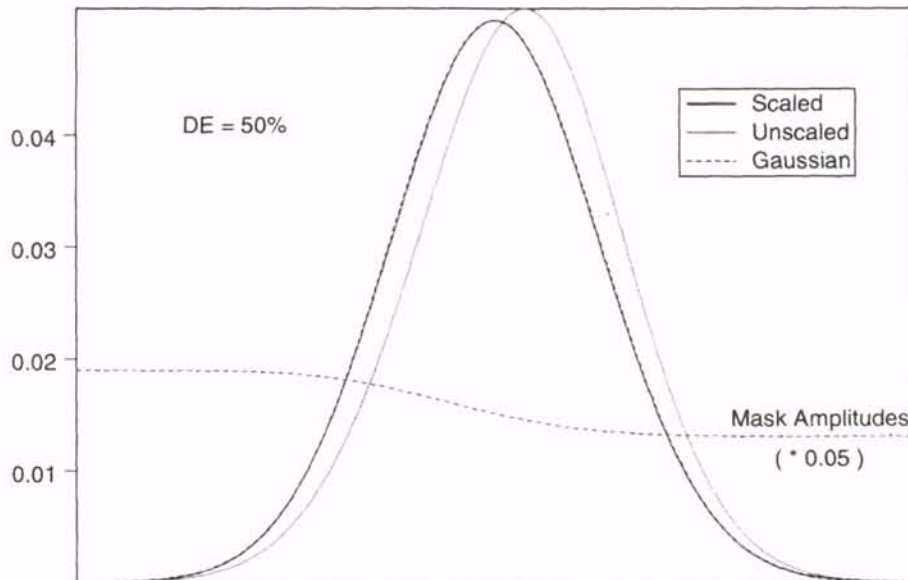
Figure 7.4. Scaled and unscaled diffracted profiles (overall D.E. = 50%)

This figure clearly shows that for an overall diffraction efficiency of 50% using a modified subject beam can give a diffracted profile that is almost exactly Gaussian. No longer is the output shifted, and the asymmetry has been removed. This is a considerable improvement over the unscaled case of chapter 6 (shown in the figure). Additionally, the

required amplitude mask has a relatively simple transmission profile which makes fabrication straightforward. As will be shown below, this may not always be the case.

When the overall diffraction efficiency is increased to 75%, as in Figure 7.5 below, we again get a diffracted profile which is nearly Gaussian. In generating this figure 20 discrete sections were used, which left some difference between the apodized plot and the Gaussian curve, however if necessary the agreement could be improved by using more sections. Again the results show considerable improvement over the unscaled curve obtained from chapter 6.
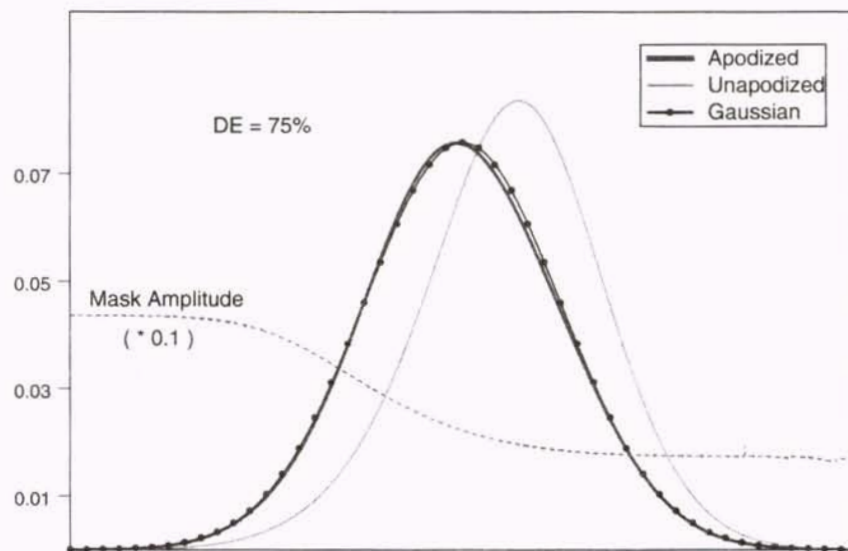
Figure 7.5. Scaled and unscaled diffracted profiles (overall D.E. = 75%)

In looking at the amplitude mask we see that a larger difference in transmission levels is required to correct the distortion at this efficiency, however the change in

transmission is gradual and it should not be overly difficult to make the mask. Thus the method seems practical for efficiencies up to 75%.

The final example of a Gaussian subject beam is for an overall diffraction efficiency of 85%. This is shown below, and it contains a few interesting points.
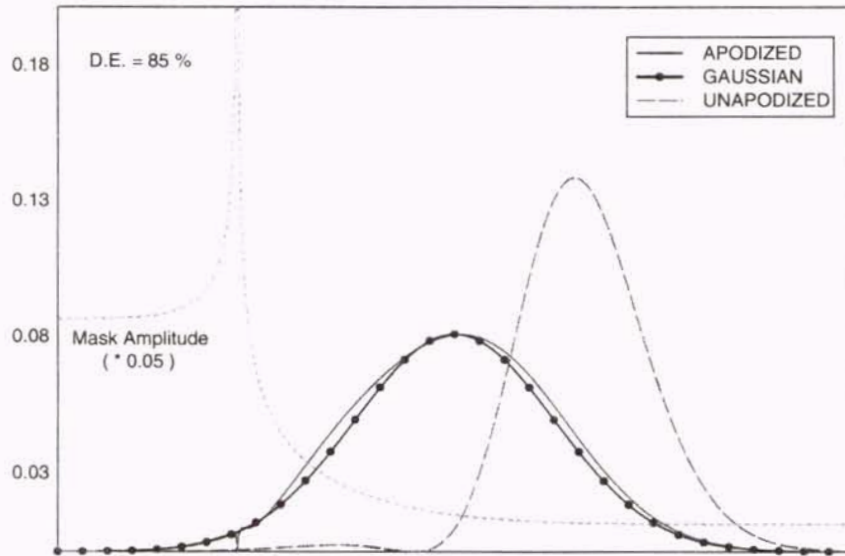
Figure 7.6. Scaled and unscaled diffracted profiles (overall D.E. = 85%)

The first thing indicated by this figure is that although the corrected profile matches the Gaussian curve better than that of the unapodized grating, the match is not as good as it was for lower efficiencies. We see therefore that as the diffraction efficiency increases, correcting for the corresponding distortion becomes more difficult (it would require a large number of discrete sections and a better curve fitting algorithm). Even if more sections were added to the analysis however, an exact Gaussian match could not be obtained since the diffracted profile must still drop to zero at 83%. This is the second

important issue illustrated by Figure 7.6, and it reinforces the idea that this is a fundamental limit to the diffraction efficiency with which an arbitrary profile can be obtained.

In looking at the apodization function of Figure 7.6 we see immediate problems. The transmittance varies by more than a factor of 100. Additionally, this variation occurs over a range of about $\sigma / 10$, which makes this mask all but impossible to fabricate.

Case 2:   Array of Five Gaussian Subject Beams

The second geometry considered is an array of five Gaussian beams. Again, three cases will be covered (diffraction efficiencies of 50%, 75% and 85%), the results of which are shown below in Figures 7.7, 7.8 and 7.9. In all of these cases a very simple apodization function was used – five uniform transmission layers (one for each of the diffracted beams). The transmission level for each of the sections are shown in the figures by their respective beams.

Total Diffraction Efficiency: 50% (10% / Beam)
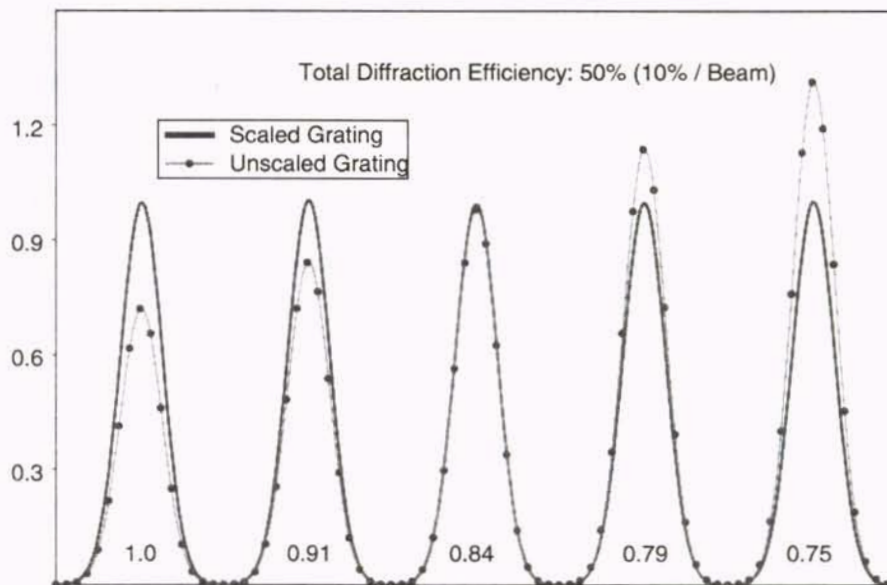
Figure 7.7. Scaled and unscaled profiles – 5 Gaussian Beams (overall D.E. = 50%)

In this first figure the diffraction efficiency is 50%. It is clear from the figure that at this diffraction efficiency the overall method can be used to make the outputs of all beams approximately uniform. This addresses one of the data storage issues described earlier (the threshold problem). Additionally, the beams are not shifted (the other concern), however at this efficiency the shift was also minimal in the unapodized plot. Since the apodization function is just five uniform levels fabrication of the mask is quite simple.

When the diffraction efficiency is increased to 75% we get the results shown below. Again we see that the beam diffracted from the apodized grating closely matches the desired profile (five uniform Gaussian beams). The separation of the maxima of these beams is relatively constant, which implies that the shifting of the beams has also

been corrected. As in all of these cases, the simplicity of the apodization function makes the mask quite easy to manufacture, increasing the practicality of its use.
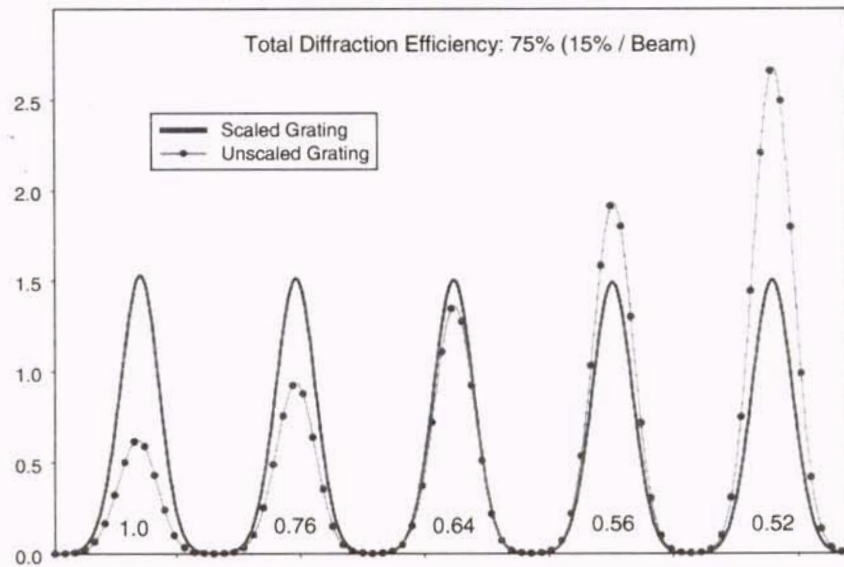
Figure 7.8. Scaled and unscaled profiles – 5 Gaussian Beams (overall D.E. = 75%)

The final case has a diffraction efficiency of 85%. The diffracted profile of this grating is shown below in Figure 7.9. With this efficiency the first four beams have the desired profile, however the fifth beam is considerably distorted and again drops to zero at 83%. We do see significant improvement over the un-scaled grating (where the third through fifth beams were highly distorted and the fourth beam was virtually non-existent), which implies that the method is still worthwhile. Although the profile of the fifth beam could be improved by using a more complex amplitude mask (one where the transmittance varies across the final section), the zero drop would still not be removed, so it is questionable whether it would be worth the additional complexity.
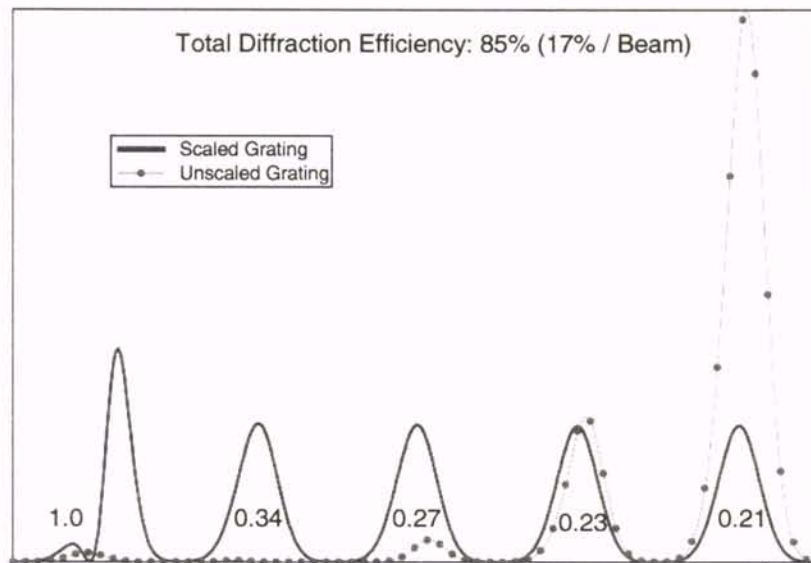
Figure 7.9. Scaled and unscaled profiles – 5 Gaussian Beams (overall D.E. = 85%)

Case 3:   Truncated Uniform Subject Beam

The final case to be covered is the uniform subject beam. Figure 7.10 shows the resulting profiles (both apodized and unapodized) for all three diffraction efficiencies. Due to the simplicity of the profile all efficiencies have been included in one plot.

As was the situation in the single Gaussian and multi-Gaussian cases we see that at 50% diffraction efficiency we were able to obtain the desired profile (a diffracted beam with uniform amplitude) quite nicely, with a relatively simple amplitude mask. In a similar manner when the diffraction efficiency was raised to 75% the profile still remained substantially uniform, again with a reasonable amplitude mask.

125
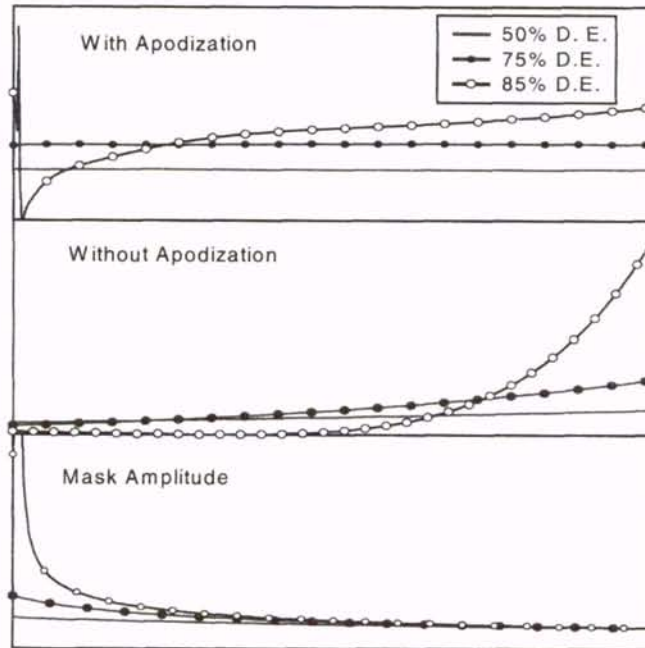
Figure 7.10.   Beam profiles for a uniform desired beam (various D.E.'s)

When the efficiency was increased beyond 85% however, the profile could no longer be made uniform, even with the complex amplitude mask shown in the figure. As in the other cases, the amplitude dropped to zero at an overall diffraction efficiency of 83%. Because the amplitude masks in the 50% and 75% cases were so simple (the transmission varied by no more than 13%), it seems that this method is practical for correcting these profiles.

Based on all of these examples we can state that the distortion that is naturally present in an overlap hologram can be substantially reduced by applying an amplitude mask to one of the recording beams (at least in the linear region of Figure 6.10). Additionally, somewhat general beam shaping can be performed.   Along with the

examples presented here, this was also shown in [1] where a similar method was used to alter the diffracted beam profile.

Although the plots with multiple Gaussian beams indicate that two of the issues in holographic data storage can be addressed, there are problems associated with correcting them in this manner. This is because the required amplitude mask depends on the number of illuminated bits. Thus the ideal mask would change for each bit pattern stored. It may be possible however, to choose a profile that is based on some average number of bits being illuminated (such as assuming that every other bit is illuminated). In this way the effects would be reduced (although not eliminated). It may be that the best way to correct these problems is that which was suggested in chapter six - operating at a low enough efficiency where the first Born approximation holds.

Finally, the ability to perform somewhat arbitrary beam shaping brings up the possibility of an interesting device – a variable beam shaper. The idea is to angularly multiplex several holograms with different diffracted profiles in one area. Then by rotating the stored hologram a person could choose any one of a number of profiles for use in an optical system or projection applications (such as shaped laser pointers). This would avoid the necessity of entirely replacing the shaping element when a new profile is desired.

7.3   Summary

A method was developed for obtaining a diffracted beam from an overlap hologram with any desired profile. This method is based on using an amplitude mask

127

over the subject beam to vary the grating strength with position. The required amplitude masks are simple (realizable) when the desired profile is close to that of the recording subject beam, thus this should be used when possible. It was shown that the method can be used to correct for most of the beam depletion effects discussed previously, however the splitting of the beam at 83% can still not be alleviated, thus it represents the true limit at which a good reproduction of the recording beam can be obtained.

CHAPTER 8

MULTILINE FILTERS

Throughout this work the filtering properties of various hologram types have been discussed. In all cases the thickness of the medium resulted in a filter with an extremely narrow linewidth. This narrowness can be exploited to record multiple holograms within the same volume. Although the effects of this multiplexing could have been analyzed as each type of hologram was presented, it seems more appropriate to separate the topic into a single chapter on multi-line filters. Thus, all filters considered here will have multiple resonances (either in angle or wavelength). Analysis will begin with multiplexed transmission gratings, continue with reflection gratings, and end with a discussion of novel thin film alternatives which can give similar results.

8.1   Angularly Multiplexed Transmission Gratings

Table 2.1 listed the width of the angular response of a transmission grating as a function of its thickness. For our materials this width is on the order of

0.05°, which is narrow enough for these devices to be used in such applications as high density data storage systems and multi-line filtering systems with close angular spacing.

In any of these applications the separation necessary for the individual holograms to be uncoupled is an important point. It is determined from the properties of the individual gratings, along with those of the replay beam, and the particular application at hand (which determines how "independent" the gratings must be). Although this decoupling can easily be obtained by putting "extra" angular spacing into the separation (to ensure minimal cross coupling), this is not a desirable solution as it wastes precious bandwidth and reduces the number of gratings which can be multiplexed. Determining the required separation is clearly an electromagnetics problem, and an investigation of it will be given shortly.

Along with finding the necessary angular separation between holograms there is an additional concern which must be addressed when multiplexing gratings in many real materials - holograms that are recorded later tend to degrade those which were recorded previously. This effect is obviously material specific, however it has been well studied in some of the more common materials [1-4]. Results from these works will be discussed presently, along with a brief discussion of the effect in PTR glass.

A number of approaches have been used to investigate cross coupling in multiplexed gratings, three of which will be discussed here. The first, which was devised by Tu et. al. [5], uses a multiple scattering approach to predict the outputs. This method can handle any number of gratings (although only the dual grating case was discussed in the paper), along with multiple coupling between orders. It appears to be the most

powerful of the methods described here, however it has the disadvantage of being a purely numerical approach which makes obtaining physical insight difficult.

A simpler method was used by Case [6] who (following a procedure similar to Kogelnik) derived a set of coupled wave equations for two superposed gratings which shared a common reference angle. By restricting his derivation to exact Bragg incidence he was able to obtain closed for solutions, although only for the overall diffraction efficiency of each beam.

Finally, Alferness and Case [7] gave a method which could handle any number of gratings (although again limiting the discussion to the dual grating case) which was based on thin grating decomposition. Their paper presented several cases of grating strengths and angular separations, and showed that considerable cross-coupling effects can exist. From the results obtained they were able to determine that a reasonable estimate of the Bragg-angle difference necessary to effectively decouple the gratings was 1.5 times the width between the maximum and the first zero of a single grating. Their solution was obtained by slicing the grating into a number of thin sections, and using scalar diffraction theory to determine the transmission of each one. The output of each slice was then used as the input to the next until the grating was traversed. Although it is questionable whether scalar diffraction theory is valid for these gratings, considerable insight is nonetheless gained from the analysis.

In this section we present the derivation of coupled wave equations for two or more multiplexed gratings. The resulting expressions will be valid for on or off-Bragg incidence. The basic procedure is to insert the permitivity variation corresponding to a pair of superposed sinusoids into the wave equation and to solve it under the assumption

of four propagating waves (a reference beam and three signal beams), while allowing any of these waves to propagate off-Bragg. This will result in a set of four coupled equations from which an overall solution can be found.

The permitivity variation which results when a pair of sinusoids are recorded in a linear non-absorbing material can be written as:

$$\varepsilon_0 + \varepsilon_1 \cos(\bar{K}_1 \cdot \bar{r}) + \varepsilon_2 \cos(\bar{K}_2 \cdot \bar{r}) \tag{8.1}$$

Plugging this into the wave equation and assuming that the replay field is given by:

$$\bar{E} = R(z)e^{-j(\bar{\rho}\cdot\bar{r})} + S_1(z)e^{-j(\bar{\sigma}_1\cdot\bar{r})} + S_2(z)e^{-j(\bar{\sigma}_2\cdot\bar{r})} + S_3(z)e^{-j(\bar{\sigma}_3\cdot\bar{r})} \tag{8.2}$$

results in the following set of coupled wave equations.

$$\frac{\partial R}{\partial z} + j\frac{\kappa_1}{c_R}S_1 + j\frac{\kappa_2}{c_R}S_2 = 0$$

$$\frac{\partial S_1}{\partial z} + j\frac{\vartheta_1}{c_{S1}}S_1 + j\frac{\kappa_1}{c_{S1}}R + j\frac{\kappa_2}{c_{S1}}S_3 = 0$$

$$\frac{\partial S_2}{\partial z} + j\frac{\vartheta_2}{c_{S2}}S_2 + j\frac{\kappa_2}{c_{S2}}R + j\frac{\kappa_1}{c_{S2}}S_3 = 0 \tag{8.3}$$

$$\frac{\partial S_3}{\partial z} + j\frac{\vartheta_3}{c_{S3}}S_3 + j\frac{\kappa_2}{c_{S3}}S_1 + j\frac{\kappa_1}{c_{S3}}S_2 = 0$$

where :

$$\kappa_1 = \frac{\pi n_1}{\lambda}, \quad \kappa_1 = \frac{\pi n_2}{\lambda}, \quad \vartheta_k \equiv \frac{(\beta^2 - |\sigma_k|^2)}{2\beta}$$

These equations are quite complicated, however they are also quite general. In the special case where the recording beams share a common reference angle and incidence is strictly on-Bragg, they reduce exactly to those of [6]. Additional gratings can be included in the analysis, which will add two equations to expression 8.1 for each new grating. Because of the complexity of these equations, closed form solutions are not available. Since we are dealing with transmission gratings however, the boundary conditions are such that a solution can be obtained using a simple Runge-Kutta numerical integration. Although this is less desirable than an explicit solution, it does make incorporating such things as finite beams (through the angular spectrum of plane waves) relatively straightforward.

In order to present results, specific examples must be selected. For comparison purposes, individual gratings similar to those used by Alferness and Case in [6] and [7] will be used. The analysis will include replay with plane waves as well as Gaussian beams. The plane wave results that will be given are significantly different from those of [7]. The differences arise because the authors chose an unusual method of calculating the angular separation between the gratings, which is atypical of traditional multiplexing systems. The result is that the actual separation of the gratings in their system is much higher than they listed. Because of this, their analysis does not apply to typical multiplexing schemes. In the following analysis, the results of multiplexing (and measuring the angular separation) in the traditional manner are presented.

Figure 8.1 below gives the resulting angular selectivities when two gratings are multiplexed with a separation of 5°. Results are shown for replay with a plane wave, as well as with two different Gaussian beams. The index modulation of each grating is

chosen such that $\kappa=\pi/2$, thus the plane wave diffraction efficiency of each grating in the absence of the other would be 100%. Again, these gratings are approximately the same as those used by Alferness and Case for comparison purposes, thus they have a HWHM slightly below 1.15°.
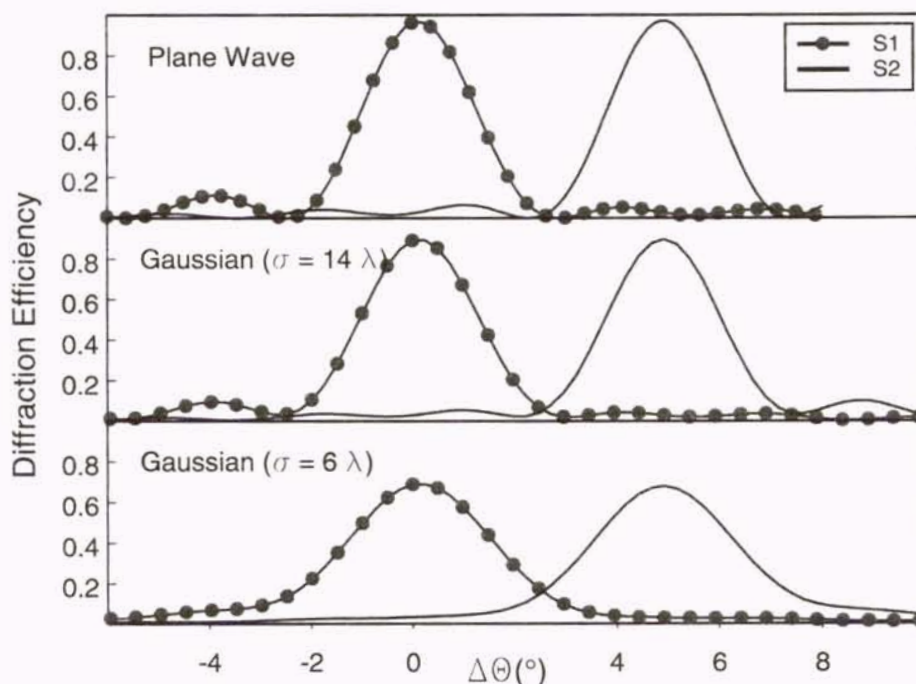
Figure 8.1. Angular selectivity of two multiplexed gratings - $\Delta\Phi=5°$ for various inputs

These plots show several important effects. The first is that the maximum diffraction efficiency of both beams has dropped from what they would be if only a single grating were present. Secondly, the location of each of these maxima has shifted slightly outward from their recording locations of 0° and 5°. Additionally, the responses are no longer symmetrical around their maximum values. All of these effects are present regardless of the type of incident beam, however they become more pronounced when the input beam is finite. Part (c) of this figure shows another important effect. At no point

134

does the diffraction efficiency of both beams drop to zero. If it did then a measurement of the sum of the diffracted beams could easily distinguish the presence of two gratings. Although measuring this sum in Figure 8.1c would show two maxima, the high diffraction efficiency between the peaks (>40%) would make distinguishing the individual gratings difficult. A final effect (which is barely visible in Figure 8.1) is that the maximum diffraction efficiency of each of the beams has become unequal.

All of these effects are even more pronounced in Figure 8.2 where the separation has been narrowed to 3°. In this case measuring the total diffracted power with either Gaussian beam would show efficiency plots with a single peak. Clearly when the replay beam is Gaussian (or at all finite) the effects of the inter-grating coupling are worse, and a wider separation than that suggested by Alferness and Case should be used.
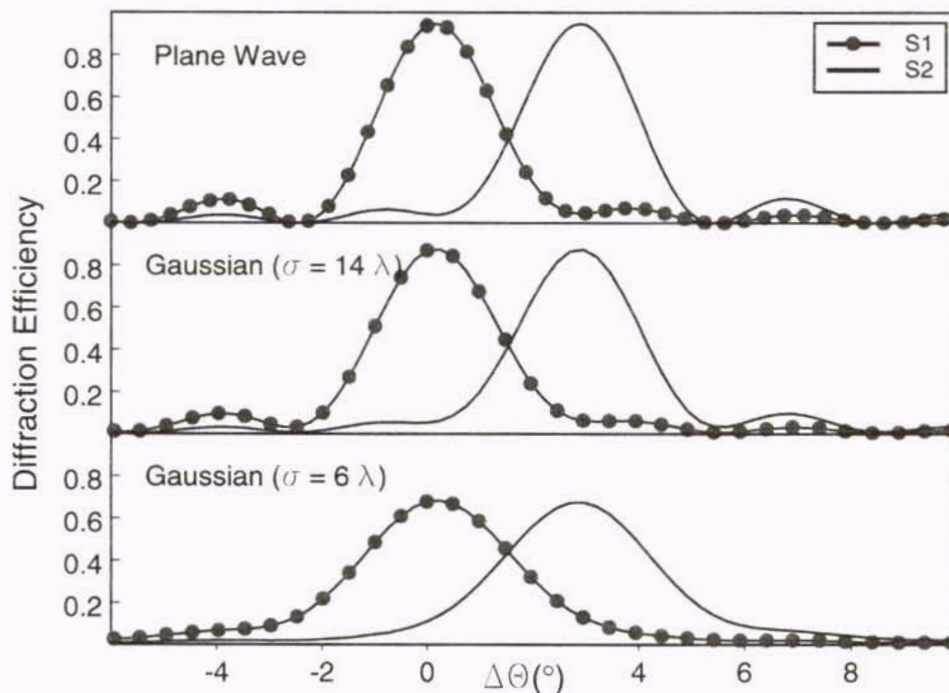
Figure 8.2. Angular selectivity of two multiplexed gratings - $\Delta\Phi=3°$ for various inputs

The Gaussian beams used for these analyses were extremely narrow. This was necessary for any beam effects to be apparent due to the thin material used (the HWHM of these filters was about 1.15°). The gratings considered in the present work are much thicker, so we can expect to see similar effects even with relatively wide beams. Figure 8.3 shows this by giving the replay characteristics of a pair of gratings (with separations of 0.1° and 0.05°) recorded in lithium niobate and replayed with a plane wave and with a typical He-Ne laser. The FWHM each grating was about 0.025°. In this case the effects seen above are present even with the wide beams. For these parameters a separation of about 4 times the FWHM is adequate for independent replay with Gaussian beams.
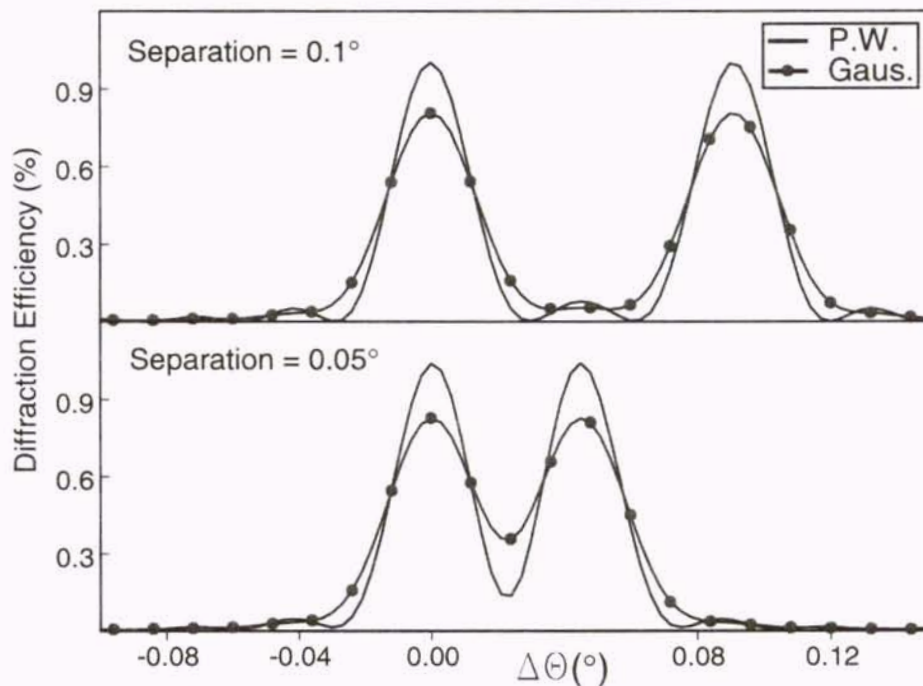
Figure 8.3. Angular selectivity of multiplexed gratings - P.W. & typical gaussian inputs

As in previous chapters an experimental verification of these effects was performed. In this case five gratings were recorded in lithium niobate with angular separations of 0.3° and 0.4° (in air). Replay was made with a helium neon laser. The results are shown below in Figure 8.4, which clearly indicates the presence of cross coupling effects. For the particular grating and beams used in the experiment a separation of 0.4 degrees seems adequate for independence. The D.C. background present in these plots is most likely due to scattering within the crystal.
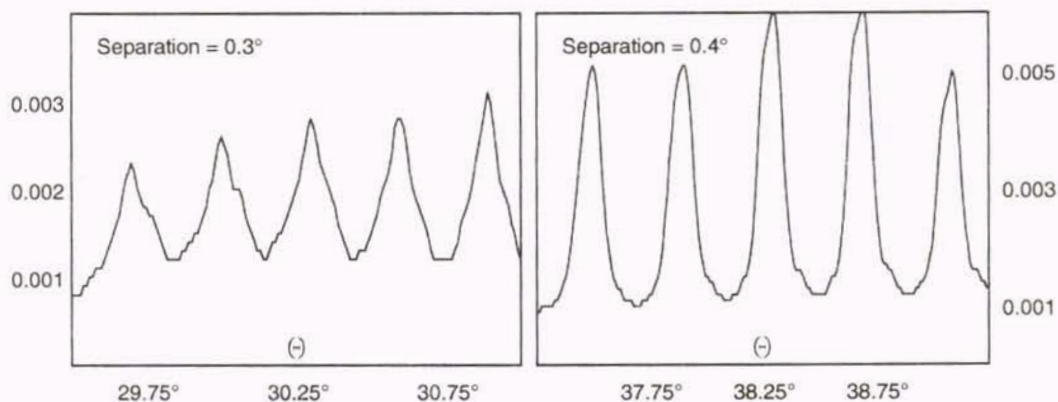
Figure 8.4. Experimental verification of multiplexing in LiNbO$_3$ - $\Delta\Phi$=0.3°, 0.4°

All of the previous analyses were concerned with finding the required separation between multiplexed gratings, however as discussed earlier another issue exists when multiplexing – the gratings recorded later tend to degrade the ones recorded earlier. This effect can be seen in Figure 8.5 below, which shows the angular selectivity of 10 gratings multiplexed in lithium niobate. The gratings were recorded in order from left to right, and the exposure times and recording intensities were kept constant throughout.

Obviously the latter gratings have a much higher diffraction diffraction efficiencies than those recorded earlier, indicating that overwriting has taken place.
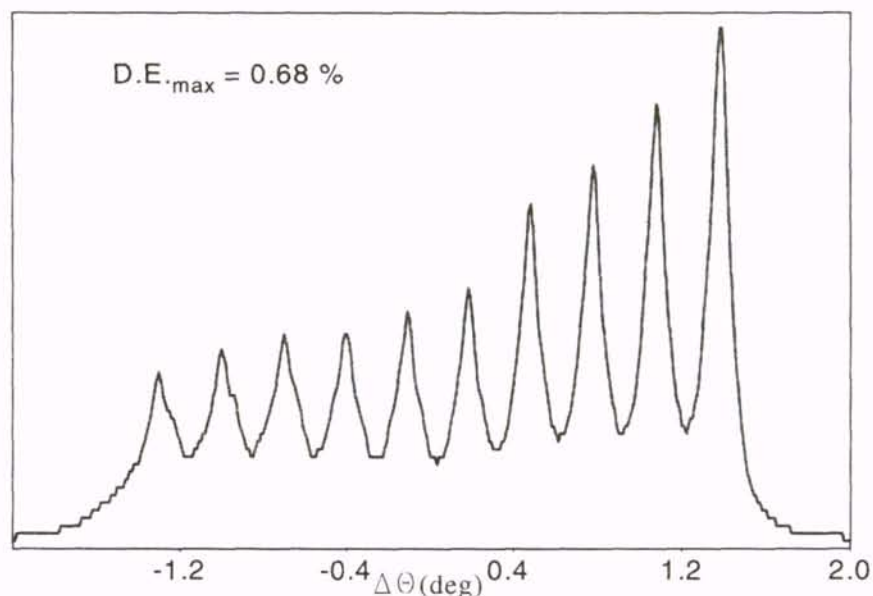
Figure 8.5. Effect of grating overwriting when multiplexing in $LiNbO_3$ - experimental

In this plot we see that the gratings with higher diffraction efficiency have a correspondingly higher signal to noise ratio (SNR). For this reason it is important to equalize the individual grating's efficiencies. The overwriting that is seen in Figure 8.5 is due to the writing process in lithium niobate, where the motion and trapping of charge carriers sets up a space charge field along with a corresponding modulation of the refractive index. In these materials later recordings re-orient the distribution of charges, thus changing the induced field and adding an additional index modulation – at the expense of previous recordings.

Since this process is reasonably well understood, scheduling methods have been developed which vary the exposure time or intensity, leading to much better uniformity in

multiplexed gratings [1]. Using these methods over 10,000 holograms have been recorded in lithium niobate with notable uniformity across their efficiencies [8]. Nevertheless the requirement for this scheduling, along with other complications in LiNbO$_3$ add significant complications to the multiplexing process.

Because a different mechanism is used to obtain an index modulation in PTR glass, it is desirable to see if similar effects occur when multiplexing in it. To investigate this, several sequential exposures were made in a sample of PTR glass, after which all exposures were developed (heated) concurrently. The recording intensity and time was kept constant for all exposures. Figure 8.6 shows an example where two gratings were multiplexed and heated several times. In this experiment a large separation angle was used in order to remove any effects due to cross coupling.
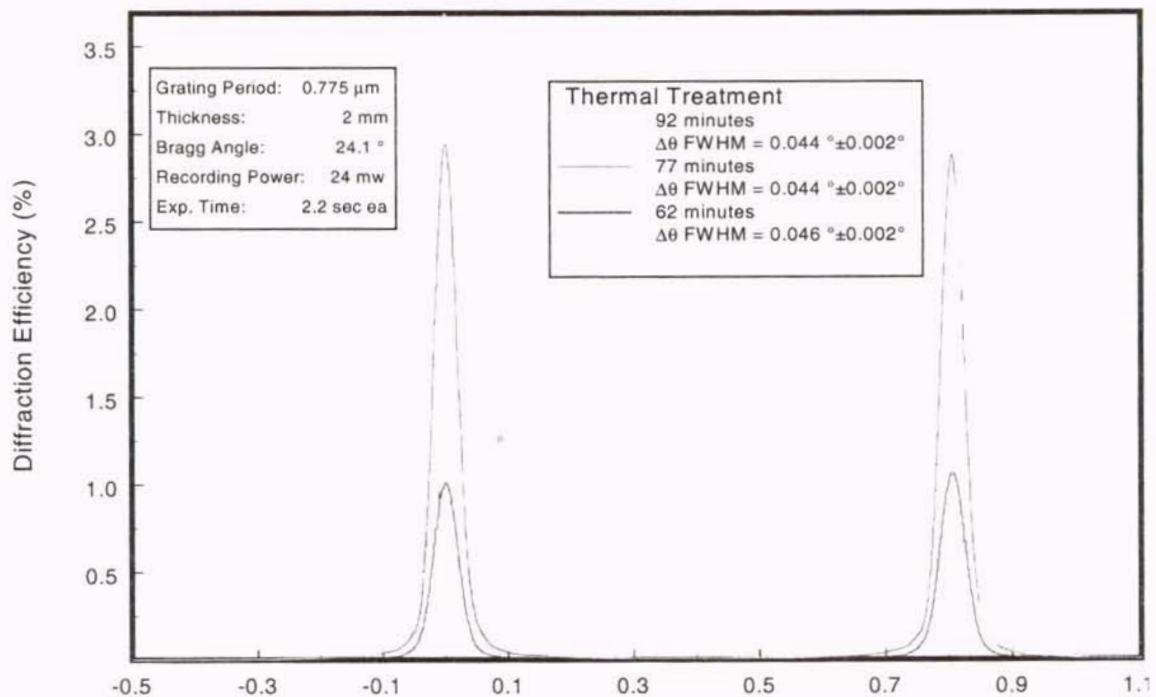
Figure 8.6. Multiplexing two gratings in PTR glass – experimental

As can immediately be seen, each of the gratings has approximately the same line shape and maximum diffraction efficiency, even across several heatings. In addition scattering is seen to be negligible. This is quite desirable as it implies no special compensation or scheduling is required when using this material.

The number of multiplexed gratings was then increased to fifteen, and the angular separation was reduced significantly. The results are shown in Figure 8.7 below. Again it appears that the diffraction efficiencies are relatively constant (only one of the fifteen gratings differs in diffraction efficiency by more than 25%), and cross talk remains minimal. Thus it appears that minimal degradation of earlier gratings is occurring during mutltiplexing – at least at this efficiency. This is in agreement with the understanding of the physical processes at hand, however it is anticipated that degradation would occur as the efficiency was increased to near the limit of the dynamic range.
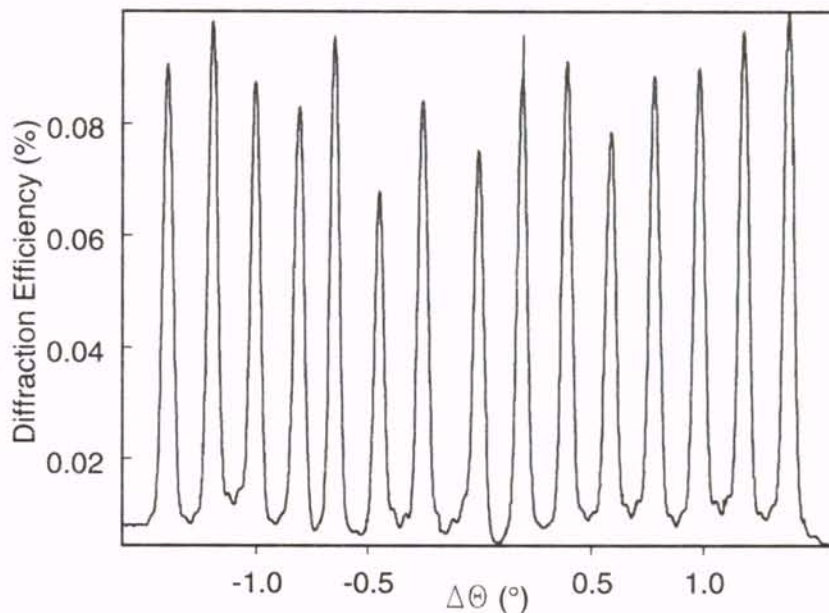
Figure 8.7. Multiplexing fifteen gratings in PTR glass – experimental

140

8.2   Angularly Multiplexed Reflection Gratings

Just as in the case of transmission gratings, additional multiplexing effects can be investigated when the gratings are of the reflection type. Performing this analysis in the manner used for transmission gratings would be quite difficult, since the boundary conditions of the differential equations exist at two different points, which is not straightforward to solve even with numerical methods. It is possible to avoid this complexity by limiting the analysis to cases where all of the gratings are unslanted, for which simpler solution methods exist.

When all of the gratings are of the unslanted reflection type the chain-matrix method described by Moharam and Gaylord [9] can be used to find the solution. This method slices the resulting index profile into a number of uniform sections and calculates the diffracted and transmitted fields as the beam propagates through the individual sections. Unlike when Alferness and Case used a similar concept for solving transmission gratings, the slices in this analysis are uniform, therefore the accuracy can be made arbitrarily high by increasing the number of slices. This section will present some of the results obtained using this method. Since these gratings are typically used as wavelength filters all plots will be in terms of wavelength deviations.

Figure 8.8 shows the wavelength response of two multiplexed gratings illuminated by a uniform plane wave. The grating wavelength separation is varied between 1 nm and 0.2 nm, the index modulation is kept constant at 0.000295, and the base index was chosen to be 1.5. These values were selected because they are typical for gratings recorded in either PTR glass or in doped fibers (for use in fiber Bragg

141

gratings)[10]. In this example the center wavelength was 1 μm, and normal incidence was assumed which gave a period of about 0.3 μm. Ten thousand periods were used, thus these normalized lengths represent actual lengths of approximately 3 mm. Again this is well within the reaches of either PTR glass or fiber Bragg gratings.
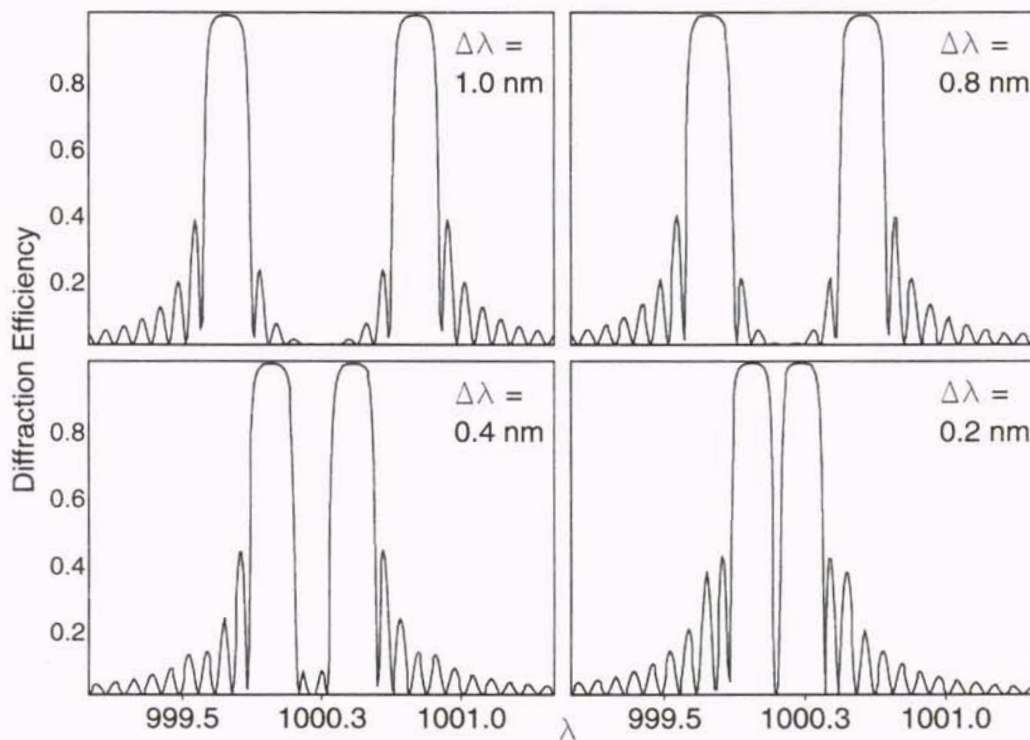
Figure 8.8. Wavelength response of multiplexed reflection gratings - various Δλs

The first thing we can note from these plots is that effects due to the presence of more than one grating definitely exist. This is obvious from the response between the two peaks. If no cross coupling or interference effects were present each individual plot would be symmetric and the intensity between the peaks would be the sum of the individual responses. Clearly, however they are not. The intensity in this area is

considerably lower than what would be present with only a single grating. This effect could result from either cross coupling between the gratings or from the individual responses from the grating interfering destructively. In fact it is most likely due to a combination of these. Similar effects are seen outside of the peaks, but the level is considerably less since the intensity of one of the responses is much lower than the other.

An interesting (and possibly useful) occurrence is seen in the fourth part of the figure (where the separation is 0.2 nm). Between the two reflection peaks is a narrow transmission peak. It is possible that this could be used as a wavelength transmission filter (recall that typical wavelength filters are used in the reflection geometry). The narrowness of this linewidth (on the order of angstroms) as well as its symmetry are both desirable properties. Additionally, the fact that this operates as a transmission filter brings up the possibility of using this filter in non-traditional areas of a WDM communications system.

Another advantage of this filter is its simplicity, which allows it to be scaled for generating multi-line narrowband wavelength transmission filters. This is shown below in Figure 8.9 where one, two, three and four gratings have been multiplexed to make filters with zero, one, two and three transmission peaks respectively. Each of the gratings in all of the figures was 10,000 periods long and had an index modulation of 0.000295. The period of the first grating was $L_1=0.333\mu m$, while those of the other gratings were: $1.0002\ L_1$, $1.0004\ L_1$ and $1.0006\ L_1$ respectively.
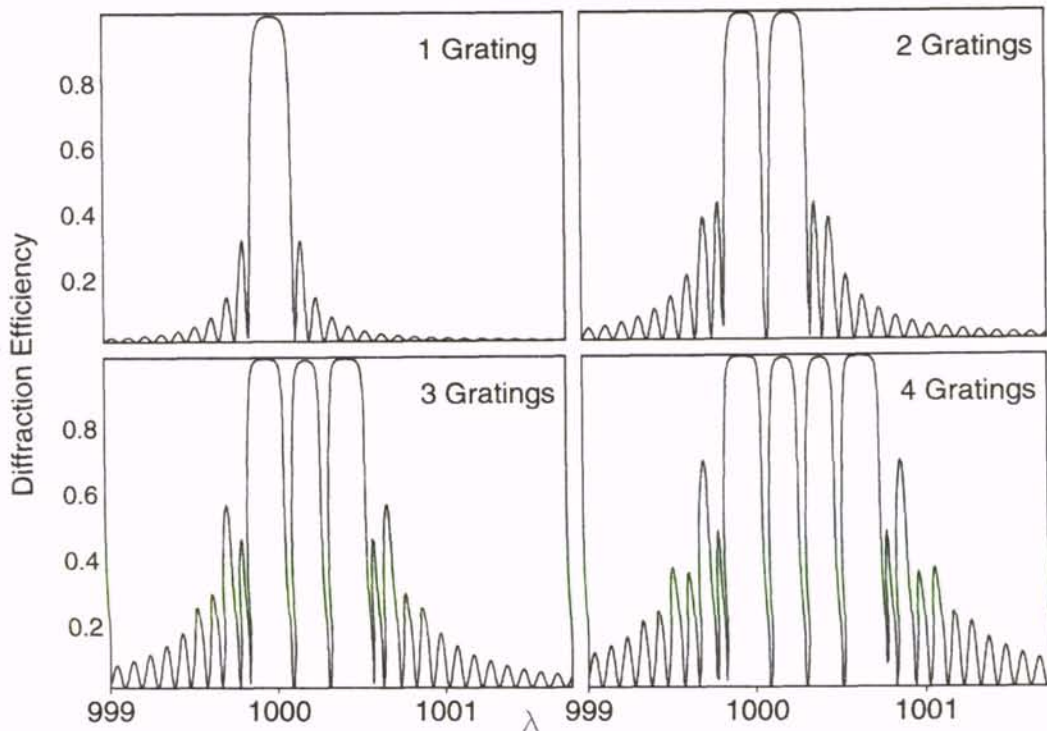
Figure 8.9. Wavelength response of one to four multiplexed reflection gratings

Clearly this filter has some interesting possibilities in wavelength division multiplexing (WDM) systems, however its response also has many undesirable characteristics which must be addressed. These include the fact that the rejection band on either side of the transmission peak is relatively narrow, and the transitions are not extremely sharp. From the early chapters of this work we know that the first of these can be corrected by increasing the index modulation (which increases the width of the main reflection lobe). In a similar way we can turn to traditional grating theory to find the solution to the second problem - increasing the number of periods in the grating (this has been used extensively in fiber Bragg gratings [11] since extremely thick gratings are easy to obtain). Figure 8.10 below shows the effect of doubling the grating thickness to 20000

periods (approximately 6 mm). As these devices get longer we are obviously entering an area best handled by gratings recorded in optical fibers.
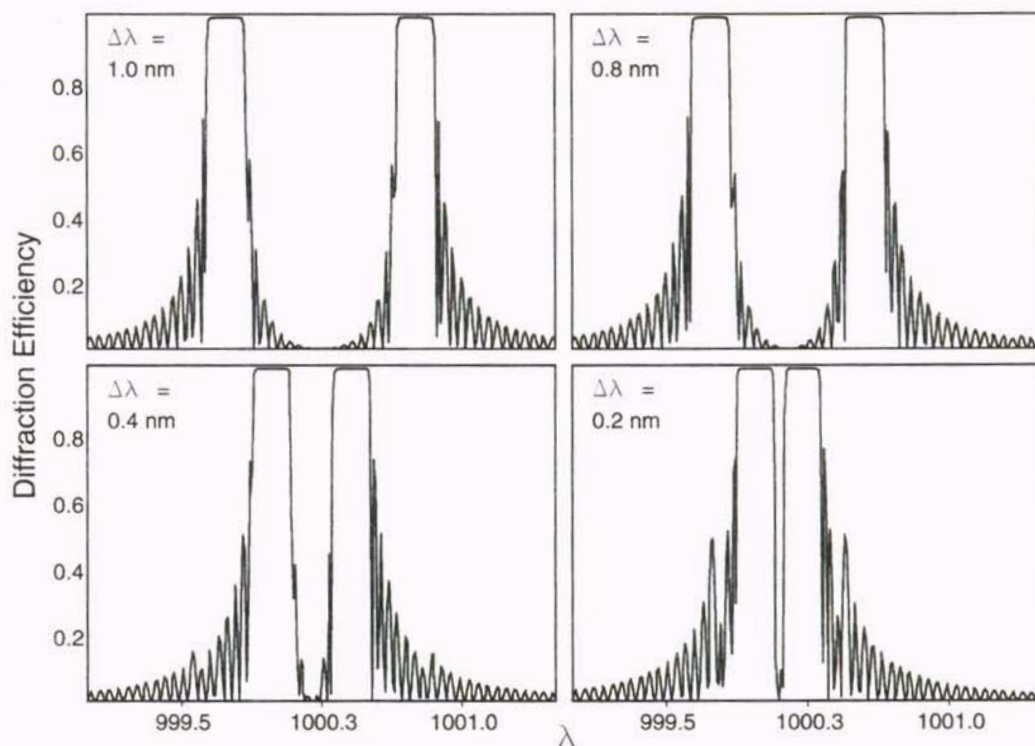
Figure 8.10. Wavelength response of two doubly thick multiplexed reflection gratings

At this point we have achieved the desired sharp transitions, however a new (and perhaps worse) problem has developed – multiple sidelobes. Although the amplitude of these sidelobes has not increased for any given amount of dephasing, the width of each one has narrowed, and considerably more of them exist. Furthermore sidelobes have turned up between the main peaks. The problem of reducing sidelobes has been addressed often in traditional as well as in fiber gratings [12-15]. The solution that is typically used (applying an apodization function) actually dates back into the early days

145

of microwave theory [16]. In this method the grating strength is varied as a function of position. Several possible functions have been investigated [17], the most popular of which are the raised cosine and the gaussian profiles. Since the raised cosine gives better sidelobe suppression it will be used in the following analysis. Cases where the thickenss is 5,000, 10,000, 20,000 and 40,000 periods are considered.
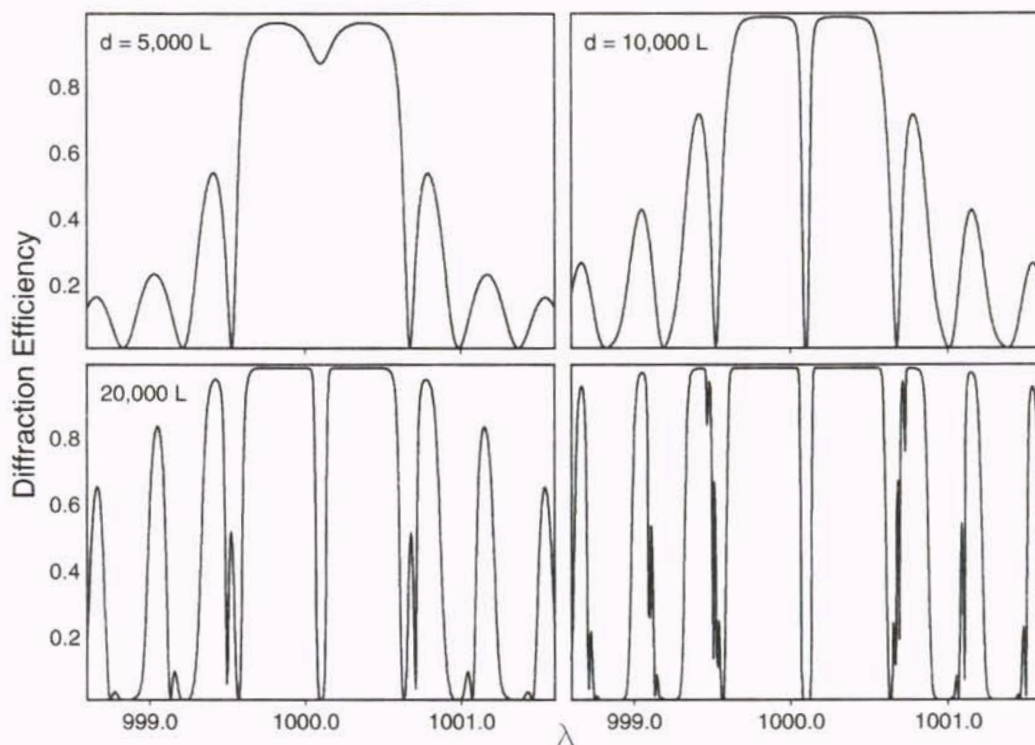
Figure 8.11. Response of multiplexed apodized reflection gratings – various thicknesses

Again we see that traditional methods have solved the problem. In all but the first case narrow linewidth with sharp transitions were attained. In the 5000 period plot the transitions were not sharp enough for a transmission spike to exist (the 0.2nm separation was originally selected based on a 10,000 period grating).

146

Since it appears that all undesirable characteristics of can be corrected, an investigation of manufacturing tolerances seems appropriate. Figure 8.12 shows one example of this - the wavelength response as the phase between the gratings is varied between 0 and π (i.e. one grating is kept as a cosine while the other is varied between a cosine and a sine). In this figure all gratings are 10,000 periods long and have an index modulation of 0.000295.
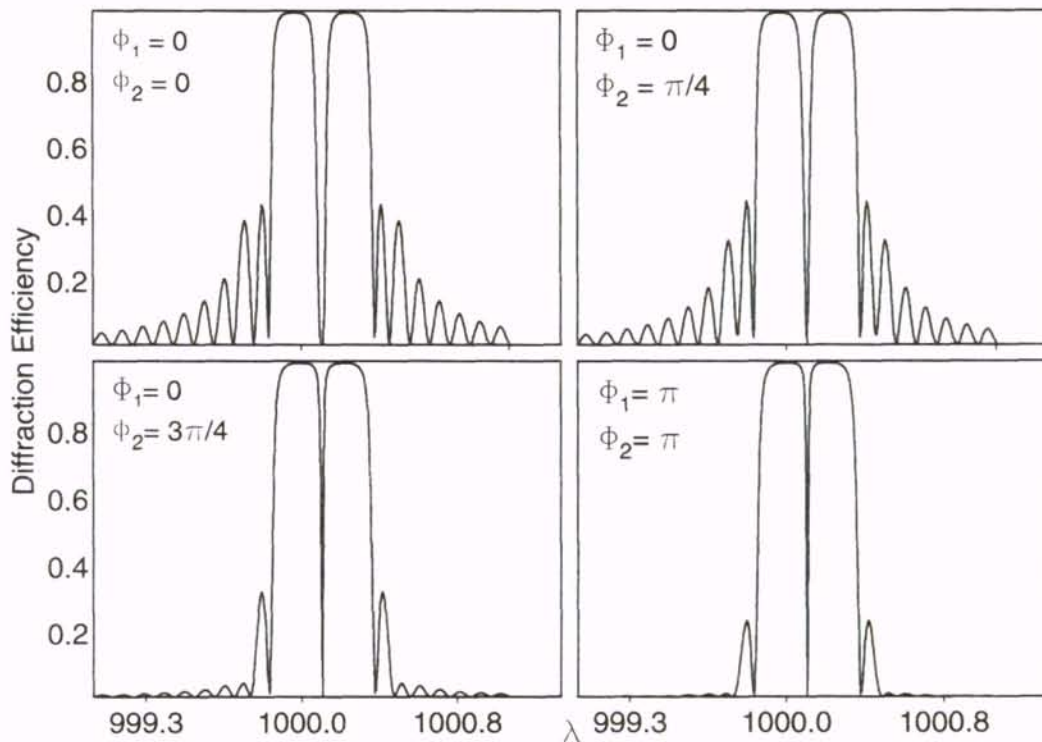
Figure 8.12. Response of a filter with different relative grating phases

Figure 8.12 indicates a serious potential problem with these devices. The width of the transmission peak is strongly dependent on the phase difference of the gratings. Since the phase of each grating is determined by that of the recording beams this overall

difference cannot be controlled. Thus these filters appear impractical for large scale use, since these operations require many filters with similar characteristics (as opposed to a single filter with an extremely narrow linewidth). It should be pointed out that similar effects were predicted in doubly multiplexed fiber Bragg gratings [17].

Before abandoning the idea of a multi-line filter it seems worthwhile to see if similar characteristics could be obtained from a filter manufactured in another manner. Some insight may be found by considering the mathematics of the interference pattern. Assuming that the gratings are of equal strength, the index distribution of the resulting combination of gratings becomes

$$F(z) = A\cos(K_1 z) + A\cos(K_2 z + \Phi) \tag{8.4}$$

With some simple algebra this can be written as

$$F(z) = 2A\cos\left[\left(\frac{K_1 z + K_2 z}{2}\right) + \frac{\Phi}{2}\right]\cos\left[\left(\frac{K_1 z - K_2 z}{2}\right) - \frac{\Phi}{2}\right] \tag{8.5}$$

In this form the pattern looks like a single grating (whose period is equal to the average of the two holographic periods), with a sinusoidal apodization placed upon it. Using a single mask with this amplitude distribution would alleviate the problems of reproducibility, however it is highly unlikely that such a mask could be manufactured (it would require an amplitude which varies between 0 and 1 in a sinusoidal fashion, along with several pi phase transitions located with extreme accuracy).

148

It is reasonable to assume that the narrow transmission peaks result more from the phase transitions (zero crossings) than from the sinusoidal amplitude distribution. Thus it may be possible to get the same effect with a mask of uniform amplitude, but with a π phase transition whenever equation 8.5 changes sign. Gratings of various lengths with this profile were modeled, and the wavelength characteristics are given below.
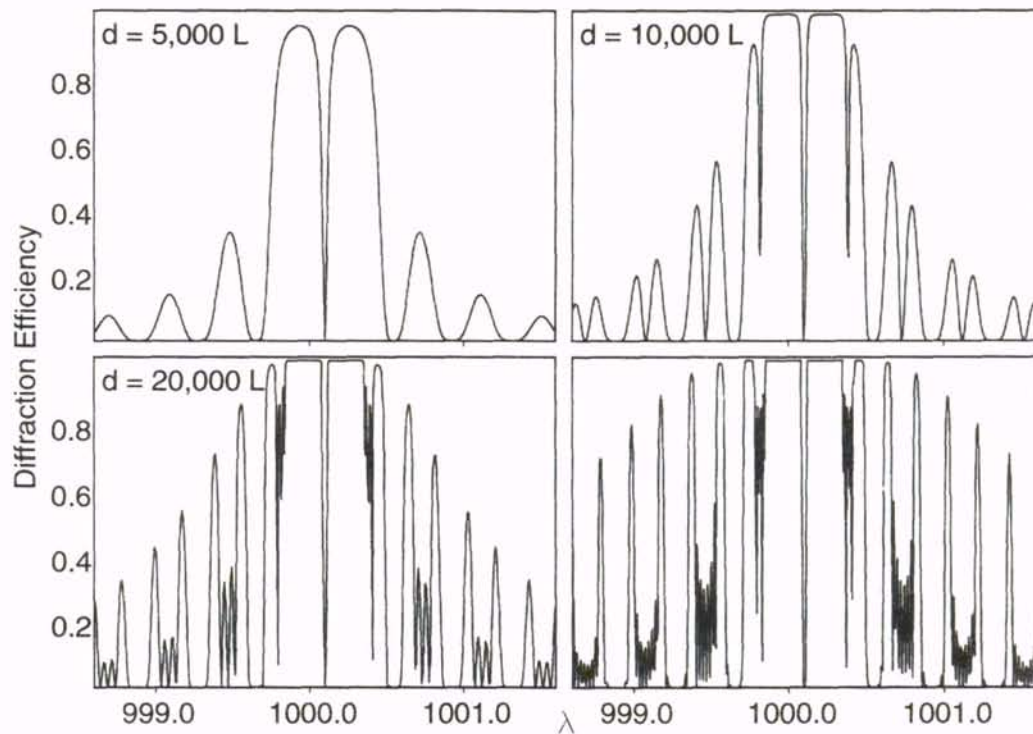
Figure 8.13. Response of a filter with simplified (±1) profile

This figure clearly shows that the simpler device does contain the desired transmission peaks. Although large sidelobes are also present in the responses, adding an apodization function will again lower their value. Even though the mask required to make this filter is simpler than the previous one, and could probably be manufactured

(making our multi-line filter truly a possibility), the number of phase transitions may make its attainment cost prohibitive.

By recognizing that these transitions are simply π phase shifts, we can see the origin of these transmission peaks. Essentially we have a number of reflection gratings each separated by a π phase shift. The entire ensemble is therefore acting as an array of Fabry-Perot interferometers, with the grating sections as the mirrors and the phase shifts as the spacers. Knowing this, and recognizing that these components can be made from thin films makes a thin film version of the entire device seem possible. Thus the final section of this chapter will discuss the characteristics of multi-line filters created from thin film layer stacks, and will describe how to vary the parameters to meet desired characteristics.

8.3   Thin Film Stack Multi-line Filters

Since it appears that the multi-line filters we have been discussing are essentially Fabry-Perot interferometers with multiple mirrors, a review of Fabry-Perot interferometer theory should yield additional insight into their behavior. We will begin this with the traditional (two mirror) case, while postponing the discussion of the multi-mirror situation until later.

The characteristics of a standard Fabry-Perot interferometer has been well studied and the results are readily available [18]. The transmittance of this device can be expressed as:

$$T = \frac{(1-R_1)(1-R_2)}{1+R_1 R_2 - 2\sqrt{R_1 R_2} \cos\delta}$$ (8.6)

where R_1 and R_2 are the reflectances of mirrors 1 and 2 respectively, and δ is the round trip phase difference of the spacer layer. A schematic drawing of the device along with the characteristics resulting from this equation (assuming $R_1 = R_2 = R$) are shown below in Figure 8.14. for several values of $r = \sqrt{R}$. Clearly the transmission peaks become narrower and sharper as the value of R increases.
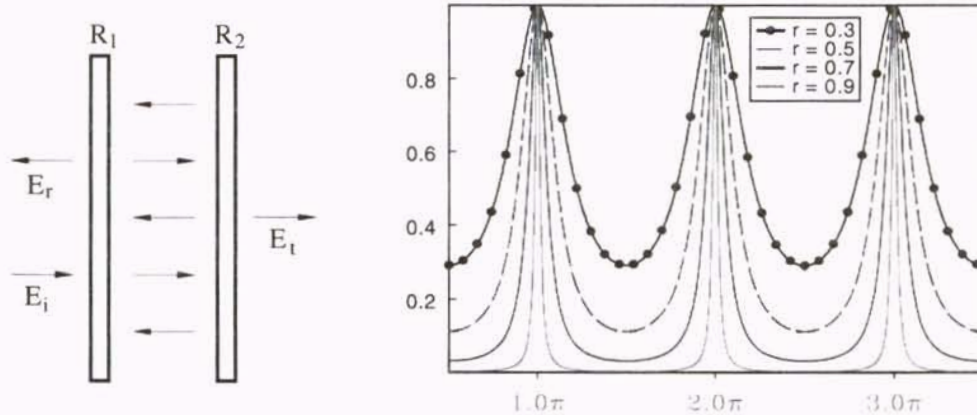
Figure 8.14. Fabry-Perot interferometer and resulting transmission characteristics

In addition to showing the transmission characteristics, equation 8.6 also gives the requirements necessary for the transmission of the device to reach 100%. The first of these is that the phase difference must be π (which is actually just the requirement for a transmission peak), while the second is that the mirrors have equal reflectivity. It will be shown presently that similar requirements hold when the interferometer is made with multiple mirrors.

Fabry-Perot Interferometers (FPIs) can be made in a number of ways, including using free standing mirrors separated by air or even a simple block of optical material with parallel sides (where the thickness determines the phase difference, and the reflections at the boundaries act as the mirrors). Another way to make a Fabry-Perot interferometer is with thin film stacks. In this method two stacks of quarter wave layers are joined by a half wave layer. Each of these stacks then acts as a mirror, and the half wave layer performs the duties of the spacer. Since the optical thickness of the spacer is exactly one half wave only at the design frequency, maximum transmission is limited to that wavelength with symmetric reductions occurring as this is varied.

Despite the similarities between the different versions of these devices, subtle differences exist. First, since the reflectivity of these thin film stacks is determined by the indices of refraction and the number of layers, the requirement for equal mirrors when using thin-films is that each stack must have the same number of layers. Additionally a real advantage to using thin films becomes evident when constructing multi-mirror versions. They are made by simply adding a half wave spacer and an additional stack to the previous design for each new mirror desired, which is quite simple.

Before continuing, a method to describe the various high-low layer combinations will be given. This is done by listing in order the index (either H or L) of the individual quarter wave sections. Thus, HLHS refers to a low index quarter wave section between two high index sections atop a substrate, while in HLLHS the low index section is a half wave thick. For compactness repeated H/L combinations will be designated as $(HL)^n$ where n is the number of repeated pairs. Thus the combination H L H L H H L H L H S is written as $(HL)^2 HH (LH)^2 S$.

152

An example of a thin film multi-line filter is shown below, along with the responses of several different versions. The individual combinations used are given in each plot. In every case the high and low indices of refraction were 2.0 and 1.5.
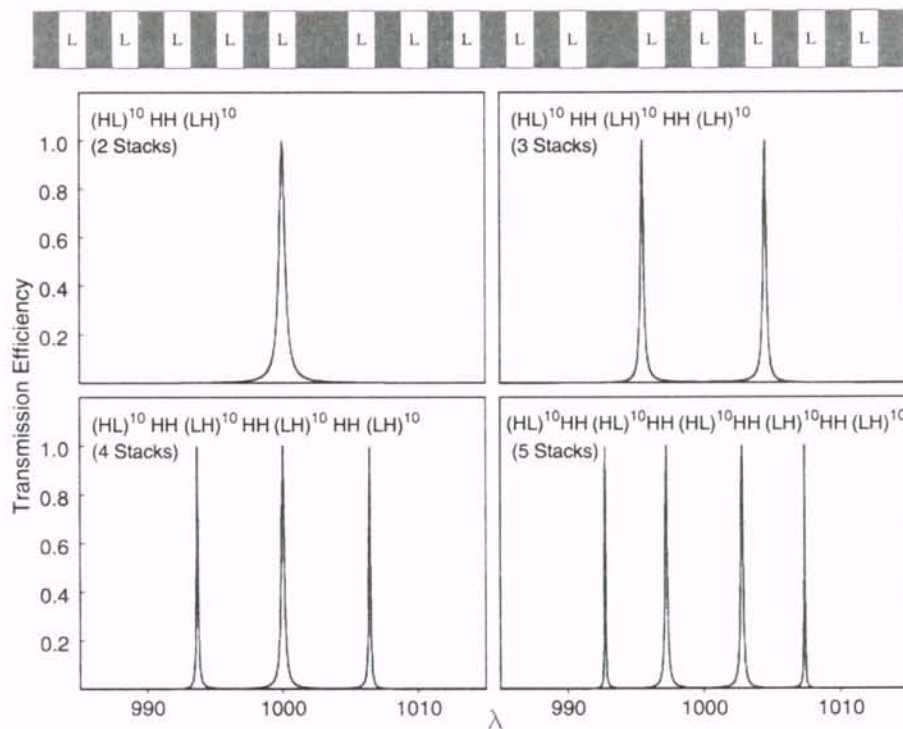
Figure 8.15. Example and response of thin film filters with multiple H/L stacks

Clearly this is the behavior that we have been seeking – several narrow transmission peaks separated by regions of near zero transmission. Additionally these responses contain no sidelobes, which is a big advantage over holographic alternatives.

These plots may give the impression that adding additional stacks automatically results in additional transmission peaks. Although this probably seems obvious, it is actually not even true in the general case. In order to determine how the filter will behave when a stack is added, one can use either multi-mirror Fabry-Perot interferometer theory [19] or traditional thin film filter theory [20,21]. With regard to the number of

153

transmission peaks, they give similar results (even though they are arrived at in entirely different manners). For the remainder of this discussion the multi-mirror FPI theory will be used since it appears to give more insight.

Van de Stadt and Muller [19] presented an analysis of multi-mirror FPIs which gave expressions for the transmissivity of three and four mirror versions of these devices. In this work it was shown that symmetric structures are required for the transmissivity to reach 100% (which is equivalent to the need for equal mirrors in the traditional case). This requirement has been seen to hold in the thin film case, thus only symmetric structures will be presented here. Furthermore, due to the complexity of the equations this analysis will be limited to cases with either three or four mirrors. The insight gained from this can however be generalized if additional mirrors are used.

The diagram of a three mirror FPI is shown below in Figure 8.16. The resulting transmission characteristics (assuming $R_1=R_3$ and $\delta_1=\delta_2=\delta$) are given by :

$$T = \frac{(1-R_1)^2 (1-R_2)}{(1-R_1)^2 (1-R_2)+\left[\sqrt{R_2}(1+R_1)-2\sqrt{R_1}\cos(\delta)\right]^2} \qquad (8.7)$$

where R_1, R_2 and R_3 are the mirrors reflectivities, and δ_1 and δ_2 are the round trip phases between the mirrors. This expression is a corrected version of that found in [19], where an error was made in determining the sign of the reflection coefficients at either side of the mirror. This resulted in a positive value for the cosine term, and a shift in the resonance location.

By setting equation 8.7 to 1 and recalling that the minimum value of the cosine is
−1 we can determine that the requirement for unity transmission is:

$$r_2 \leq \frac{2r_1}{\left(1 + r_1^2\right)}$$

(8.8)

If the equality sign in this expression holds then a single peak (of amplitude 1) will occur
at $\delta=\pi$. If the left side is less than the right side, two peaks (each with magnitude 1)
symmetric around $\delta=\pi$ will be present. If the right side of equation 8.8 is greater than the
left then a single peak will occur at $\delta=\pi$, however its magnitude will be less than unity.
These relationships are shown in Figure 8.16 below, which gives the transmission
characteristics of a three mirror Fabry-Perot interferometer as the phase difference δ is
varied. Several values are shown for the reflectivity of the inner mirror, while the
reflectivity of the outer mirrors remain constant at 70 percent.
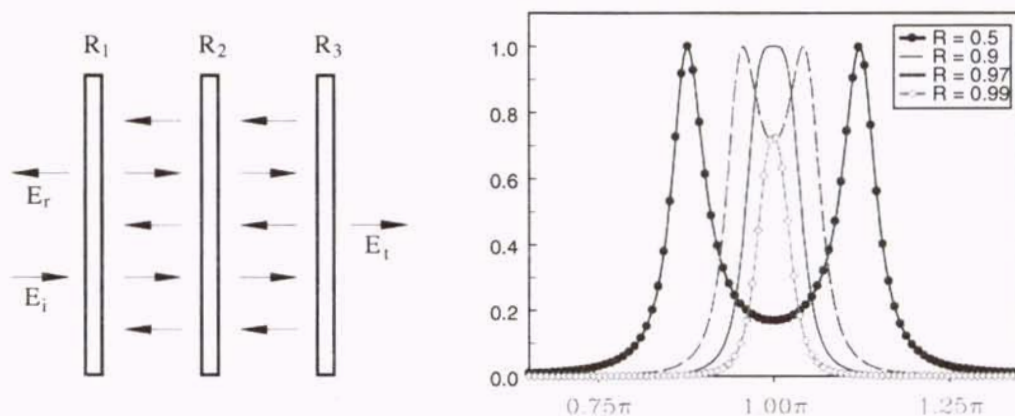
Figure 8.16. Three mirror FPI and resulting transmission characteristics

The following figure shows the responses of thin film versions of the 3-mirror FPI structure. The high and low indices of refraction used were 2.0 and 1.5. The outer stacks were kept constant at 13 layers, while inner ones were varied between 17 and 33 layers. The resulting reflectivities are such that in parts a and b of Figure 8.17 the requirements of equation 8.8 are met and two unity transmission peaks occur, while in part d of the figure these requirements are not met and a single of lower amplitude is found. Figure 8.17c represents the equality in relationship 8.8, and the single peak has 100% transmission. Clearly multi-mirror FPI theory accurately predicts the number of peaks.
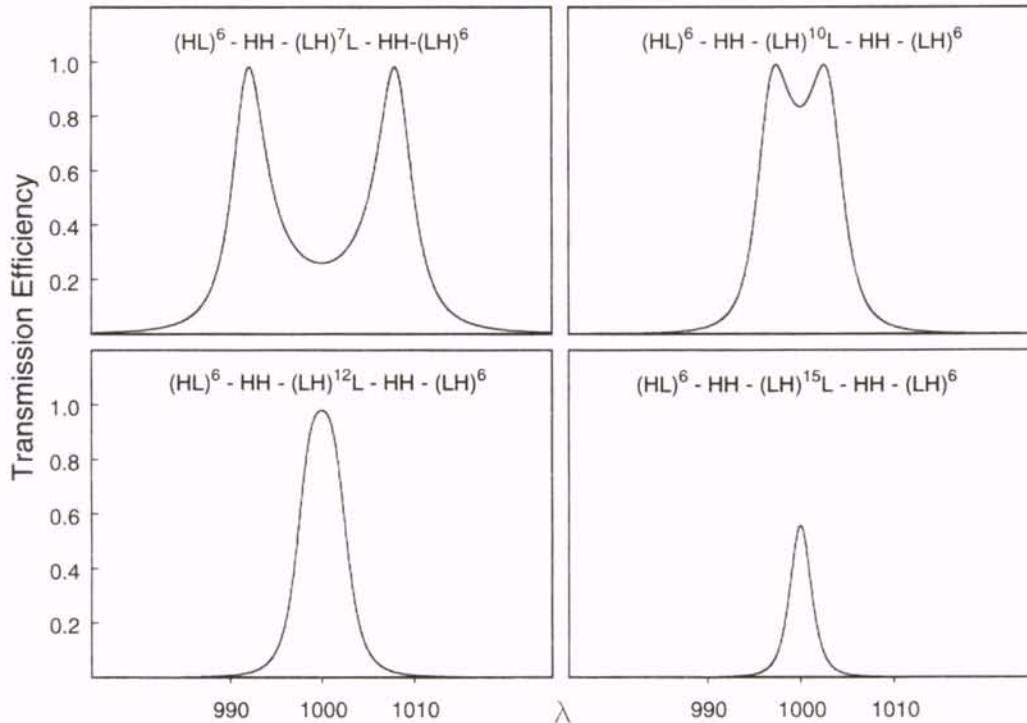
Figure 8.17. Response of a three-stack structure with various layers in each stack

A similar (albeit more complicated) expression was given in [19] for the transmission of the 4-mirror Fabry-Perot interferometer. Assuming a symmetric structure (i.e. $R_1=R_4$, $R_2=R_3$, and $\delta_1=\delta_2=\delta_3$), the transmission can be expressed as:

156

$$T = 1 / \left\{ 1 + 2r_1^2 \left(1 + \cos\delta\right)\left(2\cos\delta + B - 1\right)^2 / \left[\left(1 - r_1^2\right)^2 \left(1 - r_2^2\right)^2 \right] \right\}$$

where :   (8.9)

$$B = \left(1 + r_1^2 + r_1\, r_2\right) r_2 / r_1$$

and r is $\sqrt{R}$. Using this equation the condition for unity transmission becomes:

$$r_2^2 + r_2\left(r_1 + 1/r_1\right) \le 3 \qquad\qquad (8.10)$$

Again we can use these to investigate the filter response under various conditions. Figure 8.18 shows this response as the inner mirror reflectivity is varied while the outer mirror reflectivities are kept constant at 70%. In this situation an inner mirror reflectivity of 99.5% equates both sides of expression 8.10, and we see in the figure that all of the peaks again sit atop each other. We also see that this results in a response with the flattest top.
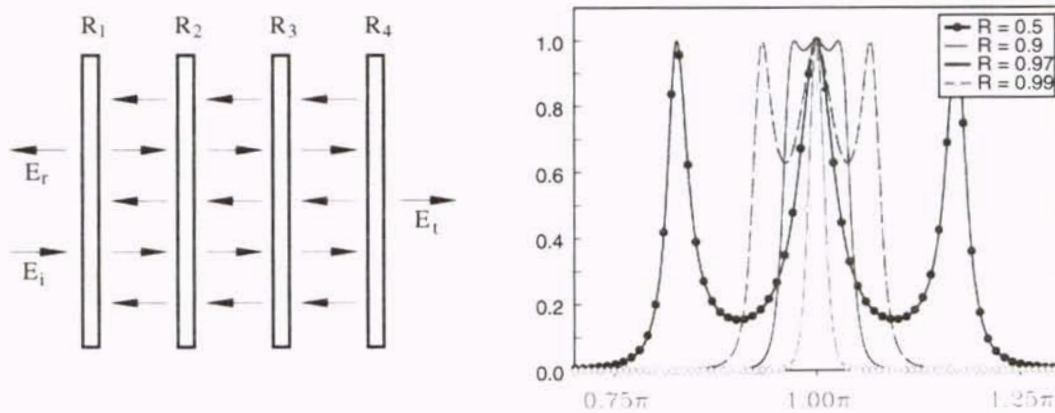
Figure 8.18.   Four-mirror FPI and resulting transmission characteristics

It is interesting to note that the plots for which the additional peaks have collapsed upon themselves (the equality signs in equation 8.8 and 8.10) correspond to traditional multi-cavity designs when using thin films. These designs use multiple half wave layers (hence the name multi-cavity), to give a filter with a single peak that is extra flat. This additional flatness is also present in Figures 8.16 and 8.18.

We saw above that as the reflectivities of the mirrors in an FPI are increased, the transmission peaks become narrower. Furthermore, we know that the reflectivity of a stack of quarter-wave layers increases with the number of layers. Thus we would anticipate that the widths can be controlled by varying the number of layers in each stack. This is examined for four-stack devices in Figure 8.19. In this investigation the number of layers in the inner stacks was kept constant at 17, while those of the outer stacks were varied between 13 and 25. Again the high and low indices of refraction were 2.0 and 1.5.
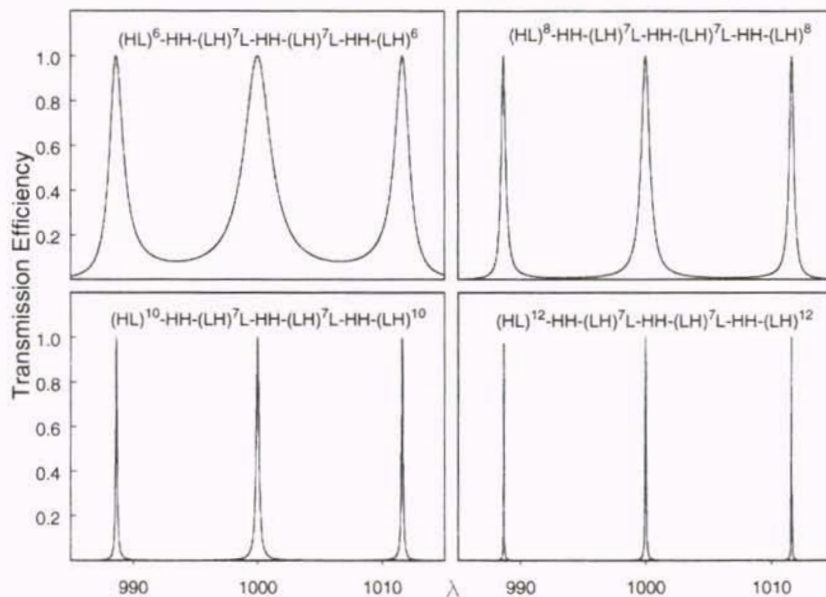
Figure 8.19. Varying the number of layers in the outer stacks of a multi-line filter

158

The behavior exhibited by this filter exactly matches what was anticipated. It thus appears that these devices do indeed act like typical multi-mirror FPIs. An additional observation from Figure 8.19 is that the response does not drop all the way to zero between the transmission peaks in the first two cases. The value of the transmission in these areas is equal to that of the individual stacks since the phase difference prevents constructive interference. For the number of layers and indices used in this example the reflectivity is approximately 88%, which explains the 12% minimum in the graph. Clearly this figure indicates that the number of layers in the outer stack determines the linewidths of the filter (by setting the reflectivity of the stack). Again expressions for these linewidths will be given at the end of this chapter.

Since it has been shown that the outer stacks determine the width of the transmission peaks, it is natural to question what the inner stacks do. Figures 8.16 and 8.18 suggest that the reflectivity of these layers will determine the quantity and the location of the transmission peaks. Figure 8.17 verified that the number of transmission peaks was indeed found from these stacks. In order to see if these stacks also determine the location Figure 8.20 has been included. In this figure the number of layers in the outer stacks was held constant at 21 (representing a total reflectivity of about 98.7%), while the number in the inner stacks was varied between 17 and 23. As in all previous examples the high and low indices of refraction were chosen as 2.0 and 1.5, which gave outer stacks with reflectivities that varied between 96% and 99.3%.
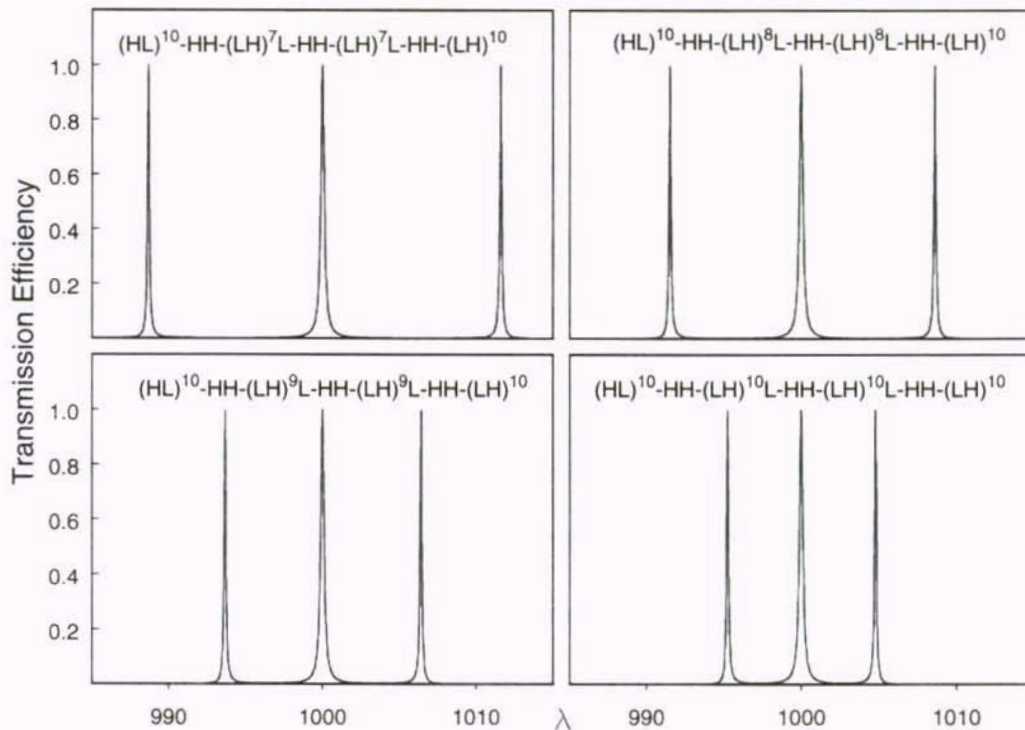
Figure 8.20. Varying the number of layers in the inner stacks of a multi-line filter

This figure indeed indicates that the number of layers in the inner stacks determines the location of these resonances. In order to separate these transmission peaks further the number of layers is decreased, while it is increased to bring them closer. Since this behavior was predicted by multi-mirror interferometer theory, it appears that we can design devices to meet desired criteria. Unfortunately however, the exact locations of the resonances in Figure 8.20 do not correspond to those predicted from equation 8.9. A similar difference in location is found when three stacks are included in the analysis. Furthermore it seems that no simple scaling parameter can be used to reconcile these locations.

It turns out that accurate predictions of the location of these resonances can be found using traditional methods of thin film optics. One approach is to replace the entire device (including all spacer layers) with an equivalent one composed of only one or two layers. The typical Fabry-Perot rules are then applied to this simplified structure (i.e. a transmission peak is found when the optical phase difference is an integral multiple of π at some wavelength). The difficulty lies in determining these equivalent parameters. In the case of symmetric stacks however, a means of calculating them was given by Epstein [22] and extended for practical use by Thelen [23]. This method was described by Macleod [24], so only a brief review of the results will be given below.

The phase shift experienced by a wave as it propagates through a layer of thickness d at an angle θ is given by $\delta = 2\pi n d \cos\theta/\lambda$. Using this, the fields at the input to the layer can be written in terms of those at the output as:

$$\begin{bmatrix} E_{in} \\ H_{in} \end{bmatrix} = \begin{bmatrix} \cos\delta & (j\sin\delta)/Y \\ jY\sin\delta & \cos\delta \end{bmatrix} \begin{bmatrix} E_{out} \\ H_{out} \end{bmatrix} \tag{8.11}$$

where Y is the admittance of the layer. We can define an input optical admittance for the entire stack as: $Y_T = H_{in}/E_{in}$. Adding additional stacks merely adds additional matrices to equation 8.11, thus for a stack of q layers the resulting characteristic matrix is:

$$[M] = \prod_{k=1}^{q} \begin{bmatrix} \cos \delta_k & (j \sin \delta_k)/Y_k \\ j Y_k \sin \delta_k & \cos \delta_k \end{bmatrix} \tag{8.12}$$

Consider the special case of a three layer symmetric ($n_1=n_3$, $\varphi_1=\varphi_3$) stack. In this situation the components of equation 8.12 become:

$$M_{11} = M_{22} = \cos 2\delta_1 \cos \delta_2 - \frac{1}{2}\left(\frac{Y_2}{Y_1} + \frac{Y_1}{Y_2}\right)\sin 2\delta_1 \sin \delta_2$$

$$M_{12} = \frac{j}{Y_1}\left[\sin 2\delta_1 \cos \delta_2 + \frac{1}{2}\left(\frac{Y_2}{Y_1} + \frac{Y_1}{Y_2}\right)\cos 2\delta_1 \sin \delta_2 + \frac{1}{2}\left(\frac{Y_1}{Y_2} - \frac{Y_2}{Y_1}\right)\sin \delta_2\right] \tag{8.13}$$

$$M_{21} = j Y_1\left[\sin 2\delta_1 \cos \delta_2 + \frac{1}{2}\left(\frac{Y_1}{Y_2} + \frac{Y_2}{Y_1}\right)\cos 2\delta_1 \sin \delta_2 - \frac{1}{2}\left(\frac{Y_1}{Y_2} - \frac{Y_2}{Y_1}\right)\sin \delta_2\right]$$

Now if two quantities φ_E and Y_E are defined by:

$$\cos \varphi_E = M_{11} = M_{22} \quad \text{and} \quad M_{12} = \frac{j \sin \varphi_E}{Y_E} \tag{8.14a,b}$$

Then since $M_{11} M_{22} - M_{12} M_{21} = 1$ it follows that:

$$M_{12} = j Y_E \sin \varphi_E \tag{8.15}$$

162

These quantities have the same form as a single layer of phase thickness φ_E and admittance Y_E. Thus equations 8.13 and 8.14 can be solved to obtain values for these equivalent parameters. At the particular wavelength for which the layers have a quarter-wave optical thickness the equivalent parameters are given by:

$$\phi_E = 3\frac{\pi}{2} \qquad \text{and} \qquad Y_E = \frac{\eta_1^2}{\eta_2} \qquad (8.16)$$

At wavelengths away from this value the expressions for the equivalent parameters get quite complicated. From equations 8.13, 8.14 and 8.15 the following are obtained:

$$Y_E = \left(\frac{M_{21}}{M_{12}}\right)^{\frac{1}{2}} = \left(\frac{Y_1^2\left[\sin 2\delta_1 \cos\delta_2 + \frac{1}{2}\left(\frac{Y_1}{Y_2}+\frac{Y_2}{Y_1}\right)\cos 2\delta_1 \sin\delta_2 - \frac{1}{2}\left(\frac{Y_1}{Y_2}-\frac{Y_2}{Y_1}\right)\sin\delta_2\right]}{\sin 2\delta_1 \cos\delta_2 + \frac{1}{2}\left(\frac{Y_1}{Y_2}+\frac{Y_2}{Y_1}\right)\cos 2\delta_1 \sin\delta_2 + \frac{1}{2}\left(\frac{Y_1}{Y_2}-\frac{Y_2}{Y_1}\right)\sin\delta_2}\right)^{\frac{1}{2}}$$

$$\varphi_E = a\cos\left[\cos 2\delta_1 \cos\delta_2 - \frac{1}{2}\left(\frac{Y_1}{Y_2}+\frac{Y_2}{Y_1}\right)\sin 2\delta_1 \sin\delta_2\right] \qquad (8.17)$$

This expression is multi-valued so the value chosen is the one nearest to $2\delta_1+\delta_2$, (the actual sum of the phase thicknesses). Figure 8.21 shows the characteristics of this phase for a stack of 5 high and low layers (with indices of 2.0 and 1.5 respectively) as the incident wavelength is varied. In this plot we see that the phase varies by $\pm \pi/2$ around the value 2.5π $(3\delta_1+2\delta_2)$. We also see that there are no values for $\lambda<0.85\lambda 0$ or $\lambda>1.15\lambda 0$

163

(actually the values are imaginary). Mathematically this occurs because $M_{11} < -1$ (i.e. cos $\varphi_E < 1$) in equation 8.13a. In these regions the equivalent index (from equation 8.17) approaches infinity. These represent transmission stop bands, and occur when the stack becomes highly reflecting. As with φ_E, η_E is also imaginary in this region.
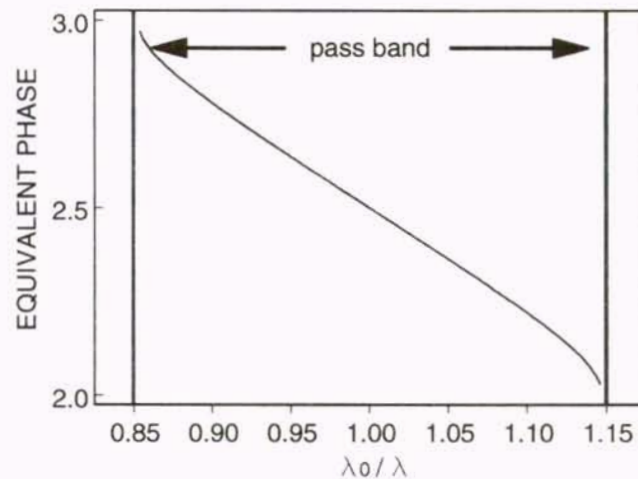
Figure 8.21. Equivalent phase thickness of a stack of layers – $n_H = 2.0$, $n_L = 1.5$

Although the above derivation was made for a stack of three layers it is actually valid for any symmetric stack. To see this, consider a stack of five alternating high and low layers: H-L-H-L-H. By following the above procedure we can replace the 3 symmetric center layers (L-H-L) with a single equivalent layer E. This leaves H-E-H. This too is a symmetric layer, so it can be reduced to a single layer. Clearly this can be repeated any number of times to reduce a symmetric stack to a single layer. It should be noted that these equivalent layers are not exact replacements for the true stacks, and there are some phenomena that cannot be accounted for in this way [24].

164

From these formulas we can see that at the design wavelength, a stack of q alternating high and low index layers will have an effective phase and index given by:

$$\varphi_E = q\frac{\pi}{2} \quad \text{and} \quad n_E = \frac{n_H^{(q+1)/2}}{n_L^{(q-1)/2}} \tag{8.18}$$

The phase behavior seen in Figure 8.21 above is typical for any symmetric stack. Regardless of the number of layers present the phase will vary by $\pm\pi/2$ across a narrow transmission band which can be given approximately by [24]:

$$\Delta\lambda = \frac{4(n_H - n_L)n_L^{(q-1)/2}}{\pi n_H^{(q+1)/2}}\lambda_0 \tag{8.19}$$

Because Fabry-Perot theory tells us that the phase difference must be an integral multiple of π for a transmission peak to occur, then these peaks cannot occur when only a single stack is present (since $(q-1)\pi/2 < \varphi < (q+1)\pi/2$, φ never reaches an integral number of π). If however an exact copy of this stack is placed behind the first then the combination will have the same effective index, but double the phase difference. Thus it will vary by $\pm\pi$ around the center point of $q\pi/2$, and a transmission peak will be found at $q\pi$ (again peaks can not occur at $(q-1)\pi$ and $(q+1)\pi$ because these are in the high reflection (stop band) region). Adding a third stack would give an overall phase difference of 3π centered around $q\pi/2$, which would result in two peaks. This procedure can be repeated to add additional peaks.

165

This procedure has been used by Macleod to design dual line transmission filters for wideband applications (i.e. two transmission bands hundreds of nanometers wide separated by large bands), however appears untested in narrowband filtering applications.

Although this method gives the location of the transmission peaks, a much simpler method can be used to obtain the same information, along with information on the FWHM of the peaks. Again it is based on the admittance of a stack of layers, but its simplicity gives more insight into the problem.

The normalized admittance of any layer versus position can be written as:

$$Y(z) = Y_0 \frac{Y(0) - jY_0 \tan(k_z z)}{Y_0 - jY(0) \tan(k_z z)} = Y_0 \frac{Y(0) + jY_0 \tan(-k_z z)}{Y_0 + jY(0) \tan(-k_z z)} \quad (z \text{ is negative })$$

(8.20)

In traversing an almost quarter wave (including higher order) layer this admittance is:

$$Y(z) = Y_0 \frac{Y(0) - jY_0 \cot \delta_{\pi/2}}{Y_0 - jY(0) \cot \delta_{\pi/2}} = Y_0 \frac{jY(0) \tan \delta_{\pi/2} + Y_0}{jY_0 \tan \delta_{\pi/2} + Y(0)} \quad p = 0,1,2 \cdots$$

where:

(8.21)

$$\delta_{\pi/2} \equiv \pi \left(p + \frac{1}{2} \right) \left[\frac{\Delta n}{n} - \frac{\Delta \lambda}{\lambda_0 + \Delta \lambda} - \frac{\Delta n}{n} \frac{\Delta \lambda}{\lambda_0 + \Delta \lambda} \right]$$

while in traversing a half wave (including higher order) layer it becomes:

$$Y = Y_0 \frac{Y(0) + jY_0 \tan \delta_\pi}{Y_0 + jY(0) \tan \delta_\pi} \qquad p = 1,2,3,...$$

where :

(8.22)

$$\delta_\pi \equiv p\pi \left[\frac{\Delta n}{n} - \frac{\Delta \lambda}{\lambda_0 + \Delta \lambda} - \frac{\Delta n}{n} \frac{\Delta \lambda}{\lambda_0 + \Delta \lambda} \right]$$

This process can be repeated to give the following expression for a H/L pair:

$$Y = Y_s \left(\frac{Y_L}{Y_H} \right)^2 \left[\frac{(1 - Y_H / Y_L \tan \delta_{H,\pi/2} \tan \delta_{L,\pi/2}) + j(\tan \delta_{H,\pi/2} + (Y_H / Y_L) \tan \delta_{L,\pi/2})(Y_H / Y_s)}{(1 - Y_L / Y_H t \tan \delta_{H,\pi/2} \tan \delta_{L,\pi/2}) + j(\tan \delta_{H,\pi/2} + (Y_L / Y_H) \tan \delta_{L,\pi/2})(Y_s / Y_H)} \right]$$

(8.23)

At the wavelength for which these layers are exactly $\lambda/4$ ($\delta=0$) this formula reduces to:

$$Y = \left(\frac{Y_L}{Y_H} \right)^2 Y_s$$

(8.24)

Again this can again be repeated for a stack of M quarter wave layers:

$$Y = \left(\frac{Y_L}{Y_H} \right)^{2M} Y_s$$

(8.25)

If $\delta_{\pi/2} \leq 0.01$ and $Y_S = Y_H$ equation 8.23 becomes:

$$Y_{out} \cong Y_s \left(\frac{Y_L}{Y_H} \right)^2 \left[1 + j \tan \delta_{L,\pi/2} \left(\frac{Y_H}{Y_L} - \frac{Y_L}{Y_H} \right) \right]$$

(8.26)

167

while for a stack of these H/L pairs it is:

$$Y_{out} \cong Y_s \left(\frac{Y_L}{Y_H}\right)^{2M} \left[1 + j\tan\delta_{L,\pi/2}\left(\frac{Y_H}{Y_L} - \frac{Y_L}{Y_H}\right)\right]^M \qquad (8.27)$$

Equations 8.20 through 8.27 give expressions for most of the combinations of high and low pair seen in thin film stacks, thus they can be used as building blocks from which the characteristics of any thin film filter can be found. For instance, Figure 8.22 below shows a device consisting of a half wave section between two quarter-wave sections sitting on a substrate:
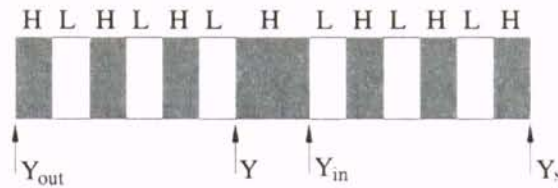
Figure 8.22. Dual stack thin film filter

Equations 8.22 and 8.24 can now be used to find the input impedance at each of the points labeled in Figure 8.22. By selecting equation 8.22 instead of 8.21 we are assuming that the phase difference in traveling through the stacks is negligible. This is in accordance with our understanding that the stacks are acting primarily as partially reflecting mirrors which would not be as sensitive to the phase. Equations 8.28 through 8.30 below give the resulting impedances.

168

$$Y_{in} = Y_S \left(\frac{Y_L}{Y_H} \right)^{2M} \tag{8.28}$$

$$Y = Y_H \frac{Y_{in} + jY_H \tan\delta_\pi}{Y_H + jY_{in} \tan\delta_\pi} \tag{8.29}$$

$$Y_{out} = Y \left(\frac{Y_H}{Y_L} \right)^{2M} = Y_s \left[1 + j \left(\frac{Y_H}{Y_S} \right) \left(\frac{Y_H}{Y_L} \right)^{2M} \tan\delta_\pi \right] \left[1 + j \left(\frac{Y_S}{Y_H} \right) \left(\frac{Y_L}{Y_H} \right)^{2M} \tan\delta_\pi \right]^{-1} \tag{8.30}$$

For small values of δ, equation 8.30 simplifies to:

$$Y_{out} \cong Y_s \left[1 + j \tan\delta_\pi \left(\left(\frac{Y_H}{Y_S} \right) \left(\frac{Y_H}{Y_L} \right)^{2M} - \left(\frac{Y_S}{Y_H} \right) \left(\frac{Y_L}{Y_H} \right)^{2M} \right) \right] \tag{8.31}$$

Additionally, if M is large it can be further reduced it to:

$$Y_{out} \cong Y_s \left[1 + j \left(\frac{Y_H}{Y_S} \right) \left(\frac{Y_H}{Y_L} \right)^{2M} \tan\delta_\pi \right] \tag{8.32}$$

By choosing the index of the substrate and incident medium to be that of the high layers we can write the reflection coefficient, and reflected and transmitted intensities as:

$$r = \frac{Y_0 - Y_{out}}{Y_0 + Y_{out}} \cong \frac{-j(Y_H / Y_L)^{2M} \tan \delta_\pi}{2 + j(Y_H / Y_L)^{2M} \tan \delta_\pi}$$

$$I_R = \delta_\pi^2 (Y_H / Y_L)^{4M} / \left[4 + \delta_\pi^2 (Y_H / Y_L)^{4M}\right] \qquad (8.33)$$

$$I_T = 1 / \left[1 + 0.25\delta_\pi^2 (Y_H / Y_L)^{4M}\right]$$

We can see from these equations that a transmission peak will occur when $\delta=0$ (i.e. at the wavelength for which the stacks are quarter wave). Additionally we can solve these equations for the following HWHM values:

$$\delta_{\pi,HWHM} = \frac{1}{2}(Y_L / Y_H)^{2M} = p\pi \left[\frac{\Delta n}{n} - \frac{\Delta\lambda}{\lambda_0 + \Delta\lambda} - \frac{\Delta n}{n}\frac{\Delta\lambda}{\lambda_0 + \Delta\lambda}\right]$$

$$\left. \frac{\Delta n}{n} \right|_{\pi,HWHM} = \left. \frac{\Delta\lambda}{\lambda} \right|_{\pi,HWHM} = \frac{2}{p\pi}(Y_L / Y_H)^{2M} = \frac{2}{p\pi}(n_L / n_H)^{2M} \quad p = 1,2,3\cdots$$

At this point all of the traditionally important filter parameters have been determined for this simple two-stack filter. A similar (albeit more complicated) method can be used for multi-stack devices, an example of which is given in Figure 8.23 below:
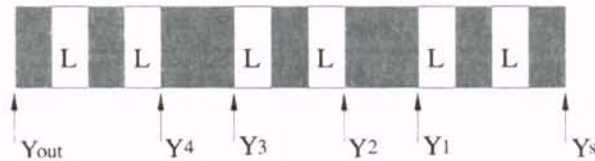
Figure 8.23. Triple stack thin film filter

In calculating the location of the transmission peaks for this filter it is possible to use the exact formulas of equations 8.21 and 8.22, but this will restrict us to a numerical solution. If however we again assume that the outer stacks act primarily as mirrors we can use the approximate forms for the quarter-wave stacks get a closed form solution for these values. Doing this yields the following values for the impedances listed in Figure 8.23:

$$Y_1 = Y_s \left(\frac{Y_L}{Y_H} \right)^{2M} = Y_s A^{-1} \quad \text{where} \quad A = \left(\frac{Y_H}{Y_L} \right)^{2M} \tag{8.35}$$

$$Y_2 = Y_H \frac{Y_1 + jY_H \tan\delta_{1,\pi}}{Y_H + jY_1 \tan\delta_{1,\pi}} = \frac{Y_s}{A} \frac{1 + jY_H Y_s^{-1} A \tan\delta_{1,\pi}}{1 + jY_H^{-1} Y_s A^{-1} \tan\delta_{1,\pi}} \tag{8.36}$$

The process is now continued to get the remaining impedances. Figure 8.23 indicates that the second stack can be considered as either $(LH)^{N-1}L$ or as $(HL)^N$. Either way will give the same results, thus it will be derived using the first notation. The resulting impedances (which are again found from the formulas 8.20 – 8.28) are:

$$Y_3 = \frac{Y_L^2}{Y_1} \left(\frac{Y_L}{Y_H} \right)^{2(N-1)} = \frac{Y_H^2}{Y_s} \frac{A}{B} \frac{1 + jY_s Y_H^{-1} A^{-1} \tan\delta_{1,\pi}}{1 + jY_s^{-1} Y_H A \tan\delta_{1,\pi}}$$

$$Y_4 = Y_H \frac{Y_3 + jY_H \tan\delta_{2,\pi}}{Y_H + jY_3 \tan\delta_{2,\pi}} = \frac{Y_H^2}{Y_s} \frac{A\left(1 + jY_H^{-1} Y_s A^{-1} \tan\delta_{1,\pi}\right) + jY_s Y_H^{-1} B\left(1 + jY_H Y_s^{-1} A \tan\delta_{1,\pi}\right)\tan\delta_{2,\pi}}{B\left(1 + jY_H Y_s^{-1} A \tan\delta_{1,\pi}\right) + jY_H Y_s^{-1} A\left(1 + jY_H^{-1} Y_s A^{-1} \tan\delta_{1,\pi}\right)\tan\delta_{2,\pi}} \tag{8.37}$$

Finally, we make the approximation that the outer stack also acts simply as a mirror, which yields:

$$Y_{out} = Y_2 \left(\frac{Y_H}{Y_L}\right)^{2M}$$

$$= A \frac{Y_H^2}{Y_s} \frac{A\left(1 + jY_H^{-1}Y_s A^{-1} \tan\delta_{1,\pi}\right) + jY_s Y_H^{-1} B\left(1 + jY_H Y_s^{-1} A \tan\delta_{1,\pi}\right)\tan\delta_{2,\pi}}{B\left(1 + jY_H Y_s^{-1} A \tan\delta_{1,\pi}\right) + jY_H Y_s^{-1} A\left(1 + jY_H^{-1}Y_s A^{-1} \tan\delta_{1,\pi}\right)\tan\delta_{2,\pi}} \qquad (8.38)$$

By setting $Y_H = Y_S$ the following simplified expression is obtained:

$$\frac{Y_{out}}{Y_H} \cong A \frac{A + jB \tan\delta_{2,\pi} - AB \tan\delta_{1,\pi} \tan\delta_{2,\pi}}{B + jA\left(B \tan\delta_{1,\pi} + \tan\delta_{2,\pi}\right)} \qquad (8.39)$$

The field and intensity reflection coefficients can be calculated from this, which gives:

$$r = \frac{1 - \dfrac{Y_{out}}{Y_H}}{1 + \dfrac{Y_{out}}{Y_H}} = \frac{\left(B + jAB \tan\delta_{1,\pi} + jA \tan\delta_{2,\pi}\right) - A\left(A - AB \tan\delta_{1,\pi} \tan\delta_{2,\pi} + jB \tan\delta_{2,\pi}\right)}{\left(B + jAB \tan\delta_{1,\pi} + jA \tan\delta_{2,\pi}\right) + A\left(A - AB \tan\delta_{1,\pi} \tan\delta_{2,\pi} + jB \tan\delta_{2,\pi}\right)}$$

$$|r|^2 = \frac{\left(BA^{-2} - 1 + B \tan\delta_{1,\pi} \tan\delta_{2,\pi}\right)^2 + \left[BA^{-1}\left(\tan\delta_{1,\pi} - \tan\delta_{2,\pi}\right) + A^{-1} \tan\delta_{2,\pi}\right]^2}{\left(BA^{-2} + 1 - B \tan\delta_{1,\pi} \tan\delta_{2,\pi}\right)^2 + \left[BA^{-1}\left(\tan\delta_{1,\pi} + \tan\delta_{2,\pi}\right) + A^{-1} \tan\delta_{2,\pi}\right]^2} \qquad (8.40)$$

$$T = \frac{4BA^2\left(1 - B \tan^2\delta_{1,\pi}\right) + A^2 B^4 \tan^2\delta_{1,\pi}\left(2B + B^2\right)}{B^2 + A^4\left(1 - B \tan^2\delta_{1,\pi}\right)^2 + 2BA^2\left(1 - B \tan^2\delta_{1,\pi}\right) + A^2 B^4 \tan^2\delta_{1,\pi}\left[2B + 1\right]^2}$$

To find the location of the resonance $|r|^2$ is set to zero. This yields:

172

$$B - A^2 + A^2 B \tan\delta_{1,\pi} \tan\delta_{2,\pi} = 0 \qquad (8.41)$$

which for small δ becomes:

$$\delta_{1,\pi} \delta_{2,\pi} = \frac{A^2 - B}{A^2 B} \cong \frac{1}{B} \qquad (8.42)$$

From this expression the following resonance location is found:

$$\frac{\Delta\lambda}{\lambda} = \frac{1}{\pi}\frac{1}{\sqrt{B}} \quad \rightarrow \quad \Delta\lambda = \lambda_0\left(\frac{1}{\pi\sqrt{B}-1}\right) \qquad (8.43)$$

To obtain the FWHM of these transmission peaks we again consider standard Fabry-Perot interferometer theory. This gives:

$$I = \frac{1}{1 + F\sin^2\phi} \quad \text{where} \quad F = \frac{4R}{(1-R)^2} \quad , \quad \phi = k_z z$$

$$\qquad\qquad (8.44)$$

$$\phi\big|_{HWHM} = \sin^{-1}\left(\frac{1}{\sqrt{F}}\right)$$

By inserting the following expressions for the admittance and reflectance of quarter-wave stacks we can get the desired linewidth:

$$Y = \frac{n_s}{n_0}\left(\frac{n_L}{n_H}\right)^{2M} \quad \text{and} \quad R = \left[\frac{1-Y}{1+Y}\right]^2$$

$$F = 0.25\left[\frac{1}{Y} - Y\right]^2 = \frac{1}{4Y^2}\left(1 - 2Y^2 + Y^4\right) \approx \frac{1}{4Y^2} \tag{8.45}$$

$$\phi\big|_{HWHM} = \frac{1}{\sqrt{F}} = \frac{2Y}{1-Y^2} = 2Y\left(1-Y^2\right)^{-1} \approx 2Y$$

Written in terms of wavelength deviations these become:

$$\Delta\phi\big|_{FWHM} = 2\pi\frac{\Delta\lambda}{\lambda} \approx 4Y \qquad \frac{\Delta\lambda}{\lambda}\bigg|_{FWHM} = \frac{4Y}{\pi} = \frac{4}{\pi}\frac{n_0}{n_s}\left(\frac{n_L}{n_H}\right)^{2M} \tag{8.46}$$

If the overall impedance of each example given in this chapter is calculated using equations 8.21 and 8.22 , the locations of the predicted transmission peaks line up exactly with those obtained rigorously (Figures 8.15, 8.19 and 8.20). Using the simplified expression of equation 8.43 however, requires a correction factor of $(1 - n_L/n_H)$ to be added, which makes the location: $\frac{\Delta\lambda}{\lambda} = \frac{1}{\pi}\frac{1}{\sqrt{B}}\cdot\left(1 - \frac{n_L}{n_H}\right)$. This has been shown numerically with much accuracy, and is believed to be a result of the phase difference experienced by the various wavelengths as they travel through the stacks, however analytical verification has not been possible.

Thus, by combining the general ideas of multimirror Fabry-Perot interferometers with the impedance concept from thin films (which itself has been borrowed from microwave theory), all of the characteristics of these filters can be understood and used.

Throughout this section an analysis has been made of multi-line thin film filters. This presentation has indicated that these are interesting and useful devices, and has highlighted the positive features. A few other points should be noted. First, the spacing between the peaks is constant in frequency (not wavelength) since the resonances occur when the phase difference is an integral number of π. The second issue is more of a problem. This is seen in Figure 8.19. This figure shows how the width of each transmission peak changes with the number of layers in each stack (with the reflectivity of the mirrors). Since the number of layers in a stack can only be changed in a discrete manner (full layers are added), there is only a discrete number of passbands that can be made without changing materials. There are two ways to address this issue. The first is to choose relatively close high and low indices. This of course will require additional layers to get any desired reflectivity. Since fabrication issues make it desirable to use the least number of layers possible this is not a preferred solution. Another way to vary the width of the transmission peaks is to use three separate materials in the filter design. This makes the analysis slightly more complicated, however it has been addressed in thin film design [25], so some insight into the problem is available.

Despite these difficulties the idea of a multiline filter with the characteristics presented still seems to remain an intriguing device.

8.4   Summary

A set of coupled wave equations for calculating the output of multiplexed transmission gratings was developed which are valid for both on and off-Bragg incidence. They are comprised of two equations for each multiplexed grating. Using them it was found that the required spacing for uncoupled gratings offered by previous analyses was insufficient. Experimental verifications of the resulting angular selectivities were performed in PTR glass and in $LiNbO_3$, where it was seen that at low modulations, the overwriting of previous gratings when multiplexing in PTR glass is minimal, while it can be considerable in $LiNbO_3$, due to the real-time nature of the process.

When multiplexing reflection gratings, a narrow transmission peak will occur midway between the Bragg wavelengths of the individual gratings. By multiplexing several gratings, additional transmission peaks are obtained. Although this device has interesting possibilities as multi-line transmission filters in DWDM systems, the strong dependence of the transmission peaks on the phase of the individual gratings (which is not controllable) makes them impractical for typical applications.

An alternative method of fabricating these devices using stacks of thin films each $\lambda/4$ thick was presented. These thin film versions have the added advantages of no sidelobes and simpler scaling (adding extra stacks to make multi-line versions is relatively straightforward). The number of transmission peaks was shown to match those of multi-mirror Fabry-Perot interferometers, and expressions were given for the location and widths of the individual transmission peaks.

CHAPTER 9

THE EFFECTS OF LINEAR ABSORPTION AND

OVER-MODULATION IN VOLUME HOLOGRAMS

Throughout this work many characteristics of volume holography have been investigated, with an emphasis on their filtering abilities. The material has generally been presented by hologram type (i.e. each chapter dealt with a particular type of hologram). In this chapter two additional issues will be covered – absorption during recording and over-modulation. They are presented together because they are either not specific to any particular type of hologram, or are not encountered under typical operating conditions.

9.1   Absorption During the recording process

The first issue to be considered is the effect of subject and reference beam absorption during recording. The result of this process will be a variation in both the grating strength and the average permittivity. These effects have been addressed by many

authors [1 - 10] with the general result being the use of a modified grating strength parameter that decays with grating position. All of these works however, considered the variation in the modulation only, completely neglecting the change in the average index that also occurs. This was pointed out by Owen and Solymar [11], who gave a solution that included this effect, and who showed that it could be significant. Their solution however, was restricted to unslanted reflection holograms. In the following paragraphs a similar derivation will be presented which is applicable to all hologram types (regardless of slant angle), and which includes the effects of the varying average permittivity. It will follow Kogelnik's notation, however it will include Owen and Solymer's handling of the absorption.

As with previous derivations, we start with two plane waves incident on a medium from the same side (transmission holograms) or from opposite sides (reflection holograms). In the situation under consideration the medium is lossy which strictly speaking results in a non-sinusoidal interference pattern. In the derivation it will be assumed that the waves are incident from an index-matched medium (allowing refraction to be neglected). It is straightforward to include refraction when performing numerical calculations, thus the plots at the end of this section will do so. The recording situation is shown in figure 9.1 below.

TRANSMISSION   **REFLECTION**
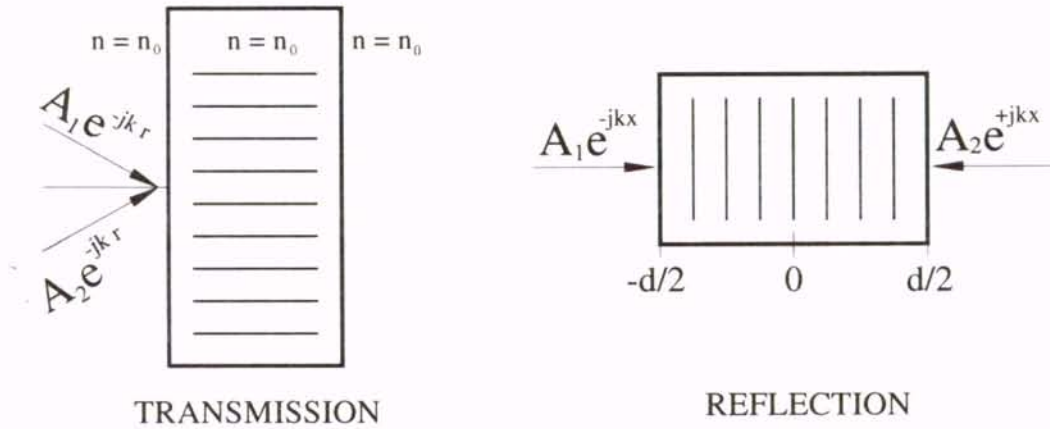
Figure 9.1. Transmission and reflection recording geometries

Transmission gratings will be considered first. From figure 9.1 we can see that within the medium the recording fields are:

$$E_1 = A_1 e^{-j\left(\vec{k}_1 \cdot \vec{r}\right)} e^{-\alpha\left(\frac{x}{\cos\theta_1}\right)} \quad \text{and} \quad E_2 = A_2 e^{-j\left(\vec{k}_2 \cdot \vec{r}\right)} e^{-\alpha\left(\frac{x}{\cos\theta_2}\right)} \quad (9.1)$$

The intensity of the interference pattern inside the medium is therefore $I = |E_1 + E_2|^2$ or:

$$|A_1|^2 e^{-2\alpha\left(\frac{x}{\cos\theta_1}\right)} + |A_2|^2 e^{-2\alpha\left(\frac{x}{\cos\theta_2}\right)} + 2|A_1||A_2| e^{-\left(\frac{\alpha}{\cos\theta_2} + \frac{\alpha}{\cos\theta_1}\right)x} \cos((k_2 - k_1)\cdot r) \quad (9.2)$$

Assuming that the material is linear, the relative permittivity after development will be proportional to this intensity ($\varepsilon_r = k*I$), thus ε_r can be written as $\varepsilon_r = \varepsilon_{r0} + \Delta\varepsilon$ or:

179

$$\varepsilon_r = \varepsilon_{r0} + \varepsilon_{r1}\, e^{-\left(\frac{\alpha}{\cos\theta_1} + \frac{\alpha}{\cos\theta_2}\right)z} \left[A\, e^{\left(\frac{\alpha}{\cos\theta_2} - \frac{\alpha}{\cos\theta_1}\right)z} + \frac{1}{A}\, e^{-\left(\frac{\alpha}{\cos\theta_2} - \frac{\alpha}{\cos\theta_1}\right)z} + \cos[(\beta_2 - \beta_1)\cdot\bar{r}] \right]$$

where : $A \equiv |A_1|/|A_2|$ is the ratio of the recording beams and $\varepsilon_{r1} \equiv 2k|A_1||A_2|$

$$(9.3)$$

This can be separated into slowly and quickly varying parts by writing $\varepsilon_r = \varepsilon_{rs} + \varepsilon_{rf}$ where:

$$\varepsilon_{rs} = \varepsilon_{r0} + \varepsilon_{r1}\, e^{-\left(\frac{\alpha}{\cos\theta_1} + \frac{\alpha}{\cos\theta_2}\right)z} \cosh\left\{ \ln(A) - \left(\frac{\alpha}{\cos\theta_1} - \frac{\alpha}{\cos\theta_2}\right)z \right\}$$

$$(9.4)$$

$$\varepsilon_{rf} = \varepsilon_{r1}\, e^{-\left(\frac{\alpha}{\cos\theta_1} + \frac{\alpha}{\cos\theta_2}\right)z} \cos[(\beta_2 - \beta_1)\cdot\bar{r}]$$

This expression for the permittivity is quite complex since the amplitude of both the sinusoidal and average portions vary with position. It must be pointed out that a plane wave will not propagate in a material with a varying average permittivity [12]. Thus the first step in finding a solution is determining the form of a wave which can propagate. The general form of this waves is assumed to be :

$$E_1 = R\, \exp\{-j[(\beta\cdot r)] + g_1(z)\} \quad \text{and} \quad E_2 = S\, \exp\{-j[(\sigma\cdot r)] + g_2(z)\} \qquad (9.5)$$

(i.e. waves that accumulate additional phase as they propagate). For these to be valid they must satisfy Maxwell's equations, thus E_1 is inserted into the wave equation with $\varepsilon_r = \varepsilon_{rs}$

and the value of $g_1(z)$ which yields a solution is found. Doing this gives the following formula for $g_1(z)$:

$$g_1(z) = \frac{-\beta\varepsilon_{r1}}{4(\varepsilon_{r0} + \varepsilon_{r1})\cos\theta} \left[\frac{e^{\ln(A)}}{\frac{2\alpha}{\cos\theta_2}} e^{-2\frac{\alpha z}{\cos\theta_1}} + \frac{e^{-\ln(A)}}{\frac{2\alpha}{\cos\theta_2}} e^{-2\frac{\alpha z}{\cos\theta_2}} + 2z \right] \tag{9.6}$$

Assuming that σ in the diffracted wave is given by $\sigma = \rho\text{-}\mathbf{K}$, the above procedure is repeated find the form of $g_2(z)$. The result is:

$$g_2(z) = g_1(z) + \delta(z)$$
$$\text{where:} \quad \delta(z) \equiv g_1(z)\left[\frac{c_r}{c_s} - 1\right] + \frac{\vartheta z}{c_s} \tag{9.7}$$

Now that we have the forms of the replay beams E_1 and E_2, we can follow Kogelnik by inserting them into the wave equation with $\varepsilon_r = \varepsilon_{rs} + \varepsilon_{rf}$. The result of this is the following rather messy equation:

$$
\begin{aligned}
&\begin{bmatrix}
R''(z) - 2j(\rho_z + g_1'(z))R'(z) - jg_1''(z)R(z) - 2\rho_z\, g_1'(z)R(z) \\[4pt]
= \left[-\beta^2\, R(z) + \beta^2 R(z) + \left(\dfrac{2\pi}{\lambda}\right)^2 \varepsilon_{r1}\left[e^{-\left(k_1^{(1)} + k_2^{(1)}\right)z} \cosh\left(\ln(A) - \Delta k^{(1)}z\right) - 1 \right] \right] e^{-j\left[(\rho\cdot r) + g_1(z)\right]} \\[4pt]
\times R(z) - g_1'^2(z)R(z) + 2k(z)\beta S(z)e^{-j\delta}
\end{bmatrix} \\[6pt]
&+\begin{bmatrix}
S''(z) - 2j(\sigma_z + g_2'(z))S'(z) - jg_2''(z)S(z) - 2\sigma_z\, g_2'(z)S(z) \\[4pt]
-\sigma^2\, S(z) + \beta^2\, S(z) + \left(\dfrac{2\pi}{\lambda}\right)^2 \varepsilon_{r1}\left[e^{-\left(k_1^{(1)} + k_2^{(1)}\right)z} \cosh\left(\ln(A) - \Delta k^{(1)}z\right) - 1 \right] e^{-j\left[(\sigma\cdot r) + g_2(z)\right]} \\[4pt]
\times S(z) - g_2'^2(z)S(z) + 2k(z)\beta R(z)e^{+j\delta}
\end{bmatrix} \\[6pt]
&+ 2k(z)\beta R(z)e^{-j\left[(\rho + K)r + g_1(z)\right]} + 2k(z)\beta S(z)e^{-j\left[(\sigma - K)r + g_2(z)\right]} = 0
\end{aligned}
$$

(9.8)

This can be simplified considerably by ignoring the last two terms (which propagate way off Bragg) and noting that g_1' and g_2' were calculated specifically to force several of the terms to zero. Additionally we note that g_1'(z) $\ll\rho_z$ and can therefore be ignored, along with the second derivatives (which is standard in coupled wave analysis). Using these approximations reduces equation 9.8 to the more manageable:

$$
\begin{aligned}
\nabla^2 E(z) + k^2 E(z) &= \left[\frac{d^2 R(z)}{dz^2} - 2j\left(\rho_z + \frac{dg_1(z)}{dz}\right)\frac{dR(z)}{dz} + 2\kappa(z)\beta S(z)e^{-j\delta(z)} \right] e^{-j\left[(\rho\cdot r + g_1(z))\right]} \\[6pt]
&+ \left[\frac{d^2 S(z)}{dz^2} - 2j\left(\sigma_z + \frac{dg_2(z)}{dz}\right)\frac{dS(z)}{dz} + 2\kappa(z)\beta R(z)e^{-j\delta(z)} \right] e^{-j\left[(\rho\cdot r + g_2(z))\right]} = 0
\end{aligned}
$$

(9.9)

Setting each portion to zero and solving for the coupled wave equations yields:

$$c_R \frac{\partial R(z)}{\partial z} = -j\kappa(z)S(z)e^{-j\delta(z)} \quad \text{and} \quad c_S \frac{\partial S(z)}{\partial z} = -j\kappa(z)R(z)e^{j\delta(z)} \qquad (9.10)$$

where c_R and c_S are the same as in standard coupled wave theory, $\delta(z)$ is defined in equation 9.7, and $\kappa(z)$ is given by the following function:

$$\kappa(z) = \frac{1}{4}\left(\frac{2\pi}{\lambda}\frac{\varepsilon_{r1}}{\sqrt{\varepsilon_{r0}+\varepsilon_{r1}}}\right)e^{-\left(\frac{\alpha}{\cos\theta_1}+\frac{\alpha}{\cos\theta_2}\right)} \qquad (9.11)$$

Other than κ being a function of position, the only difference between equation 9.10 and the typical coupled wave equations is the additional phase term $\delta(z)$. By combining these two first order coupled wave equations into a single second order one, we can see the effect of the variation in the DC term. The result of this combination is:

$$S''(z) + j\left[\frac{\partial g_1(z)}{\partial z}\left(1-\frac{c_S}{c_R}\right) - \frac{\vartheta}{c_S}\right]S'(z) + \frac{\kappa(z)^2}{c_R c_S}S = 0 \qquad (9.12)$$

Recall that this equation was derived for the transmission case. Similar results are obtained for reflection holograms by replacing E_2 in equation 9.1 with:

$$E_2 = A_2 e^{-j\left(\vec{k}_2\cdot\vec{r}_2\right)}e^{\alpha\left(\frac{z-d}{\cos\theta_2}\right)} \qquad (9.13)$$

The algebra involved when this substitution is made is the same as in the transmission case, so it will not be repeated. The result of the substitution is that g_1 becomes:

$$g_1(z) = \frac{-\beta\varepsilon_{r1}}{4(\varepsilon_{r0} + \varepsilon_{r1})\cos\theta} \left[\frac{e^{\left(\ln(A) + \frac{\alpha d}{\cos\theta_2} - 2\frac{\alpha z}{\cos\theta_1}\right)}}{\left(\frac{2\alpha}{\cos\theta_1}\right)} + \frac{e^{\left(-\ln(A) - \frac{\alpha d}{\cos\theta_2} + 2\frac{\alpha z}{\cos\theta_1}\right)}}{\left(\frac{2\alpha}{\cos\theta_2}\right)} + 2z \right]$$

$$(9.14)$$

No additional changes are required in the derivation. Thus, when dealing with transmission gratings (slanted or unslanted) equation 9.6 is used in equation 9.12, while equation 9.14 is used for all reflection gratings.

Several things should be noted from equation 9.10. First, the original dephasing parameter ($\upsilon d/2c_S$) has been replaced by one which is itself a function of the additional phase parameter g. Furthermore the concept of being on Bragg at a particular angle no longer makes sense since "Bragg" changes as we go through the material (it is a function of z). This was addressed in [11] by replacing the dephasing parameter with its average value. This seems reasonable (for small variations in g(z)), so the result becomes:

$$\xi'' = \frac{1}{d}\int_0^d g_1'(z)\left(1 - \frac{c_R}{c_S}\right) - \frac{\vartheta}{c_S}\,dz = [g_1(d) - g_1(0)]\left(\frac{c_S - c_R}{2c_S}\right) - \frac{\vartheta d}{2c_S} \qquad (9.15)$$

184

It is clear from these equations that in the case of unslanted transmission ($c_S=c_R$) there is no additional effect of the varying DC term. For an unslanted reflection grating ($c_R = -c_S$), the additional dephasing is $2g'_1(z)$. Between these two cases the effect increases, but the results are complicated because $g'_1(z)$ is a function of c_R and ε_{r1}.

In traditional coupled wave analysis the equations were solved (using standard methods) to obtain the results. Unfortunately however, no closed form solutions exist for equation 9.10; thus solutions must be found numerically (by way of a fourth order Runge-Kutta numerical integration). This has been done, and the results are shown in figures 9.2 through 9.5 below. In these figures, plots are given for gratings ranging from pure transmission ($\Phi=\pi$) through pure reflection ($\Phi=0°$). Situations are covered where the absorption coefficient causes the power to drop by 0, 50% and 75%, and include several values of Δn and A (the ratio of the recording beam intensities). For comparison purposes, models calculating absorption in the usual manner (ignoring average permittivity variation) are also included (except in the unslanted transmission case where the derivation showed that there is no additional effect of the DC variation).

Figure 9.2. Angular Selectivity vs. dephasing for $\Phi = \pi/2$ and various values of A. $\kappa d = 0.5$ (top); $\kappa d = \pi/2$ (bottom)

Figure 9.3. Angular Selectivity vs. dephasing for $\Phi = \pi/3$ and various values of A. $\kappa d = 0.5$ (top); $\kappa d = \pi/2$ (bottom)
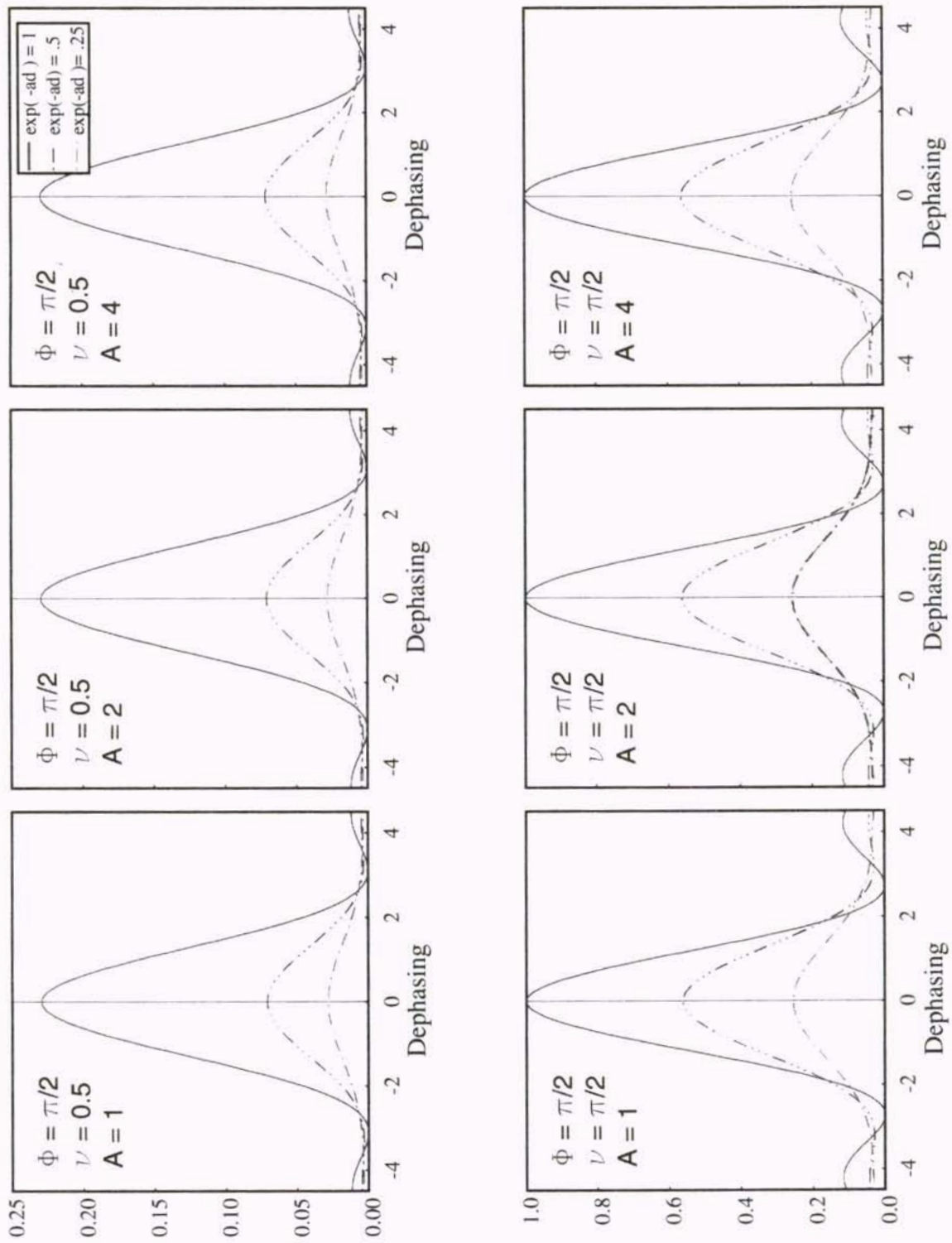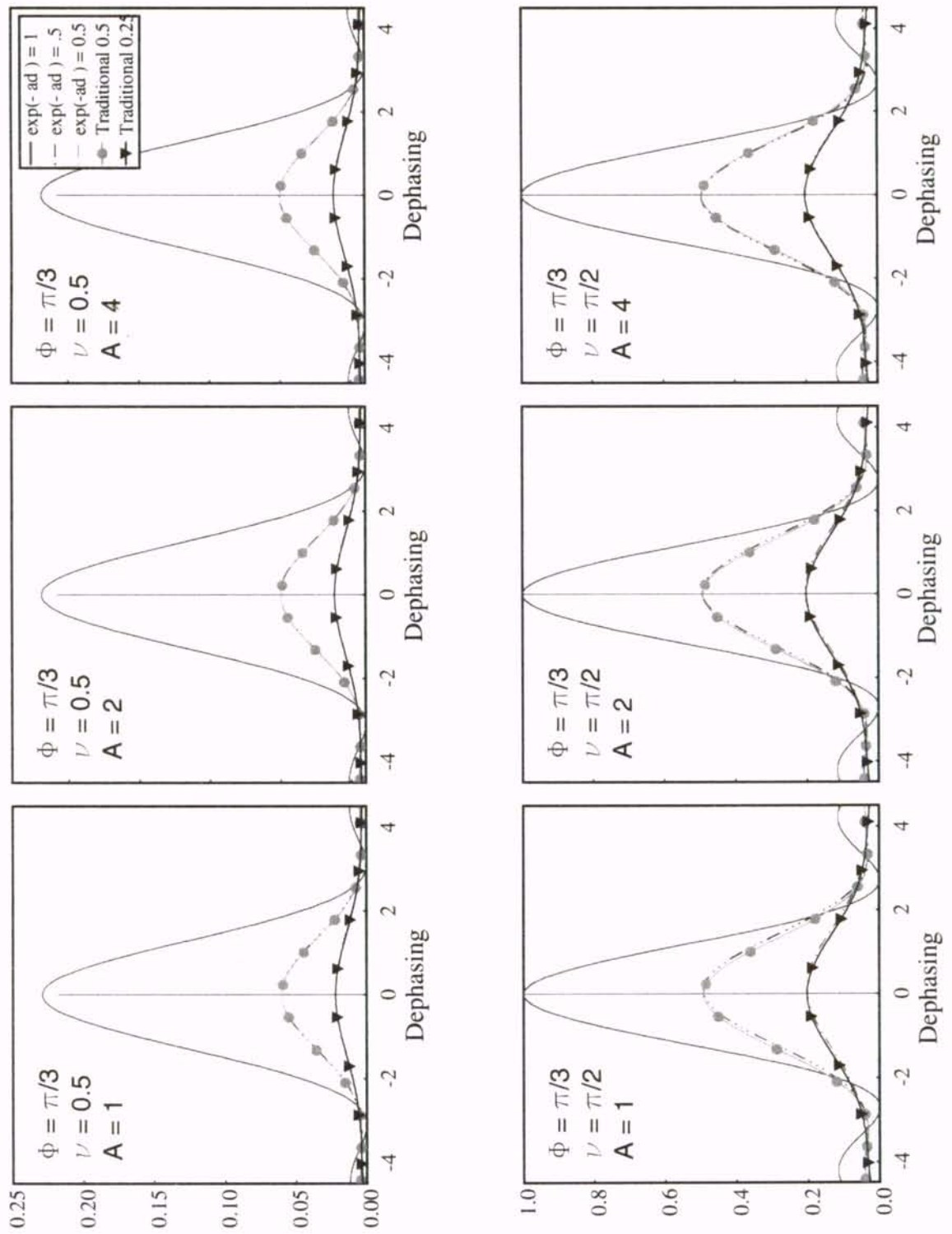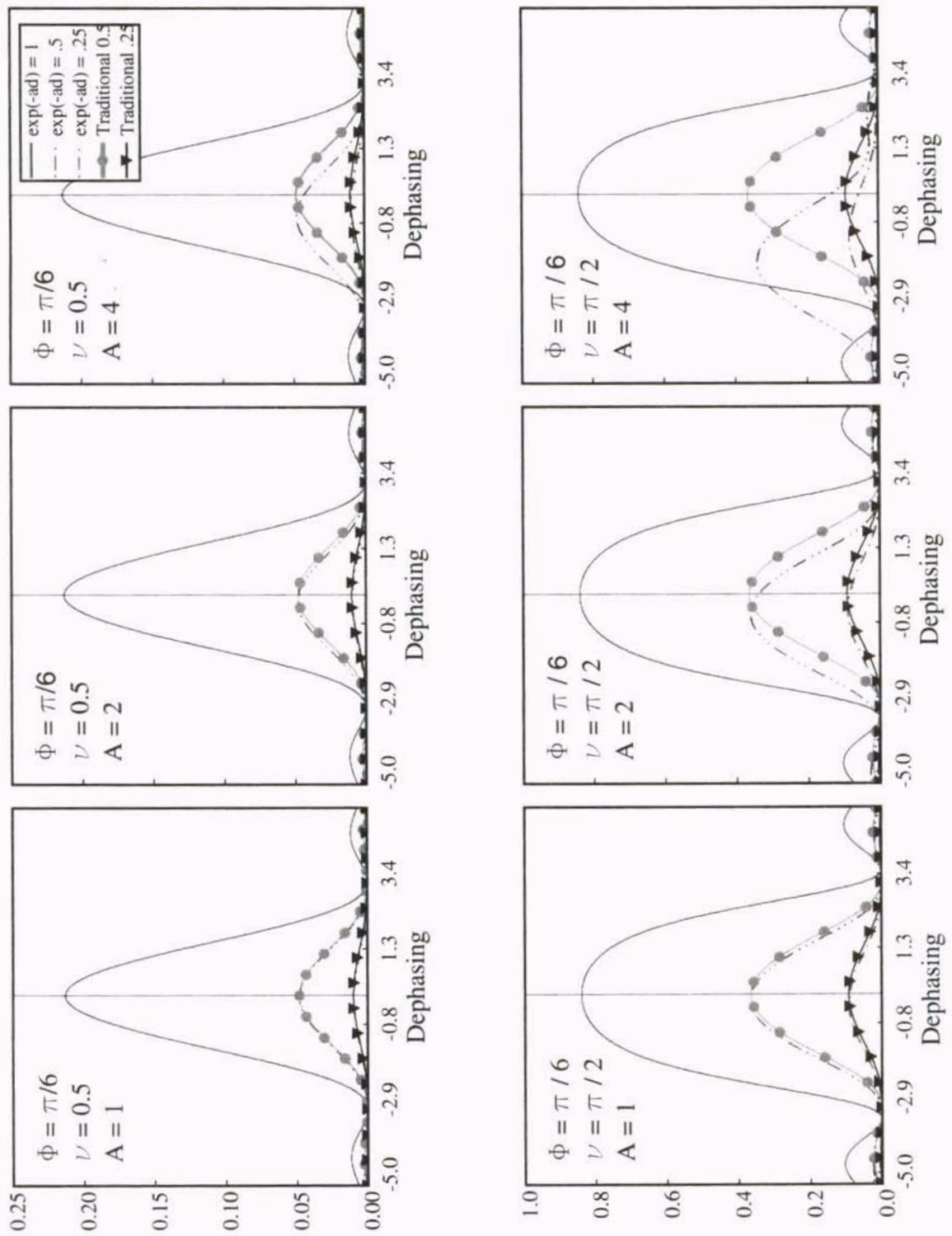
187

Figure 9.4. Angular Selectivity vs. dephasing for $\Phi=\pi/6$ and various values of A. $\kappa d = 0.5$ (top); $\kappa d = \pi/2$ (bottom)
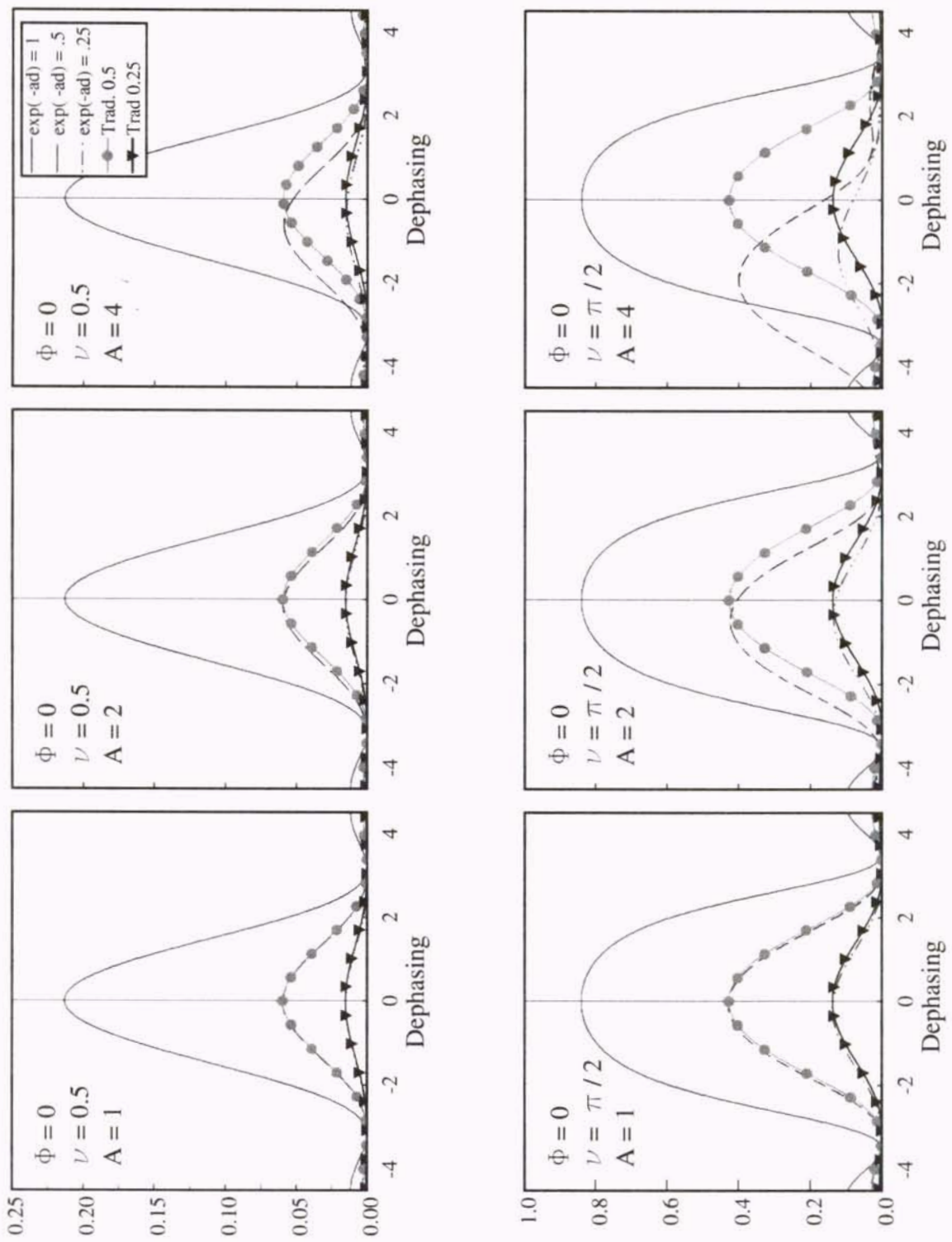
188

Figure 9.5. Angular Selectivity vs. dephasing for $\Phi=0$ and various values of A.
$\kappa d = 0.5$ (top); $\kappa d = \pi/2$ (bottom)

189

In discussing the characteristics of these gratings we first notice the obvious fact that as the absorption during recording is increased, the overall diffraction efficiency (both on and off Bragg) during replay is decreased. This occurs because as the recording beams are attenuated with propagation in the medium, the strength of the grating decreases accordingly.

Much more interesting than this is the fact that several differences exist in the diffraction properties when the variation in the average permittivity was included from those where it was ignored. These differences increase as the grating type changes from an unslanted transmission to an unslanted reflection. Additionally, they increase as both the grating strength (ν) and asymmetry factor (A) increase. With a slant angle of 60 degrees for example, there is almost no difference when the D.C. variation is included in the analysis – even with grating strengths as high as $\pi/2$. At 30 degrees however, the effect of this variation has become obvious in almost all cases. These effects manifest themselves in two ways: the peak value of the diffraction efficiency is shifted toward the negative end of the dephasing (due to the variation in "Bragg" described above), and the shape of the angular selectivity curve changes.

This change in shape causes the angular selectivity to become asymmetric with dephasing (in all cases where the variation in average permittivity was ignored the resulting angular selectivity curve was symmetric). This was also the case in all other gratings considered in this work (even those which included non-uniform recording beams). Thus having a material where the average value of the permittivity varies represents a unique and interesting situation.

9.2   Over Modulated Transmission Gratings

In an earlier chapter the characteristics of transmission gratings were investigated with the emphasis placed on their filtering characteristics. The chief means of examining these characteristics has been to look at the dependency of the diffraction efficiency on some sort of dephasing (either angular or wavelength). Although the actual shape of these characteristics differed somewhat, they all followed the general trend of being maximum on-Bragg, dropping as they deviated from Bragg, and increasing again slightly at a sidelobe.

It turns out however that gratings with large modulation values (higher than about 0.8π) can experience significantly different behavior. This behavior is not typically seen since the modulation strength is usually kept below 0.5π where an incident plane wave would have 100% diffraction efficiency. The behavior of these over-modulated gratings is important however, as their characteristics are distinct enough to allow the possibility of interesting new devices. In this section we present the characteristics of these gratings.

The behavior of transmission gratings at all strengths can be determined from equation 2.21, which is repeated below.

$$\eta = \frac{v^2}{\xi^2 + v^2} \sin^2\left(\sqrt{v^2 + \xi^2}\right) \qquad (9.16)$$

When the grating is operated on-Bragg this simplifies to $\eta = \sin^2(v)$. In typical situations ($v < 0.5\pi$) deviation from Bragg incidence causes the first term in equation 9.16 to drop while the second one rises. The magnitude of this drop is much higher than the rise however, which results in the grating having the behavior discussed in chapter 4 (a drop as we deviate from Bragg, followed by a slight increase at a sidelobe. As the value of v increases however, the change in the first term of equation 9.16 will be lower for any given amount of dephasing, and the effects of the sine term may dominate.

In these regions one of three types of behavior can be seen, first the grating can act as it does for low modulation strengths. This occurs when the variation in the sin term is minimal with additional dephasing (near $\pm\pi/2$). The second type of behavior that can occur is that the grating starts out acting in the traditional manner (the diffracted power decreases as we deviate from Bragg), however at a side lobe the power is higher than its on-Bragg value. The final type of behavior is when the diffracted power increase immediately as we deviate from Bragg. The choice of which behavior is seen is determined by the modulation strength, and varies cyclically.

We will determine the value of the grating strength where each region begins separately, starting with region 3. On one side of these points the grating increases with ξ while on the other side it decreases with ξ, thus they can easily be found by setting the derivitive of equation 9.16 to zero in the limit as ξ approaches zero. This derivative is given by:

$$\frac{d\eta}{d\xi} = \frac{v^2 \left[\sqrt{v^2 + \xi^2} \, \sin\left(2\sqrt{v^2 + \xi^2}\right) - 2\sin^2\left(\sqrt{v^2 + \xi^2}\right) \right]}{\left(\xi^2 + v^2\right)^2} \qquad (9.17)$$

This becomes zero in the limit as ξ approaches 0 at $v = m\pi$ and at $v \approx m\pi/2$ for arbitrary integers m. Taking the second derivative shows that the values around $m\pi/2$ are merely inflection points, and it is the values at $m\pi$ which are the minima and maxima that we are seeking. Thus at grating strengths of $m\pi$ where m is an integer the diffraction efficiency will decrease as the incidence deviates from Bragg.

In order to find the second type of behavior (when the diffraction efficiency falls as we deviate from Bragg incidence, but then exceeds the Bragg level at a sidelobe), we first note that sidelobes occur at $\sqrt{v^2 + \xi^2} \approx \dfrac{m\pi}{2}$ where m is any odd integer greater than one. The sidelobe with the highest efficiency (the first one) occurs with the lowest value of m such that $\dfrac{m\pi}{2} > v$. Since this efficiency is given by $\left(1 + \xi^2/v^2\right)^{-1}$ (see equation 9.16), we can combine the last two equations to give the following value:

$$\eta_{sidelobe} = \frac{4v^2}{m^2 \pi^2} \qquad (9.18)$$

Region 2 areas are found by solving for grating strengths when this value exceeds that found on-Bragg, i.e.:

$$\frac{4v^2}{m^2\pi^2} > \sin^2(v) \quad \text{or} \quad \frac{2v}{m\pi} > |\sin(v)| \quad \text{or} \quad v\csc(v) > \frac{m\pi}{2} \tag{9.19}$$

Each of these regions are shown below in figure 9.6. In this figure each pattern denotes a different type of behavior, and the pattern repeats every π.
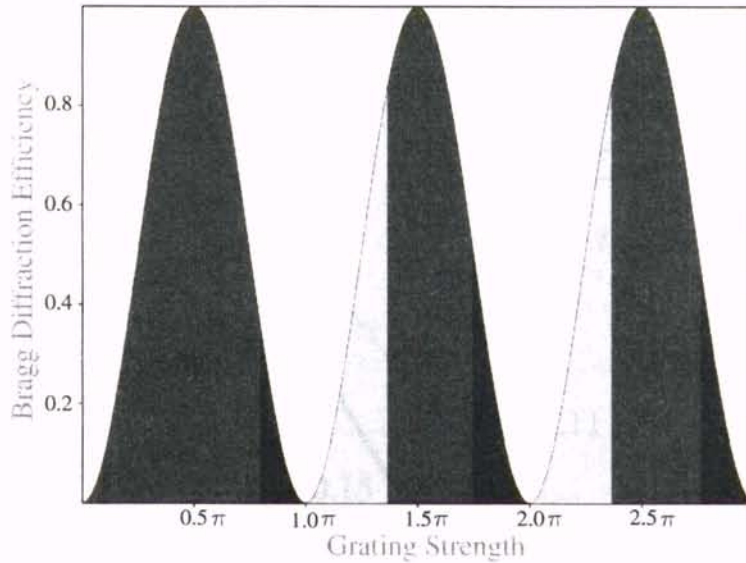
Figure 9.6. Regions of transmission grating behavior

Figures 9.7, 9.8 and 9.9 below show in detail how the grating acts in each of the regions for incidence by plane waves as well as with various Gaussian beams. The grating strengths were chosen such that the on-Bragg diffraction efficiency of an incident plane in all figures is constant at 15%. In figure 9.7 we see the behavior in region 1 (which is the typical region where these devices are operated). Here the grating acts as described above, starting with a maximum diffraction efficiency on-Bragg that decreases

as our incidence deviates from this. The geometry factors (see chapter 4) used in these analyses are 1 and 3. Essentially the behavior is as described throughout chapter 4 where it was shown that as the width of the replay beam was decreased, the on-Bragg diffraction efficiency decreased also, and the off-Bragg sidelobes were washed out due to the presence of the beam's finite angular spectrum.
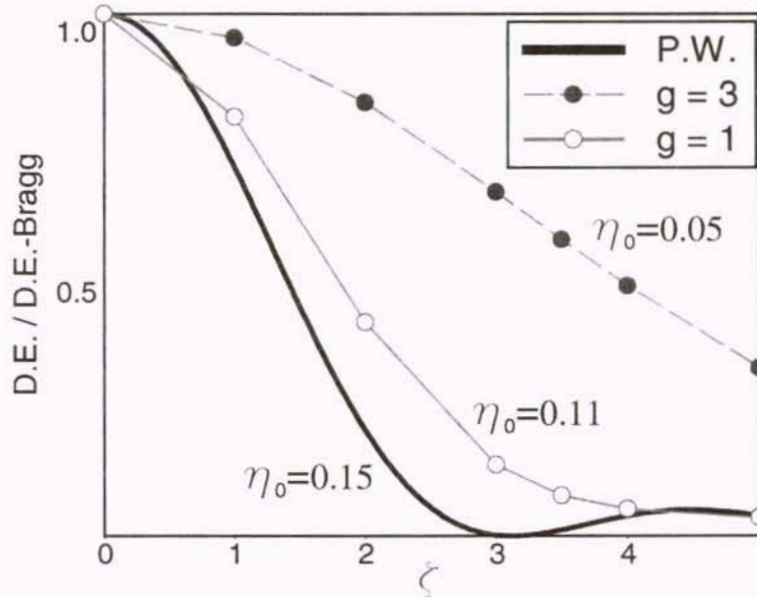
Figure 9.7. Behavior of a transmission grating in region 1

When the grating is operated in the second region however, its behavior is considerably different. In this case the width of the replay beam completely determines the behavior of the device, both on and off-Bragg. This is shown in figure 9.8 below, where we see that an incident plane wave exhibits behavior that matches the description given for this region above.
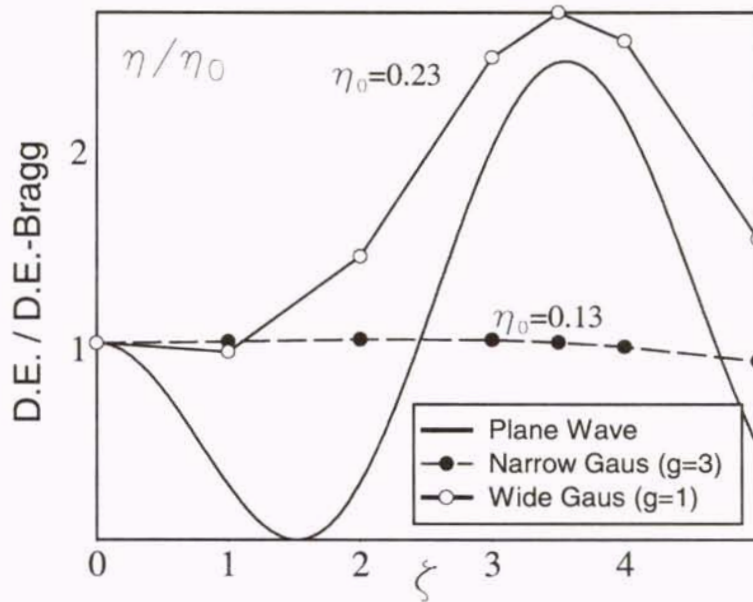
Figure 9.8. Behavior of a transmission grating in region 2

In this figure when the incident wave is a narrow Gaussian beam the diffraction efficiency remains substantially constant across the entire range of dephasing (there is no filtering present). As the beam is widened however, the diffraction efficiency increases almost immediately –acting as a plane wave would in the third region. The other important effect seen is that the on-Bragg diffraction efficiency exhibits behavior just as strange. In this case the diffraction efficiency of a wide gaussian beam is higher than that of the plane wave, while that of a narrow gaussian beam is lower than the plane wave.

Each of these effects occurs because the narrow beam has a wide enough angular spectrum for the outer components to exceed the sidelobe, while those of the wide beam just reach the sidelobe. As mentioned above the sidelobe power in these regions is higher than the Bragg value, which makes the overall diffraction efficiency of the wide Gaussian

beam higher than that of a plane wave, while that of the narrow beam is slightly lower than the plane wave.

The behavior in the final region is shown in figure 9.9, and again we see that the width of the replay beam determines how the grating will operate. In this case the grating does not exhibit good filtering characteristics for any of the beams since they all either increase with dephasing or have only a minimal drop. In this region the width of the gaussian incident beam determines whether the power will increase off-Bragg, or whether it will decrease right away (albeit in a very small amount). Both gaussian beams have a higher Bragg diffraction efficiency than the plane wave, with the narrower beam having the highest.
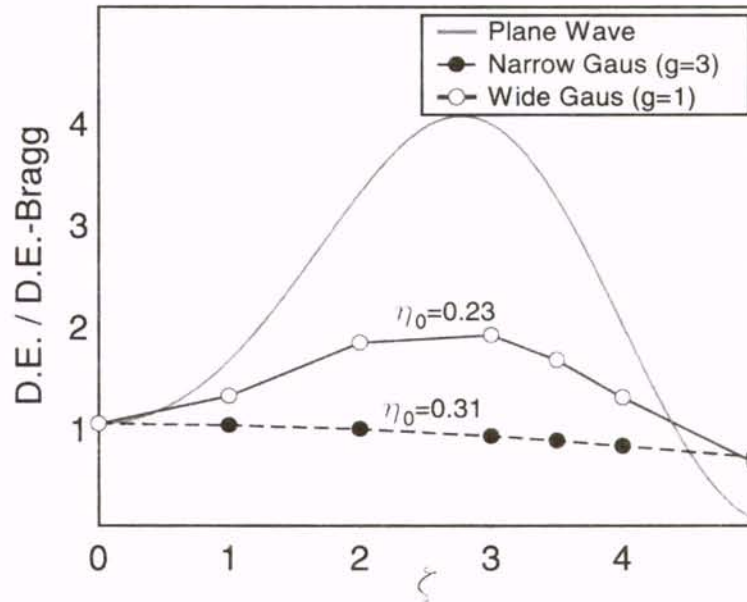
Figure 9.9. Behavior of a transmission grating in region 3

197

As before, this behavior can be explained by considering the width of its angular spectrum along with the plane wave response. That is, since the diffracted power increases with dephasing, and since the angular spectrum can be considered as dephasing (see equation 2.30), then the narrowest beams (which have the widest angular spectrum) will have a higher efficiency.

A problem occurs however, if the region in which we are experimentally operating is not exactly known. For example if when measuring the angular selectivity curve of an unknown grating the efficiency initially increases before decreasing well off-Bragg, without knowing the exact region in which we are operating, there is no way of determining whether this effect comes from being in region 3 with an extremely wide beam, or being in region two with a medium sized beam. The Bragg diffraction efficiency would give some excellent clues as to which region we are operating in, however the actual Bragg efficiency is not known (only the maximum efficiency). One method of determining the operating point is to first thoroughly characterize the replay beam, then to take accurate measurements of the spacing between the two peaks (they will be symmetric across the Bragg angle). Then by measuring the diffraction efficiency half way between the peaks one can get the actual value of the on-Bragg efficiency which should help determine where we are operating. If the position is not known after this procedure, repeating it with a different replay beam should determine it.

Since much of the emphasis of this dissertation has been on the profiles of diffracted beams, a discussion of these profiles in over-modulated transmission gratings seems to be in order. This will be limited to gratings with modulations high enough to

exhibit the third type of behavior seen above (increase in off-Bragg efficiency). Again values of one and three will be used for the geometry factor of the replay beam. Figure 9.10 shows the results of the relatively wide replay beam (g=1).
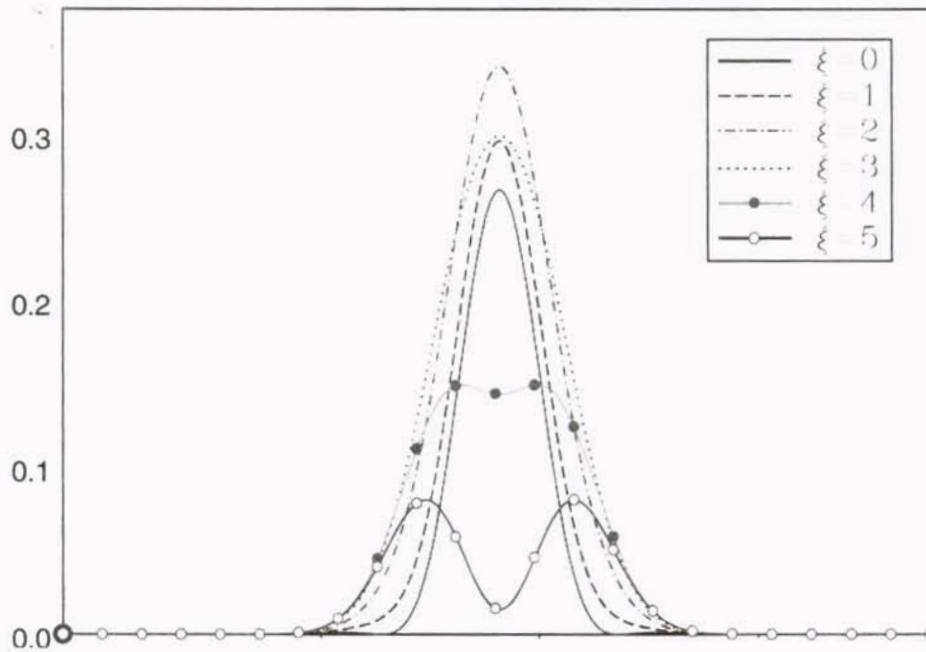
Figure 9.10. On and off-Bragg profiles of gratings in region 3 (g=1)

From the plot we see that when the wave is incident on-Bragg, the diffracted beam appears to be approximately Gaussian in shape, and as we go off Bragg the amplitude initially grows (although this time it keeps its shape unlike in the earlier transmission gratings), but finally starts to drop in efficiency (again keeping a reasonably nice shape).

In figure 9.11 below beam profiles are given when the same grating is illuminated with a much narrower replay beam. In this case the beam profile is distorted even on-Bragg, and this distortion remains through all values of the dephasing parameter.
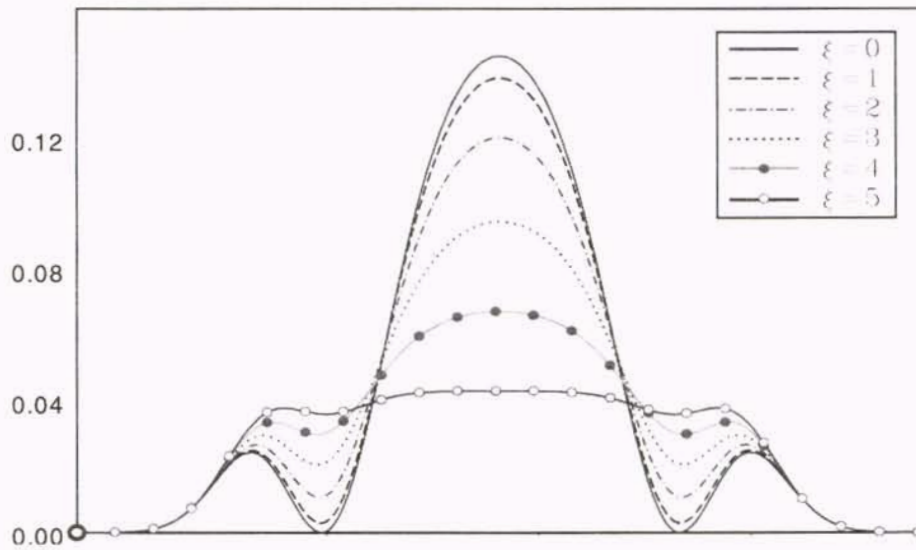
Figure 9.11. On and off-Bragg profiles of gratings in region 3 (g=3)

Since transmission gratings are not typically operated in these regions, these non-traditional off-Bragg effects are not well known. It may turn out however that these areas could be used to improve on off-Bragg diffracted profiles or efficiencies, or even to take advantage of the symmetry of the side lobes to make dual line filters.

9.3   Summary

The effect on the angular selectivity of the induced variation in average index which occurs when gratings are recorded in the presence of absorption was investigated for all hologram types. It was seen that no additional effects occur for un-slanted transmission gratings, however effects are present in all other grating types. These

include an asymmetric angular selectivity as well as a shift in the location of the maximum. Generally these effects increase as the grating type changes from unslanted transmission to unslanted reflection.

Finally, the characteristics of plane waves incident on over-modulated transmission gratings was investigated. It was seen that they can experience one of three types of off-Bragg behavior. These include the traditional behavior, an immediate increase in off-Bragg amplitude, or an immediate drop followed by a latter increase in this amplitude to an overall off-Bragg maximum. These effects make the response of the gratings (both in terms of angular selectivity and beam profile) strongly dependent on the particular replay beam chosen, and results in a difficulty in characterizing the gratings.

CHAPTER 10

CONCLUSIONS

The purpose of this work was to design and analyze volume holographic optical elements. Toward this end transmission, reflection and overlap gratings were studied. These analyses used several models including standard coupled wave analysis, two-dimensional coupled wave analysis and plane wave decomposition. A Gaussian amplitude was used for the beam profile studies because this approximates the distribution of a standard laser.

High efficiency transmission gratings were recorded in photo-thermo-refractive glass and analyzed with various replay beams. Several differences were found between the measured angular selectivity and that predicted by coupled wave analysis. These were shown to result from the angular spectrum of the replay beam.

In order to investigate these finite beam effects further, a closed form solution for the off-Bragg diffracted and transmitted beam profiles of a Gaussian replay beam was derived. Using this solution it was shown that additional distortion (above that which happens on-Bragg) occurs when the devices are operated off-Bragg. This distortion (both on and off-Bragg) arises due to non-uniformity in the attenuation of the angular spectrum

components. It was shown that when $\sigma > 1.6\lambda d/nL$ (or equivalently $g < 0.32$) the effects of the beam profile on the angular selectivity are minimal, and near plane-wave responses are found. All profile predictions were experimentally verified.

The diffraction efficiency and resulting profiles of finite beams diffracted from uniform reflection gratings on and off-Bragg were investigated. With regard to the diffraction efficiency the effect of replaying with a finite beam will be an overall decrease in the level. For a typical He-Ne laser beam this decrease will be about 5% when the grating strength is near $\pi/2$.

The profiles diffracted from these reflection gratings will generally be distorted versions of the replay beams, with the distortion increasing as incidence deviates from Bragg. Unlike in transmission gratings however, the fidelity of the on-Bragg profiles improve as the grating strength is increased. This happens because stronger reflection gratings have wider regions of approximately uniform attenuation. Again, all profile predictions were verified experimentally.

The profiles of the diffracted and transmitted beams of overlap gratings were investigated using known solutions. When the grating is replayed with the exact recording reference beam the resulting diffracted beams will be shifted and distorted versions of the recording subject beam. These effects arise because the first Born approximation is violated (substantial beam depletion occurs). At overall efficiencies of up to 50% they are minimal, however they become pronounced at 75%. When the overall efficiency exceeds 83% the diffracted beam splits into more than one section. This distortion and shift can be problematic in data storage systems.

A method was developed for obtaining a diffracted beam from an overlap hologram with any desired profile. This method is based on using an amplitude mask over the subject beam to vary the grating strength with position. The required amplitude masks are simple (realizable) when the desired profile is close to that of the recording subject beam, thus this should be used when possible. It was shown that the method can be used to correct for most of the beam depletion effects discussed previously, however the splitting of the beam at 83% can still not be alleviated, thus it represents the true limit at which a good reproduction of the recording beam can be obtained.

A set of coupled wave equations for calculating the output of multiplexed transmission gratings was developed which are valid for both on and off-Bragg incidence. They are comprised of two equations for each multiplexed grating. Using them it was found that the required spacing for uncoupled gratings offered by previous analyses was insufficient. Experimental verifications of the resulting angular selectivities were performed in PTR glass and in $LiNbO_3$, where it was seen that at low modulations, the overwriting of previous gratings when multiplexing in PTR glass is minimal, while it can be considerable in $LiNbO_3$, due to the real-time nature of the process.

When multiplexing reflection gratings, a narrow transmission peak will occur midway between the Bragg wavelengths of the individual gratings. By multiplexing several gratings, additional transmission peaks are obtained. Although this device has interesting possibilities as multi-line transmission filters in DWDM systems, the strong dependence of the transmission peaks on the phase of the individual gratings (which is not controllable) makes them impractical for typical applications.

An alternative method of fabricating these devices using stacks of thin films each $\lambda/4$ thick was presented. These thin film versions have the added advantages of no sidelobes and simpler scaling (adding extra stacks to make multi-line versions is relatively straightforward). The number of transmission peaks was shown to match those of multi-mirror Fabry-Perot interferometers, and expressions were given for the location and widths of the individual transmission peaks.

The effect on the angular selectivity of the induced variation in average index which occurs when gratings are recorded in the presence of absorption was investigated for all hologram types. It was seen that no additional effects occur for un-slanted transmission gratings, however effects are present in all other grating types. These include an asymmetric angular selectivity as well as a shift in the location of the maximum. Generally these effects increase as the grating type changes from unslanted transmission to unslanted reflection.

Finally, the characteristics of plane waves incident on over-modulated transmission gratings was investigated. It was seen that they can experience one of three types of off-Bragg behavior. These include the traditional behavior, an immediate increase in off-Bragg amplitude, or an immediate drop followed by a latter increase in this amplitude to an overall off-Bragg maximum. These effects make the response of the gratings (both in terms of angular selectivity and beam profile) strongly dependent on the particular replay beam chosen, and results in a difficulty in characterizing the gratings.

APPENDIX A

RIEMANN'S SOLUTION OF 2$^{\text{nd}}$ ORDER LINEAR

HYPERBOLIC DIFFERENTIAL EQUATIONS

APPENDIX A

RIEMANN'S SOLUTION OF 2nd ORDER LINEAR

HYPERBOLIC DIFFERENTIAL EQUATIONS

Chapter 6 of this work required the solution of a 2nd order, linear, hyperbolic differential equation. The solution given was obtained by inserting boundary conditions into a known solution form. At this time the derivation of that form will be given, based on a description by Courant and Hilbert [1].

A.1   Riemann's Method of Solution

Any second order linear hyperbolic differential equation can be written in terms of an operator L as:

$$L[u] = u_{rs} + au_r + bu_s + cu = f \tag{A.1}$$

where a, b, c and f are functions of the independent variables r and s (the principle coordinate axes), u_r and u_s represent the derivative of u with respect to the variables r and

s respectively, and u_{rs} is the second partial derivative with respect to r and s. The solution to this equation is an expression for u in terms of the function f and the initial conditions at some point P. These initial conditions are assumed known along some curve c. The coordinate system and these points are shown in Figure A.1.
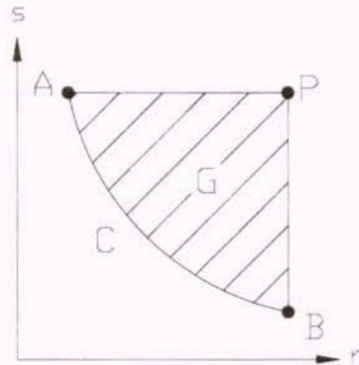
Figure A.1. Coordinate system used for the problem

To begin the solution it is noted that by definition the adjoint of an operator L (written as L*) has the requirement that $vL[u] - uL^*[v]$ must be a divergence expression. For equation A.1 this adjoint is given by:

$$L^*[v] = v_{rs} - (av)_r - (bv)_s + cv \qquad (A.2)$$

Therefore:

$$vL[u] - uL^*[v] = vu_{rs} - uv_{rs} + (uav)_r + (ubv)_s \qquad (A.3)$$

which can be written in one of two forms:

$$\begin{aligned} vL[u] - uL^*[v] &= (vu_r + buv)_s - (uv_s - auv)_r &\quad (a) \\ &= (vu_s + auv)_r - (uv_r - buv)_s &\quad (b) \end{aligned} \qquad (A.4)$$

As noted above, these are divergence expressions. To solve them we need to integrate over the area G. In doing so we will make use of Green's Theorem:

$$\int_\Gamma (A\,dx + B\,dy) = \iint_G \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) dx\,dy \qquad (A.5)$$

where Γ denotes the perimeter of the surface R. Using A.5 along with A.4 yields:

$$\iint_G (v\,L[u] - u\,L^*[v])\,dr\,ds = \int_\Gamma (v\,u_r + b\,u\,v)\,dr + (u\,v_s - a\,u\,v)\,ds \qquad (a)$$

$$= \int_\Gamma (v\,u_s + a\,u\,v)\,ds + (u\,v_r - b\,u\,v)\,dr \qquad (b)$$

$$(A.6)$$

Form1: We rewrite equation A.6a as:

$$\iint_G (v\,L[u] - u\,L^*[v])\,dr\,ds = \int_{AB} (v\,u_r + b\,uv)\,dr + (u\,v_s - auv)\,ds$$

$$+ \int_{BP} (u\,v_s - auv)\,ds - \int_{AP} (v\,u_r + b\,uv)\,dr \qquad (A.7)$$

But:

$$\int_{AP} (v\,u_r + b\,uv)\,dr = u(P)\,v(P) - u(a)\,v(a) - \int_{AP} u(v_r - bv)\,dr \qquad (A.8)$$

Thus we can write equation A.7 as:

$$u(P)\,v(P) = u(a)\,v(a) + \int_{AP} u(v_r - bv)\,dr + \int_{BP} u(v_s - av)\,ds$$

$$+ \int_{AB} v(u_r + bu)\,dr + u(vs - av)\,ds + \int_G (vf - u\,L^*[v])\,dr\,ds \qquad (A.9)$$

209

This expression can be simplified considerably if the function v in the equation satisfies the following criteria:

$$L_{rs}^{*}[R]=0 \qquad R_r = bR \text{ on } AP \qquad R_s = aR \text{ on } BP \qquad R[P]=1 \qquad \text{(A.10)}$$

If these requirements are met the function is called a Riemann function and will be denoted as R. If this holds equation A.9 simplifies to:

$$u(P)=u(a)v(a)+\int_{A}^{B}[R(u_r +bu)dr +u(R_s -aR)ds]+\iint_{G} Rf\,dr\,ds \qquad \text{(A.11)}$$

In a similar manner, had we started with equation A6b, we would get form 2, ie:

$$u(P)=u(B)v(B)-\int_{A}^{B}[R(u_s +au)ds +u(R_r -bR)dr]-\iint_{G} Rf\,dr\,ds \qquad \text{(A.12)}$$

Note that if the initial curve (A-B) is a right angle given by r=α, s=β (fig. A.2):
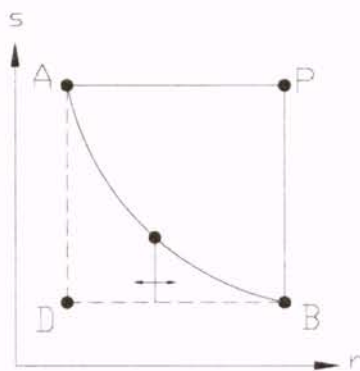
Figure A.2. Coordinate system when the initial curve degenerates into a right angle

then equations A.11 and A.12 become:

$$u(P) = u(A)R(A) - \int_D^A u(R_s - ar)\,ds + \int_D^B R(u_r + bu)\,dr + \iint_G Rf\,dr\,ds \qquad (a)$$

$$u(P) = u(B)R(B) + \int_D^A R(u_s + au)\,ds - \int_D^B u(R_r - bR)\,dr - \iint_G Rf\,dr\,ds \qquad (b)$$

(A.13)

Finally note that for most of the problems considered here a = b = f = 0, thus:

$$u(P) = u(A) - \int_D^A u\,R_s\,ds + \int_D^B R u_r\,dr \qquad (a)$$

$$u(P) = u(B) + \int_D^A R u_s\,ds - \int_D^B u\,R_r\,dr \qquad (b)$$

(A.14)

Since each of these forms is equally valid, the choice of which to use comes from the boundary conditions available from the problem.

APPENDIX B

SPECIAL FOURIER TRANSFORM PAIRS

APPENDIX B

SPECIAL FOURIER TRANSFORM PAIRS

In chapter four, a derivation was given which showed the equivalence of plane wave decomposition and 2-D coupled wave analysis in certain situations. This required (in equation 4.25) the use of three complex Fourier transform pairs. These transforms are derived below.

B.1   Development of the Fourier transform pairs

The following three inverse fourier transforms are needed:

$$\Im^{-1}\left[\Theta(\sigma)G(\sigma+\overline{\sigma}_0)\right] \qquad \Im^{-1}\left[\Theta(\sigma)\frac{\partial G(\sigma+\overline{\sigma}_0)}{\partial P}\right] \qquad \Im^{-1}\left[\Theta(\sigma)(\sigma+\overline{\sigma}_0)G(\sigma+\overline{\sigma}_0)\right]$$

These are found from:

$$\Im^{-1}\left[G(\sigma)\right] \qquad \text{(a)}$$
$$\Im^{-1}\left[\frac{\partial G(\sigma)}{\partial P}\right] \qquad \text{(b)} \qquad\qquad\qquad \text{(B.1a,b,c)}$$
$$\Im^{-1}\left[j\sigma G(\sigma)\right] \qquad \text{(c)}$$

along with combinations of the following well known fourier transform theorems:

<p style="text-align:center">shifting theorem: $\mathfrak{I}^{-1}[F(\sigma-\sigma_0)]=e^{-j\sigma x}f(x)$</p>

<p style="text-align:center">convolution theorem : $\mathfrak{I}^{-1}[F(\sigma)H(\sigma)]=\dfrac{1}{2\pi}\int\limits_{-\infty}^{\infty}f(\alpha)h(x-\alpha)d\alpha=f(x)*h(x)$</p>

<p style="text-align:center">derivative theorem : $\mathfrak{I}^{-1}[-j\sigma F(\sigma)]=\dfrac{df(x)}{dx}$</p>

The next step is to find expressions for B.1a, B.1b and B.1c. First we do B.1a

$$\mathfrak{I}^{-1}[G(\sigma)]=\int_{-\infty}^{\infty}G(\sigma)e^{-j\sigma x}\,d\sigma=\int_{-\infty}^{\infty}G(\sigma)[\cos(\sigma)-j\sin(\sigma)]d\sigma \qquad \text{(B.2)}$$

Inserting the expression for G given by equation 4.23 gives:

$$\int_{-\infty}^{\infty}\frac{\sin\!\left(P\sqrt{(k/\sin\theta_B)^2+\sigma^2}\right)}{\sqrt{(k/\sin\theta_B)^2+\sigma^2}}\cos(\sigma)d\sigma-j\int_{-\infty}^{\infty}\frac{\sin\!\left(P\sqrt{(k/\sin\theta_B)^2+\sigma^2}\right)}{\sqrt{(k/\sin\theta_B)^2+\sigma^2}}\sin(\sigma)d\sigma \qquad \text{(B.3)}$$

Note that the first half of this equation is an even function, while the second is an odd function. Due to the integration rules of even and odd functions this becomes:

$$\Im^{-1}[G(\sigma)] = 2 \int_0^\infty \frac{\sin\left(P\sqrt{(k/\sin\theta_B)^2 + \sigma^2}\right)}{\sqrt{(k/\sin\theta_B)^2 + \sigma^2}} \cos(\sigma)d\sigma \qquad (B.4)$$

Now, from the table of integrals by Gradshteyn et.al [1] identity 3.876.1) this is:

$$\Im^{-1}[G(\sigma)] = g(x) = \begin{cases} \pi J_0\left(\dfrac{k}{\sin\theta_B}\sqrt{P^2 - x^2}\right) & [0 \leq x \leq P] \\ 0 & [x < 0, \text{ or } x > P] \end{cases} \qquad (B.5)$$

Repeating this procedure to obtain an expression for B.1b gives:

$$\Im^{-1}\left[\frac{\partial G(\sigma, P)}{\partial P}\right] = \int_{-\infty}^\infty \frac{\partial G(\sigma, P)}{\partial P} e^{-j\sigma x} d\sigma = \frac{\partial}{\partial P}\int_{-\infty}^\infty G(\sigma, P)e^{-j\sigma x} d\sigma = \frac{\partial}{\partial P} g(x) \qquad (B.6)$$

Inserting g(x) in from above gives:

$$\Im^{-1}\left[\frac{\partial G(\sigma, P)}{\partial P}\right] = \begin{cases} \dfrac{-\pi P k\, J_1\left(\dfrac{k}{\sin\theta_B}\sqrt{P^2 - x^2}\right)}{\sin\theta_B \sqrt{P^2 - x^2}} & [0 \leq x \leq P] \\ 0 & \text{otherwise} \end{cases} \qquad (B.7)$$

Finally the expression for B.1c is given by:

$$\Im^{-1}[j\sigma G(\sigma)] = -\frac{df(x)}{dx} = -\frac{d}{dx}\left[\pi J_0\left(\frac{k}{\sin\theta_B}\sqrt{P^2-x^2}\right)\right]$$

$$= \begin{cases} \dfrac{-\pi x k\, J_1\left(\dfrac{k}{\sin\theta_B}\sqrt{P^2-x^2}\right)}{\sin\theta_B\sqrt{P^2-x^2}} & [0 \le y \le P] \\ 0 & \text{otherwise} \end{cases} \qquad (B.8)$$

Combining equations B.5, B.7, and B.8 with the three transform theorems given above

yield the following complex transform pairs

$$\Im^{-1}[\Theta(\sigma)G(\sigma+\overline{\sigma}_0)] = \int_{-P}^{P}\pi J_0\left(\frac{kP}{\sin\theta_B}\sqrt{1-\left(\frac{s}{P}\right)^2}\right)a(x-s)e^{j\sigma_0 s}\,ds \qquad (B.9)$$

$$\Im^{-1}\left[\Theta(\sigma)\frac{\partial G(\sigma+\overline{\sigma}_0)}{\partial P}\right] = \left\{\exp(j\sigma_0 x)\frac{-\pi P k\, J_1\left(\dfrac{k}{\sin\theta_B}\sqrt{P^2-x^2}\right)}{\sin\theta_B\sqrt{P^2-x^2}}*a(x)\right\} \qquad (B.10)$$

$$\Im^{-1}[\Theta(\sigma)(\sigma+\overline{\sigma}_0)G(\sigma+\overline{\sigma}_0)] = j\left[\exp(j\sigma_0 x)\frac{-\pi x k\, J_1\left(\dfrac{k\sqrt{P^2-x^2}}{\sin\theta_B}\right)}{\sin\theta_B\sqrt{P^2-x^2}}*a(x)\right] \qquad (B.11)$$

where $a(x) = \Im^{-1}[\Theta(\sigma)]$ in equations B.9 through B.11.

216

LIST OF REFERENCES

Chapter 1

1. Feldman, M. R. and Erlich, A. E., *Diffractive Optics Improve Product Design*, Photonics Spectra, vol. 29, no. 9, pp. 115-120, 1995.

2. Gwynne, P. ed., *The Photonics Dictionary*, (Lauren Publishing Company, Pittsfield, 1995).

3. Kogelnik, H., The Bell System Technical Journal, *Coupled Wave Theory for Thick Hologram Gratings*, vol. 48, no. 9, pp. 2909-2945, 1969.

Chapter 2

1. Syms, R.R.A., *Practical Volume Holography*, (Clarendon Press, Oxford, 1990), Chapter 1.

2. Kogelnik, H., The Bell System Technical Journal, *Coupled Wave Theory for Thick Hologram Gratings*, vol. 48, no. 9, pp. 2909-2945, 1969.

3. Goodman, J.W., *Introduction to Fourier Optics*, (McGraw-Hill, New York, 1968), Chapter 3.

4. Chu, R. and Tamir, T., *Bragg diffraction of Gaussian beams by periodically modulated media*, Journal of the Optical Society of America, vol. 66, no. 3, pp. 220-226, 1976.

5. Chu, R. and Tamir, T., *Diffraction of Gaussian beams by periodically modulated media for incidence close to a Bragg angle*, Journal of the Optical Society of America, vol. 66, no. 12, pp. 1438-1440, 1976.

6. Benlarbi, B. Russell, P.St.J. and Solymar, L., *Bragg diffraction of finite beams by thick gratings: two rival theories*, Applied Physics B. vol. 28, pp. 63-72, 1982.

7.   Takagi, S., *Dynamical theory of diffraction applicable to crystals with any kind of small distortion*, Acta Crystallography, vol. 15, pp 1311-1312, 1962.

8.   Solymar, L. *A general two-dimenstional theory for volume holograms*, Applied Physics Letters, vol. 31, no. 12, pp. 820-822, 1977.

9.   Owen, M.P., Russell, P. St. J., and Solymar, L, *The analysis of a reflection hologram of triangular shape*, Int. J. Electronics, vol. 47, pp. 267-272, 1979.

10. Solymar, L. and Jordan, M.P. *Finite beams in large volume holograms*, Microwaves, optics and acoustics, vol. 1, pp. 89-92.

11. Russell, P. St. J, *The properties of holographic overlap gratings*, Optica Acta, vol. 26, pp. 329-347, 1979.

12. Solymar, L, and Jordan, M.P., *Two-dimensional transmission type volume holograms for incident plane waves of arbitrary amplitude distribution*, Optical and Quantum Electronics, vol. 9, pp. 437-444, 1977.

13. Russell, P.St.J., *Reconstruction fidelity from volume holograms of finite width and variable index modulation*, Journal of the Optical Society of America, vol. 69, no. 4, pp. 496-503, 1979.

14. Moharam, M.G., Gaylord, T.K., and Magnusson, R, *Diffraction characteristics of three-dimensional crossed-beam volume gratings*, Journal of the Optical Society of America, vol. 70, no. 4, pp. 437-442, 1980.

15. Jordon, M.P. and Solymar, L, *On the properties of a "finite beam" reflection-type volume hologram*, Optical and Quantum Electronics, vol. 10, pp. 503-507, 1978.

16. Russell, *Optical Volume Holography*, Physics Reports, vol. 17, no. 4, pp. 209-312, 1981.

17. Solymar, L and Cooke, D.J., *Volume holography and volume gratings*, (London; New York: Academic Press, 1981), Chapter 7.

18. Solymar, L., *Power Conservation Theorem for 2-Dimensional Volume Holograms*, Electronics Letters, vol. 12, no. 23, pp. 606-607, 1976.

Chapter 3

1.   Syms, R.R.A., *Practical Volume Holography*, (Clarendon Press, Oxford, 1990), Chapter 5.

2. Solymar, L and Cooke, D.J., *Volume holography and volume gratings*, (London; New York: Academic Press, 1981).

3. Shankoff, T.A., *Phase holograms in dichromated gelatin*, Applied Optics, vol 7, no. 10, pp. 2101-2105, 1968.

4. Chang, B.J., and Leonard, C.D. *Dichromated gelatin for the fabrication of holographic optical elements*, Applied Optics, vol. 18, no. 14, pp. 2407-2417, 1979.

5. Chen, F.S., La Macchia, J.R., and Fraser, D.B, *Holographic storage in lithium niobate*, Applied Physics Letters, vol. 13, pp. 223-225, 1968.

6. Staebler, D.L., Burke, W. J., and Amodei, J.J., *Multiple storage and erasure of fixed holograms in Fe-doped LiNbO3*, Applied Physics Letters, vol. 26, no. 4, pp. 182-184, 1975.

7. Mok, F.H., Tackitt, M.C., and Stoll, H.M., *Storage of 500 high-resolution holograms in a LiNbO3 crystal*, Optics Letters, vol. 16, no. 8, pp. 605-607, 1991.

8. Xin, An, Psaltis, D., and Burr, G., *Thermal fixing of 10,000 holograms in LiNbO3:Fe*, Applied Optics, vol. 38, no. 2, pp. 386-393, 1999.

9. Glebov, L.B., Nikonorov, N.V., Panysheva, E.I., Petrovskii, G.T., Savvin, V.V., Tunimanova, I.V., Tsekhomskii, V.A., *Polychromatic glasses – a new material for recording volume phase holograms*, Sov. Phys. Dokl., vol. 35, no. 10, pp. 878-880. 1990.

10. Efimov, O.M., Glebov, L.B., Glebova, L.N., Richardson, K.C., Smirnov, V.I, *High-efficiency Bragg gratings in photothermorefractive glass*, Applied Optics, vol. 38, no. 4, pp. 619-627, 1999.

11. Efimov, O.M., Glebov, L.B., and Smirnov, V.I., *High-frequency Bragg gratings in a photothermorefractive glass*, Optics Letters, vol. 25, no. 23, pp. 1693-1695, 2000.

Chapter 4

1. Moharam , M.G., Gaylord, T.K., and Magnusson, R., *Bragg diffraction of finite beams by thick gratings*, Journal of the Optical Society of America, vol. 70, no. 3, pp. 300-304, 1980.

2. Russell, P.St.J., *Reconstruction fidelity from volume holograms of finite width and variable index modulation*, Journal of the Optical Society of America, vol. 69, no. 4, pp. 496-503, 1979.

3. Benlarbi, B. Russell, P.St.J. and Solymar, L., *Bragg diffraction of Gaussian Beams by thick gratings: Numerical Evaluations by Plane-Wave Decomposition*, Applied Physics B. vol. 28, pp. 383-390, 1982.

4. Chu, R. and Tamir, T., *Bragg diffraction of Gaussian beams by periodically modulated media*, Journal of the Optical Society of America, vol. 66, no. 3, pp. 220-226, 1976.

5. Chu, R. and Tamir, T., *Diffraction of Gaussian beams by periodically modulated media for incidence close to a Bragg angle*, Journal of the Optical Society of America, vol. 66, no. 12, pp. 1438-1440, 1976.

6. Chu, R., Kong, J., and Tamir, T., *Diffraction of Gaussian beams by a periodically modulated layer*, Journal of the Optical Society of America, vol. 67, no. 11, pp. 1555-1561, 1977.

7. Solymar, L, and Jordan, M.P., *Two-dimensional transmission type volume holograms for incident plane waves of arbitrary amplitude distribution*, Optical and Quantum Electronics, vol. 9, pp. 437-444, 1977.

8. Benlarbi, B. Russell, P.St.J. and Solymar, L., *Bragg diffraction of finite beams by thick gratings: two rival theories*, Applied Physics B. vol. 28, pp. 63-72, 1982.

9. Gradshteyn, I.S. and Ryzhik, I.M., *Table of Integrals Series and Products*, (Academic Press, New York, 1965), Chapter 6.

10. Kogelnik, H., The Bell System Technical Journal, *Coupled Wave Theory for Thick Hologram Gratings*, vol. 48, no. 9, pp. 2909-2945, 1969.

11. Russell, *Optical Volume Holography*, Physics Reports, vol. 17, no. 4, pp. 209-312, 1981.

Chapter 5

1. Jordan, M.P., and Solymar, L., *On the properties of a 'finite beam', reflection-type volume hologram*, Optical and Quantum Electronics vol. 10, pp. 503-507, 1978.

2. Owen, M.P., Russell, P. St. J., and Solymar, L, *The analysis of a reflection hologram of triangular shape*, Int. J. Electronics, vol. 47, pp. 267-272, 1979.

3.   Kogelnik, H., The Bell System Technical Journal, *Coupled Wave Theory for Thick Hologram Gratings*, vol. 48, no. 9, pp. 2909-2945, 1969.

4.   Moharam , M.G., Gaylord, T.K., and Magnusson, R., *Bragg diffraction of finite beams by thick gratings*, Journal of the Optical Society of America, vol. 70, no. 3, pp. 300-304, 1980.

5.   Benlarbi, B. Russell, P.St.J. and Solymar, L., *Bragg diffraction of Gaussian Beams by thick gratings: Numerical Evaluations by Plane-Wave Decomposition*, Applied Physics B. vol. 28, pp. 383-390, 1982.

6.   Chu, R. and Tamir, T., *Bragg diffraction of Gaussian beams by periodically modulated media*, Journal of the Optical Society of America, vol. 66, no. 3, pp. 220-226, 1976.

Chapter 6

1.   Siegman, A.E., *Bragg diffraction of a Gaussian beam by a crossed-Gaussian volume grating*, Journal of the Optical Society of America, vol. 67, no. 4, pp. 545-550, 1977.

2.   Courant, R. and Hilbert, D., *Methods of mathematical physics*, (Interscience Publishers, New York) 1953-62.

3.   Moharam, M.G., Gaylord, T.K., and Magnusson, R, *Diffraction characteristics of three-dimensional crossed-beam volume gratings*, Journal of the Optical Society of America, vol. 70, no. 4, pp. 437-442, 1980.

4.   Solymar, L. and Jordan, M.P. *Finite beams in large volume holograms*, Microwaves, optics and acoustics, vol. 1, pp. 89-92.

5.   Kenan, R.P., *Theory of Crossed-Beam Diffraction Gratings*, IEEE Journal of Quantum Electronics, vol. 14, no. 12, pp. 924-930, 1978.

Chapter 7

1.   Russell, P. St. J., *Volume holographic finite-beam conversion with perfect fidelity*, Optica Acta, vol. 27, no. 7, pp. 997-1008, 1980.

Chapter 8

1. Mok, F.H., Tackitt, M.C., and Stoll, H.M., *Storage of 500 high-resolution holograms in a LiNbO₃ crystal*, Applied Optics, vol. 16, no. 8, pp. 605-607, 1991.

2. Staebler, D.L., Burke, W.J., Phillips, W. and Amodei, J.J., *Multiple storage and erasure of fixed holograms in Fe-doped LiNbO₃*, Applied Physics Letters, vol. 26, no. 4, pp. 182-184, 1975.

3. Micheron, F. and Bismuth, G. *Electrical control of fixation and erasure of holographic patterns in ferroelectric materials*, Applied Physics Letters, vol. 20, no. 2, pp. 79-81, 1972.

4. Mok, F.H., *Angle-multiplexed storage of 5000 holograms in lithium niobate*, Optics Letters, vol. 18, no. 11, pp. 915-917, 1993.

5. Tu, K., Tamir, T., and Lee, H., *Multiple-scattering theory of wave diffraction by superposed volume gratings*, Journal of the Optical Society of America A, vol. 7, no. 8, pp. 1421-1435, 1990.

6. Case, S.K., *Coupled-wave theory for multiply exposed thick holographic gratings*, Journal of the Optical Society of America, vol. 65, no. 6, pp. 724-729, 1975.

7. Alferness, R., and Case, S.K., *Coupling in doubly exposed, thick holographic gratings*, Journal of the Optical Society of America, vol. 65, no. 6, pp. 730-739, 1975.

8. An, X., Psaltis, D., and Burr, G., *Thermal fixing of 10,000 holograms in LiNbO₃:Fe*, Applied Optics, vol. 38, no. 2, pp. 386-393, 1999.

9. Moharam, M.G. and Gaylord, T.K., *Chain-matrix analysis of arbitrary-thickness dielectric reflection gratings*, Journal of the Optical Society of America, vol. 72, no. 2, pp. 187-190, 1982.

10. Kashyay, R, *Fiber Bragg Gratings*, (Academic Press, San Diego, 1999), Chapter 2.

11. Othonos, A, and Kalli, K., *Fiber Bragg Gratings, Fundamentals and Applications in Telecommunications and Sensing*, (Artech House, Boston, 1999), Chapter 3.

12. Kogelnik, H., *Filter Response of Nonuniform Almost-Periodic Structures*, The Bell System Technical Journal, vol. 55, no. 1, pp. 109-126, 1976.

13. Hill, K. O., *Aperiodic Distributed-Parameter Waveguides for Integrated Optics*, Applied Optics, vol. 13, no. 8, pp. 1853-1856, 1974

14. Albert, J., Hill, K.O., Malo, B., Theriault, S., Bilodeau, F., Johnson, D.C. and Erickson, L.E., *Apodisation of the spectral response of fibre Bragg gratings using a phase mask with variable diffraction efficiency*, Electronics Letters, vol. 31, no. 3, pp. 222-225, 1995.

15. Canning, J., Psaila, D.C., Brodzeli, Z., Higley, A., and Janos, M., *Characterization of apodized fiber Bragg gratings for rejection filter applications*, Applied Optics, vol. 36, no. 36, pp 9378-9382, 1997.

16. Unger, H., *Circular Waveguide Taper of Improved Design*, The Bell System Technical Journal, vol. 37, pp. 899-912, 1957.

17. Kashyay, R, *Fiber Bragg Gratings*, (Academic Press, San Diego, 1999), Chap. 5.

18. Born, M. and Wolf, E., *Principles of Optics 6th edition*, (Cambridge University Press, Oxford, 1980), Chap. 7.

19. Van de Stadt, H. and Muller, J., *Multimirror Fabry-Perot Interferometers*, Journal of the Optical Society of America A, vol. 2, no. 8, pp. 1363-1370, 1985.

20. Pelletier, E., and Macleod, H.A., *Interference filters with multiple peaks*, Journal of the Optical Society of America, vol. 72, no. 6, pp. 683-687, 1982.

21. Macleod, H.A, *The Design of Interference Filters with Multiple Peaks*, 37th Annual Technical Conference Proceedings, Society of Vacuum Coaters , pp 53-58, 1994.

22. Epstein, L.I., *Design of Optical Filters*, Journal of the Optical Society of America, vol. 42, no. 11, pp 806-810, 1952.

23. Thelen, A., *Equivalent Layers in Multilayer Filters*, Journal of the Optical Society of America, vol. 56, no. 11, pp. 1533-1538, 1966.

24. Macleod, H.A., *Thin-Film Optical Filters, 2nd ed.*, (Macmillan Publishing Company, New York).

25. Jacobs, C., *Dielectric square bandpass design*, Applied Optics, vol. 20, no. 6, ppl 1039-1042, 1981.

Chapter 9

1. Kermisch, D., *Nonuniform Sinusoidally Modulated Dielectric Gratings*, Journal of the Optical Society of America, vol. 59, no. 11, pp. 1409-1414, 1969.

2. Uchida, N., *Calculation of diffraction efficiency in hologram gratings attenuated along the direction perpendicular to the grating vector*, Journal of the Optical Society of America, vol. 63, no. 3, pp. 280-287, 1973.

3. Kubota, T., *The diffraction efficiency of hologram gratings recorded in an absorptive medium*, Optics Communications, vol. 16, no. 3, pp. 347-349, 1976.

4. Kubota, T., *Characteristics of thick hologram grating recorded in absorptive medium*, Optica Acta, vol. 25, no. 11, pp. 1035-1053, 1978.

5. Kowarschik, R., *Diffraction efficiency of attenuated sinusoidally modulated gratings in volume holograms*, Optica Acta, vol. 23, no. 12, pp. 1039-1051, 1976.

6. Morozumi, S., *Diffraction Efficiency of Hologram Gratings with Modulation Changing through the Thickness*, Japanese Journal of Applied Physics, vol. 15, no. 10, pp. 1929-1935, 1976.

7. Ninomiya, Y., *Recording characteristics of volume holograms*, Journal of the Optical Society of America, vol. 63, no. 9, pp. 1124-1130, 1973.

8. Killat, U., *Coupled Wave Theory of Hologram Gratings with Arbitrary Attenuation*, Optics Communications, vol. 21, no. 1, pp. 110-111, 1977.

9. Kogelnik, H. *Filter Response of Nonuniform Almost-Periodic Structures*, The Bell System Technical Journal, vol. 55, no. 1, pp. 109-126, 1976.

10. Lederer, F. and Langbein, U., *Attenuated thick hologram gratings. Part 1: Diffraction efficiency*, Optical and Quantum Electronics, vol. 9, pp. 473-485, 1977.

11. Owen, M.P. and Solymar, L., *Efficiency of Volume Phase Reflection Holograms Recorded in an Attenuating Medium*, Optics Communications, vol. 34, no. 3, pp. 321-326, 1980.

12. Jordan, M.P., and Solymar, L., *A note on volume holograms*, Electronics Letters, vol. 14., no. 9, pp. 271-272, 1978.

224

Appendix A

1. Courant, R. and Hilbert, D., *Methods of mathematical physics*, (Interscience Publishers, New York) 1953-62.

Appendix B

1. Gradshteyn, I.S. and Ryzhik, I.M., *Table of Integrals Series and Products*, (Academic Press, New York, 1965), Chapter 6.