4/12/96

SANDIA REPORT

SAND96–0261 • UC–410
Patent Caution
Printed February 1996

RECEIVED
MAY 17 1996
OSTI

Incorporation of Wavelength Selective Devices into Waveguides with Applications to a Miniature Spectrometer

B. R. Stallard, S. Kaushik, G. R. Hadley, I. J. Fritz,
A. J. Howard, G. A. Vawter, J. R. Wendt, R. Corless

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550
for the United States Department of Energy
under Contract DE-AC04-94AL85000

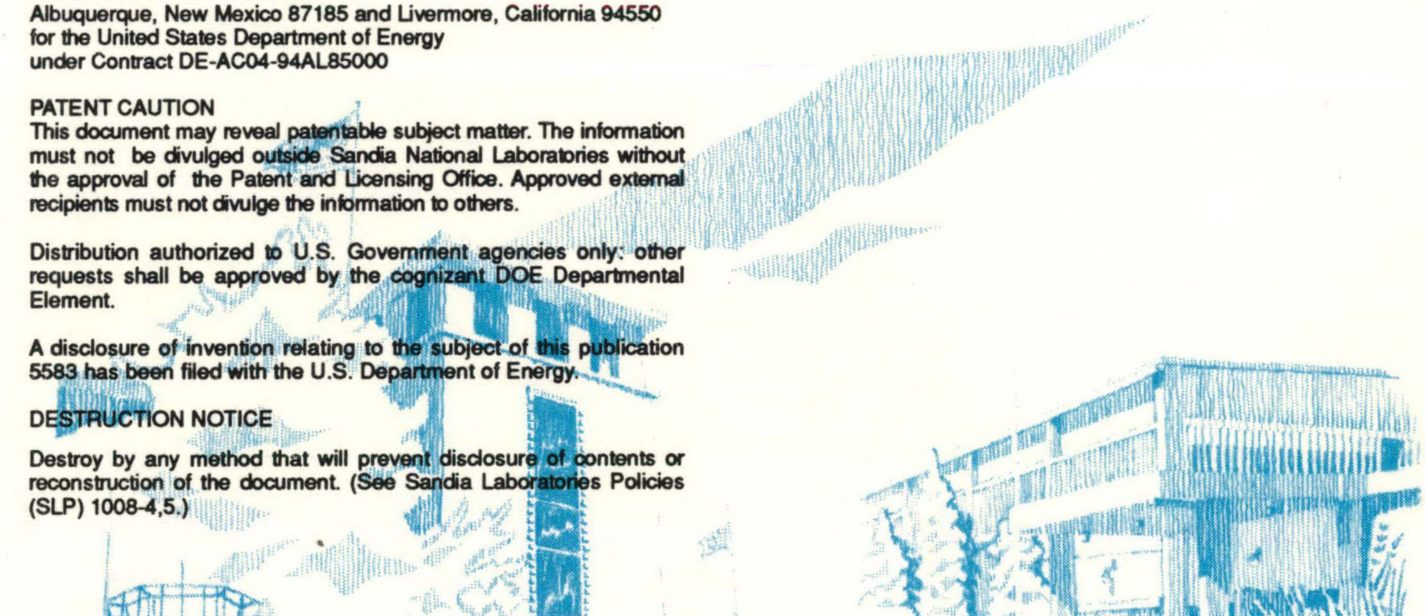
PATENT CAUTION
This document may reveal patentable subject matter. The information must not be divulged outside Sandia National Laboratories without the approval of the Patent and Licensing Office. Approved external recipients must not divulge the information to others.

Distribution authorized to U.S. Government agencies only: other requests shall be approved by the cognizant DOE Departmental Element.

A disclosure of invention relating to the subject of this publication 5583 has been filed with the U.S. Department of Energy.

DESTRUCTION NOTICE

Destroy by any method that will prevent disclosure of contents or reconstruction of the document. (See Sandia Laboratories Policies (SLP) 1008-4,5.)



MASTER

SF2900Q(8-81)

DISTRIBUTION OF THIS DOCUMENT IS LIMITED
No Automatic Distribution or Announcement
Refer all Requests Sandia

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof or any of their contractors.

SAND96-0261
Patent Caution
Printed February 1996

Distribution
Category UC-410

Incorporation of Wavelength Selective Devices into Waveguides with Applications to a Miniature Spectrometer

B.R.Stallard, S.Kaushik (1823), G.R. Hadley,
I.J. Fritz (1312), A.J. Howard, G.A. Vawter,
J.R. Wendt (1322), R. Corless, (1321-1)
Surface/Molecular
Spectro/Gas
Analysis Department
Sandia National Laboratories
Albuquerque, NM 87185

Abstract

This report pertains to a Laboratory Directed Research and Development project which was funded for FY94 and FY95. The goal was to develop building blocks for small, cheap sensors that use optical spectroscopy as a means of detecting chemical analytes. Such sensors can have an impact on a wide variety of technologies, such as: industrial process control, environmental monitors, chemical analysis in medicine, and automotive monitors. We describe work in fabricating and demonstrating a waveguide/grating device that can serve as the wavelength dispersive component in a miniature spectrometer. Also, we describe the invention and modeling of a new way to construct an array of optical interference filters using sub-wavelength lithography to tune the index of refraction of a fixed Fabry-Perot cavity. Next we describe progress in more efficiently calculating the fields in grating devices. Finally we present the invention of a new type of near field optical probe, applicable to scanning microscopy or optical data storage, which is based on a circular grating constructed in a waveguide. This result diverges from the original goal of the project but is quite significant in that it promises to increase the data storage capacity of CD-ROMs by 10 times.

Contents

Introduction

This report pertains to the Laboratory Directed Research and Development project with the same title which was funded for FY94 and FY95. The report consists of the journal articles and disclosures which are listed as Sections 2 to 6 in the Table of Contents.

The goal of this work is to develop building blocks for small, cheap sensors that will use optical spectroscopy as a means of detecting chemical analytes. Fabrication of the devices by microelectronics techniques is a key feature of the work so that the goals of *small* and *cheap* may eventually be realized. We have principally concentrated on the wavelength range of 1.5 to 1.8 μm. This gives us access to the first overtone of CH, NH, and OH stretches. Such sensors can have an impact on a wide variety of technologies, such as: industrial process control, environmental monitors, chemical analysis in medicine, and automotive monitors. Many detection problems can be solved with instruments of lower spectral resolution and lower signal-to-noise characteristics than traditional bench top equipment. The hope is that small, cheap sensors may lead to more widespread and beneficial applications of optical spectroscopic detection.

Section 2 describes work in fabricating and demonstrating a waveguide device that can serve as the wavelength dispersive component in a miniature spectrometer. Sections 3 and 4 describe the invention and modeling of a new way to construct an array of optical interference filters using sub-wavelength lithography to tune the index of refraction of a fixed Fabry-Perot cavity. This filter array can serve as the wavelength selective portion of a miniature spectrometer. Section 5 describes progress in more efficiently calculating the fields in grating devices. The new methods are particularly attractive because they are highly intuitive and allow one to write equations for complicated structures by inspection. A more complete presentation of our novel theoretical approach will be submitted to the Journal of the Optical Society of America in early 1996. In Section 6, we present the invention of a new type of near field optical probe, applicable to scanning microscopy or optical data storage, which is based on a circular grating constructed in a waveguide. Although this last result diverges from the original intent of the project, it was a natural outgrowth of our work and may prove quite significant.

This project evolved into more theory and less experiment than was envisioned at the outset. Nevertheless, the work has produced ideas for devices that appear to be truly revolutionary. We are especially excited about the very small two-dimensional filter array and its application to near-ir sensors. Also, we are excited about the near field optical probe and its potential to increase the density for CD ROM data storage by a factor of 10 or more. Fabrication of these devices is continuing with other funding. Patents and funding for commercialization are being sought.

Near-IR Spectroscopy with a Dispersive Waveguide Device

B. R. Stallard, A. J. Howard, R. Corless, G. R. Hadley, R. K. Rowe,
G. A. Vawter, J. R. Wendt, and I. J. Fritz

Sandia National Laboratories, Albuquerque, NM 87185-1405

ABSTRACT

Miniature, low cost sensors are in demand for a variety of applications in industry, medicine, and environmental sciences. As a first step in developing such a sensor, we have etched a grating into a GaAs rib waveguide to serve as a wavelength dispersive element. The device was fabricated with the techniques of metal-organic chemical vapor deposition, electron-beam lithography, optical lithography, and reactive ion-beam etching. While full integration is the eventual goal of this work, for the present, a functional spectrometer was constructed with the addition of a discrete source, sample cell, lenses, and detector. The waveguide spectrometer has a spectral resolution of 7.5 nm and a spectral dispersion of 0.11°/nm. As presently configured, it functions in the spectral range of 1500 to 1600 nm. A demonstration of the analytical capability of the waveguide spectrometer is presented. The problem posed is the determination of diethanol amine in an ethanol solution (about 10 to 100 mg/ml). This involves the detection of the first overtone of the NH stretch at 1545 nm in a moderately absorbing solvent background. The standard error of prediction for the determination was 5.4 mg/ml.

Index Headings: optical sensors, near-ir spectroscopy, optical waveguides

INTRODUCTION

Miniature, low cost sensors are in demand for a variety of applications in industry, medicine, and environmental sciences. One concept to meet this need is a miniature waveguide spectrometer built with integrated circuit technology. The proposed spectrometer has four fully integrated components, a source, a sample interaction area, a wavelength selective device, and a detector. We report the fabrication and testing of a waveguide device with wavelength dispersive characteristics suitable for incorporation into such a miniature spectrometer. For the present, this device is combined with discrete components to produce a functional spectrometer. The characteristics of the waveguide spectrometer are determined and a demonstration of its analytical application is performed.

The present work is inspired by the area of research known as wavelength-division multiplexing (WDM) which is important to the communication industry[1]. Modulated signals are imposed on distinct wavelengths of light and the signals combined for transmission. At the receiving end the wavelengths are separated in order to unravel the individual signals. Many devices have been studied for this task and some have failed because of the requirement for high spectral resolution. However, sensor applications of near-ir spectroscopy rarely require spectral resolution better than a few nm. Therefore, the discarded ideas from the WDM research may deserve further attention for development as spectroscopic sensors. The present example is one of the simplest, a grating etched into a rib waveguide. Our work is related to the work of Goldman and co-workers[2], yet is unique in that III-V materials are employed thereby opening the way for integration with a source and a detector.

EXPERIMENTAL

Fabrication of the waveguide device was accomplished in the facilities of the Compound Semiconductor Research Laboratory at Sandia National Laboratories. The primary fabrication tools were metal-organic chemical vapor deposition(MOCVD), electron beam lithography, optical contact lithograpy, and reactive ion-beam etching (RIBE).

Fig. 1 is a schematic of the optical set up used in these experiments. The waveguide device is positioned with the orientation pictured in the top of Fig. 2. The source is a model C86069-13 edge emitting diode from EG&G Optoelectronics Division, with the fiber optic pigtail and can removed. The spectrum of the source is centered at about 1550 nm with a full width half maximum (FWHM) of about 110 nm. The total optical power is about 0.5 mwatt. The cameras are model 7290 from Electrophysics, having room temperature PbS cathodes. Cameras #1 and #2 are used for alignment only, while camera #3 collects the spectroscopic data. The lens shown schematically in front of camera #3 is a spherical and cylindrical lens combination required to image the spectrally dispersed signal onto the cathode of the camera. The other three lenses are of an infinite conjugate design placed in a microscope style mounting. The image on camera #3 is captured with a Data Translation model DT2876 board in a 486 PC. The image is converted to a conventional spectrum using the line scan function of the Global Data Image software.

The interference filter, which is used for calibrating the spectrometer, was purchased from CVI Laser. It has a peak transmission (i.e. passband) for 0° incident angle at about 1570 nm and a transmission bandwidth (FWHM) of about 5.3 nm. When the optical set up of Fig. 1 is to be used as a spectrometer, the interference filter shown in Fig. 1 is replaced with a conventional glass cuvette.

Samples for the analytical demonstration were prepared by weighing reagent grade diethanol amine (DEA) into 10 ml volumetric flasks and filling to the mark with reagent grade ethanol. The commercial spectrometer used for some parts of this work was a Nicolet 800 Fourier transform (FT) instrument with a liquid nitrogen cooled InSb detector.

RESULTS AND DISCUSSION

Fabrication of the devices. Figs. 2 and 3 illustrate schematically the construction of the dispersive waveguide device. The following layers were grown on a two inch GaAs substrate by MOCVD: 2.5 μm of AlGaAs, as the bottom cladding layer, 0.2 μm of GaAs, as the guiding layer, and 0.6 mm as the top cladding layer. Electron-beam lithography was employed to pattern a second order grating with a pitch of 0.475 μm and approximately equal lines and spaces. The grating stretches 600 μm along the length of the waveguide. The pattern for the 6 μm wide rib waveguide was produced with optical lithography. Etching was done with a RIBE; first, the electron beam pattern, to a 0.2 μm depth, and then the optical pattern, to a 1.5 μm depth. Cross-sectional scanning electron microscope images were taken to verify the etch depths. For the optical experiments, the wafers were cleaved into pieces about 2 mm square containing several dispersive devices, including some waveguides without gratings which are helpful in aligning the source to the device.

Calibration and characterization. Fig. 4 shows six individual transmission curves for the bandpass filter, obtained with the optical apparatus in Fig. 1, where the bandpass filter is rotated to give varying passbands. Having previously calibrated the rotation angle versus wavelength, the pixel scale can be converted to wavelength as in Fig. 6a. Also, a knowledge of the specific optical components used for imaging allows us to calculate a spectral dispersion of 0.11°/nm for the grating device. Fig. 4 also contains information regarding the spectral resolution of the apparatus. Knowing the spectral width of the bandpass filter, we calculate (using a simple Gaussian convolution model) that the spectral broadening function of the waveguide spectrometer is about 7.5 nm (FWHM). This can be considered the spectral resolution of the waveguide spectrometer. The resolution can be altered somewhat by the optics chosen to image the grating onto the cathode of the camera. However, there is a more fundamental limitation to the resolution of the waveguide spectrometer, that is the interaction distance of the grating with the light in the waveguide. The interaction length is, in-turn, a function of the groove depth of the grating. Shallow grooves permit the light to interact with more teeth of the grating giving a higher resolving power. Yet spreading the light out over a wider portion of the grating gives, in effect, a wider slit width, to draw an analogy with a conventional grating monochrometer. A thorough understanding of the relationship of groove depth and spectral resolution will require further study.

An analytical demonstration. In order to demonstrate the application of the waveguide spectrometer we have chosen to determine the concentration of DEA in an ethanol solution over the range of about 10 to 100 mg/ml. This involves the detection of the first overtone of the NH stretch at about 1545 nm in the presence of a moderately absorbing solvent background. Although there is no specific commercial need for this determination, it represents an important class of compounds, the amines. The determination was first performed with a commercial FT spectrometer. The results are presented in Fig. 5. Part (a) is a series of representative spectra with differing amounts of DEA in a cell of 1 mm path length. Since the reference spectrum is air, both the NH peak at 1545 nm and the OH peak, primarily from the solvent, at 1580 nm are evident. Fig. 5b is the cross-validated calibration curve constructed using partial least squares

(PLS) data analysis[3-5]. The standard error of prediction (SEP), 0.41 mg/ml, is a measure of the precision of the determination assuming negligible error in the reference concentrations.

To acquire spectra with the waveguide spectrometer, we remove the interference filter from the apparatus of Fig. 1 and place a 5 mm optical sample cell in the collimated beam at the same position. Fig. 6 contains the results of the DEA determination. In this case, the reference spectrum is ethanol without DEA (rather than air as above) so that the OH peak from the solvent is removed from the spectrum. The SEP is 5.4 mg/ml; less precise that the results from the commercial FT instrument, but still a small fraction of the range of the samples. It should be noted that the precision of the determination is largely a function of the detector type and source intensity, so that absolute performance capabilities of the waveguide spectrometer should not be inferred. It is clear that the waveguide spectrometer is capable of performing a quantitative analysis for this NH containing compound.

ACKNOWLEDGMENTS

We acknowledge the support of the Laboratory Directed Research and Development Program at Sandia National Laboratories.

REFERENCES

1. J. P. Laude, *Wavelength Division Multiplexing* (Simon Schuster, New York, 1993).
2. D. S. Goldman, P. L. White, and N. C. Anheir, *Appl. Opt.* **29**, 4583 (1990).
3. M. P. Fuller, G. L. Ritter, and C. S. Draper, *Appl. Spect.* **42**, 217 (1988).
4. D. M. Haaland, R. G. Easterling, and D. A. Vopicka, *Appl. Spect.* **39**, 73 (1985).
5. E. V. Thomas and D. M. Haaland, *Anal. Chem.* **62**, 1091 (1990).

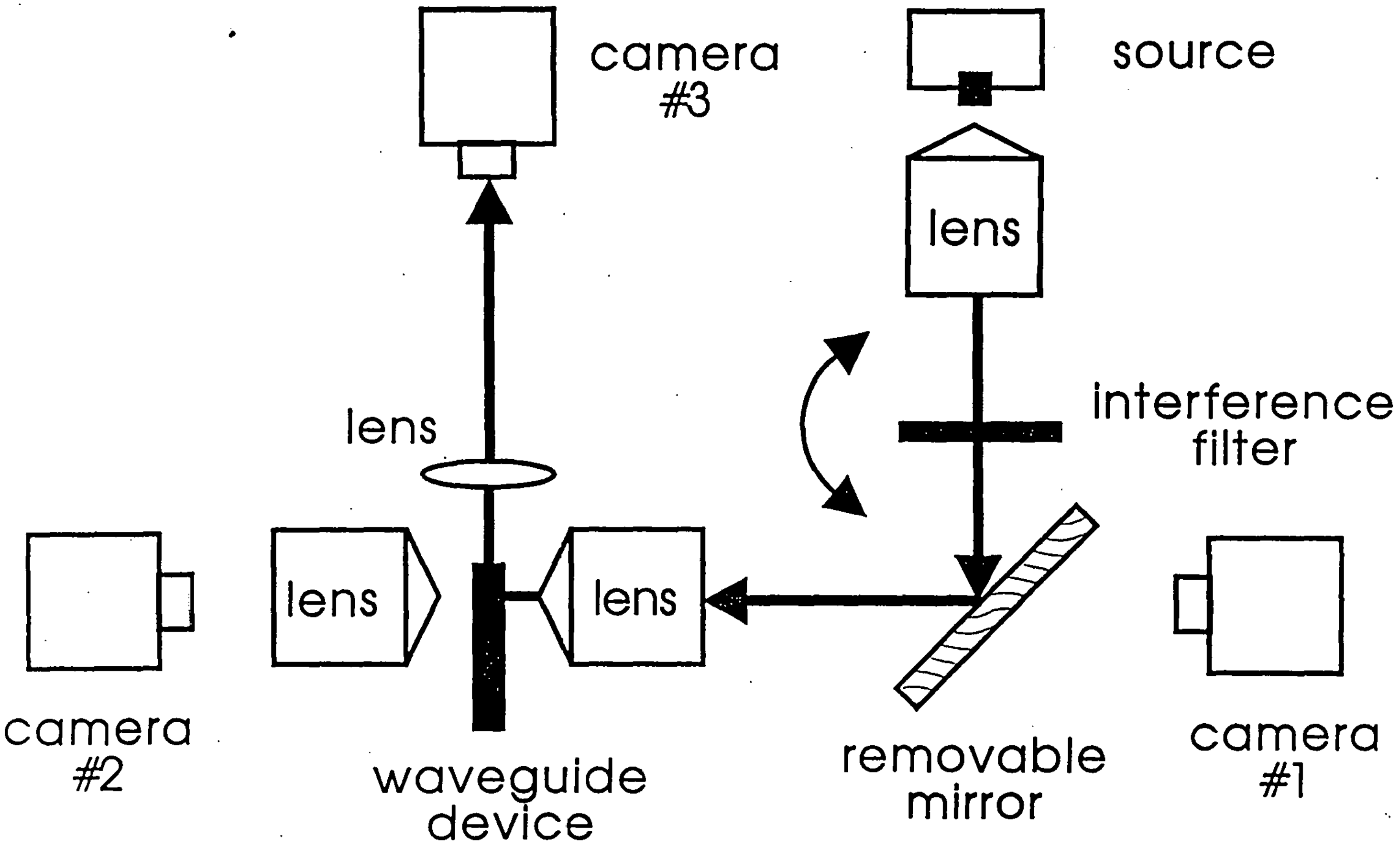


Fig. 1. A schematic drawing of the optical set up used to obtain spectra with the dispersive waveguide device. See the text for a complete explanation. The interference filter is shown in position for calibration. To acquire a sample spectrum the interference filter is replaced by a standard glass cuvette.

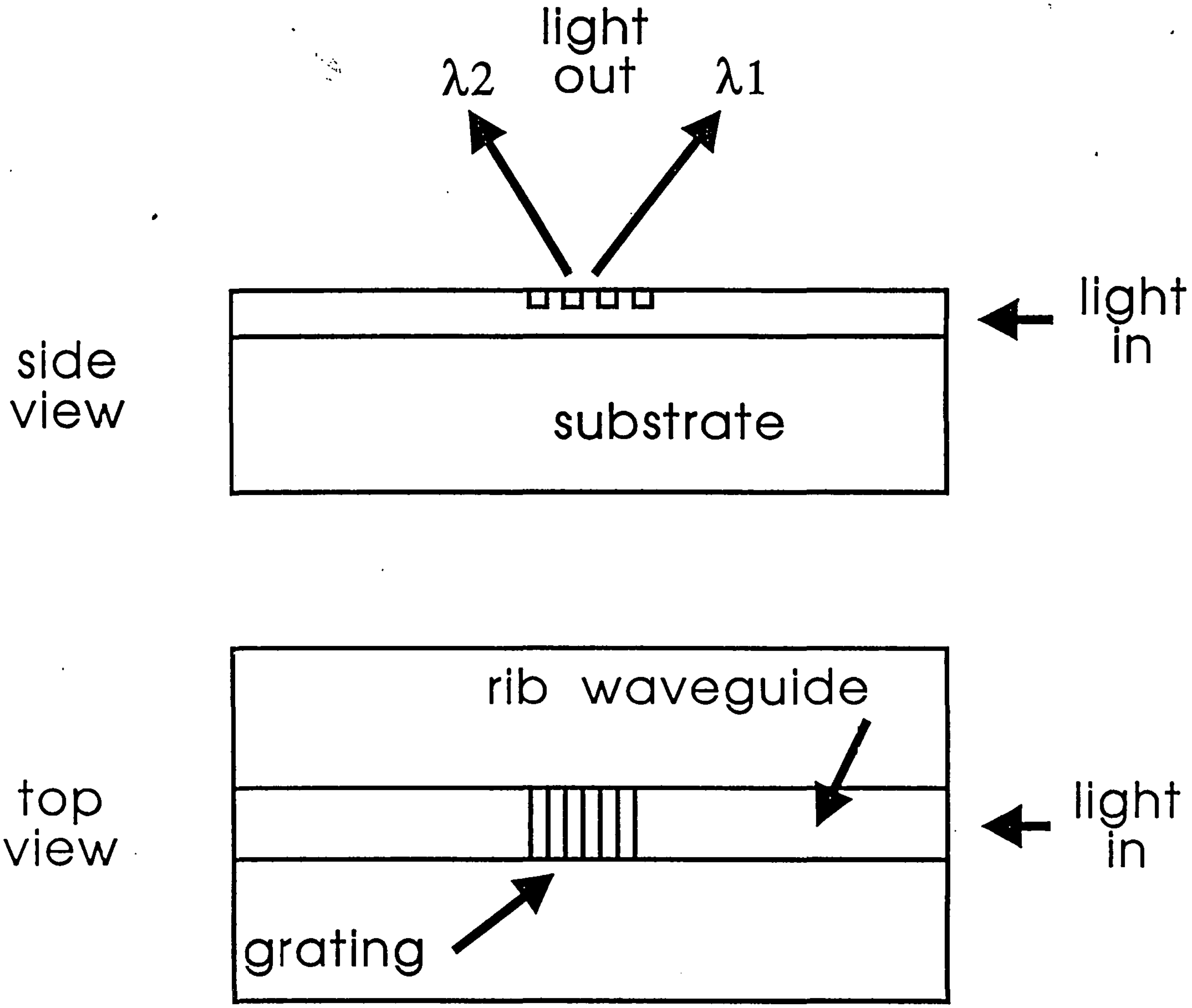


Fig. 2. A schematic drawing of the dispersive waveguide device, a grating etched into a rib waveguide as described in the text.

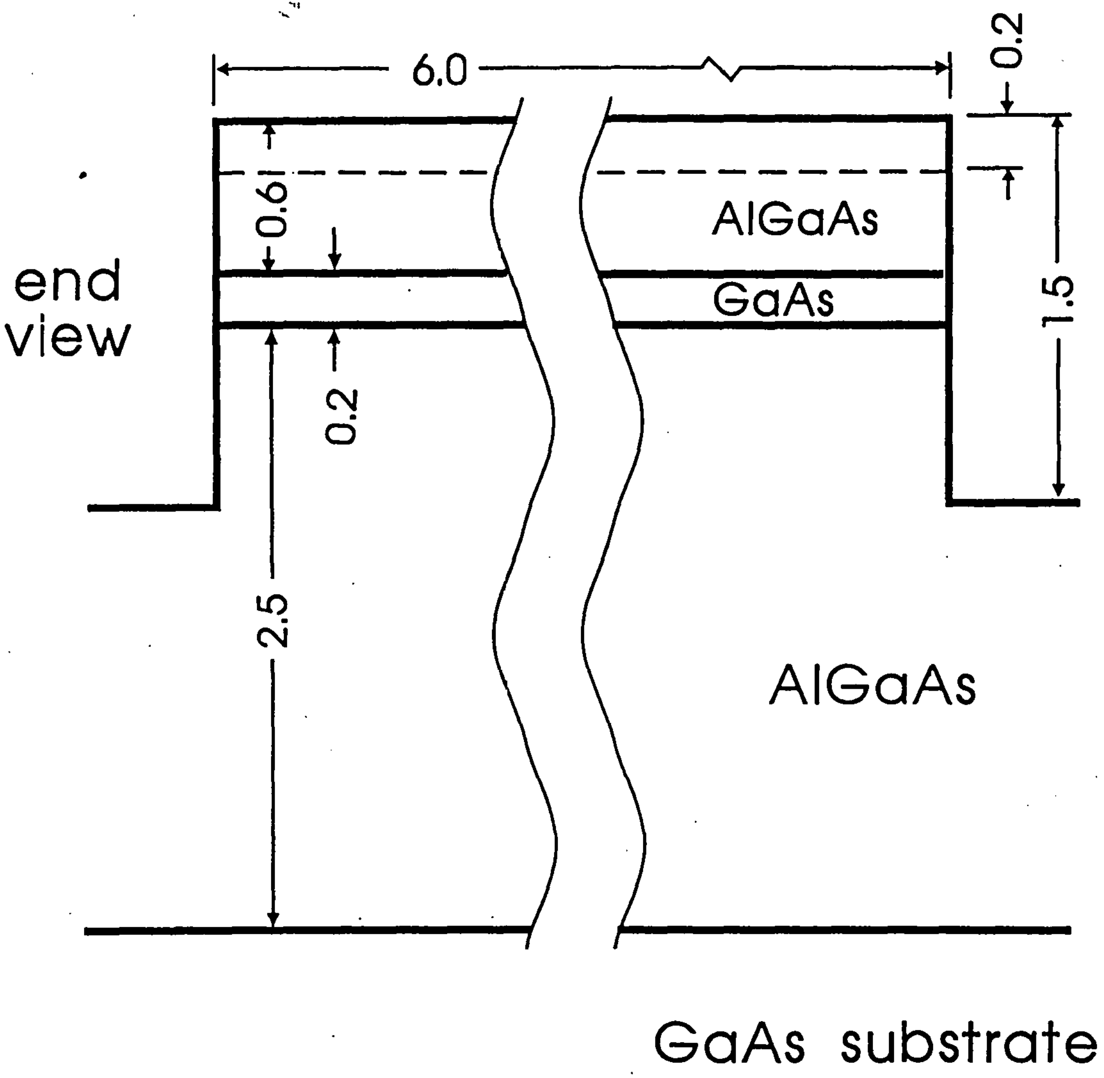


Fig. 3. An end-on, cross-sectional view of the waveguide dispersive device showing the layers and etch depths. Dimensions are in μm.

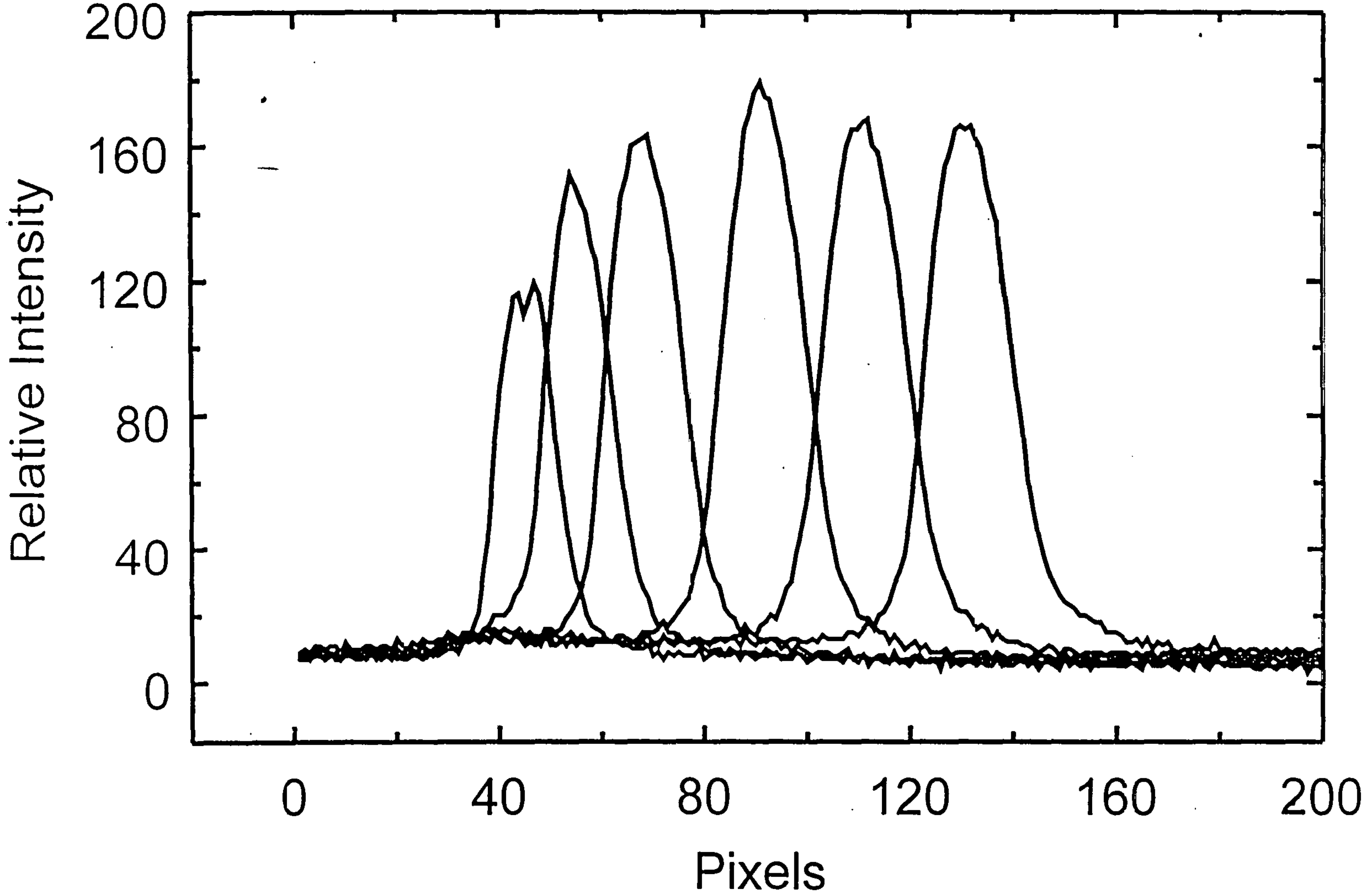


Fig. 4. A calibration plot showing six separate spectra obtained with the interference filter rotated to provide a variable passband. A traditional spectrometer is used to determine the dependence of peak position on angle for the filter (not shown). With this information one may convert the pixels to wavelength, as was done in Fig. 6a.

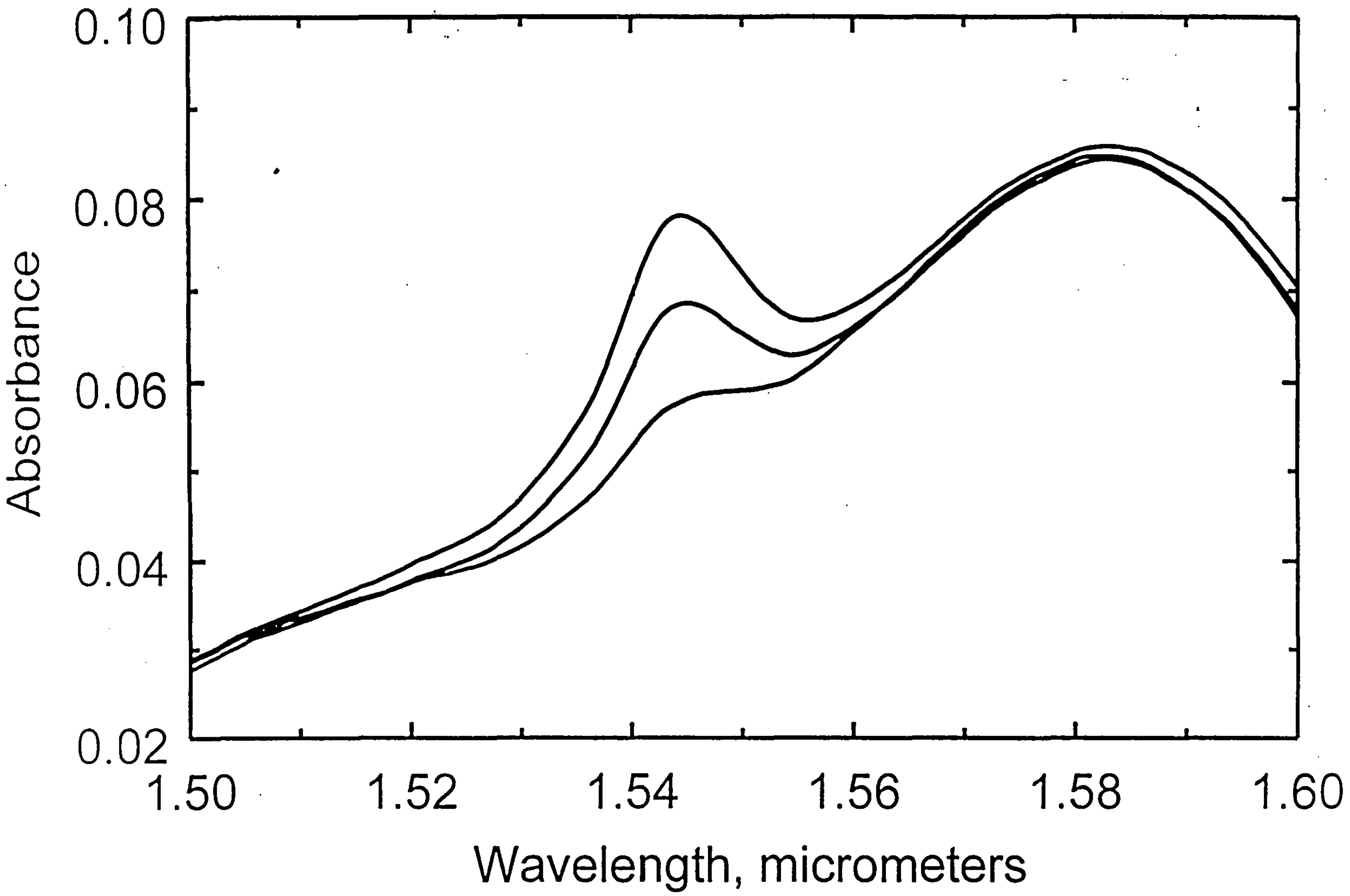


Fig. 5. (a) Representative spectra, acquired with a commercial FT instrument. The samples are DEA in ethanol, path length 1 mm, with air as the reference. The overtones of both the NH stretch at 1.545 μm and the OH stretch at 1.58 μm are evident.

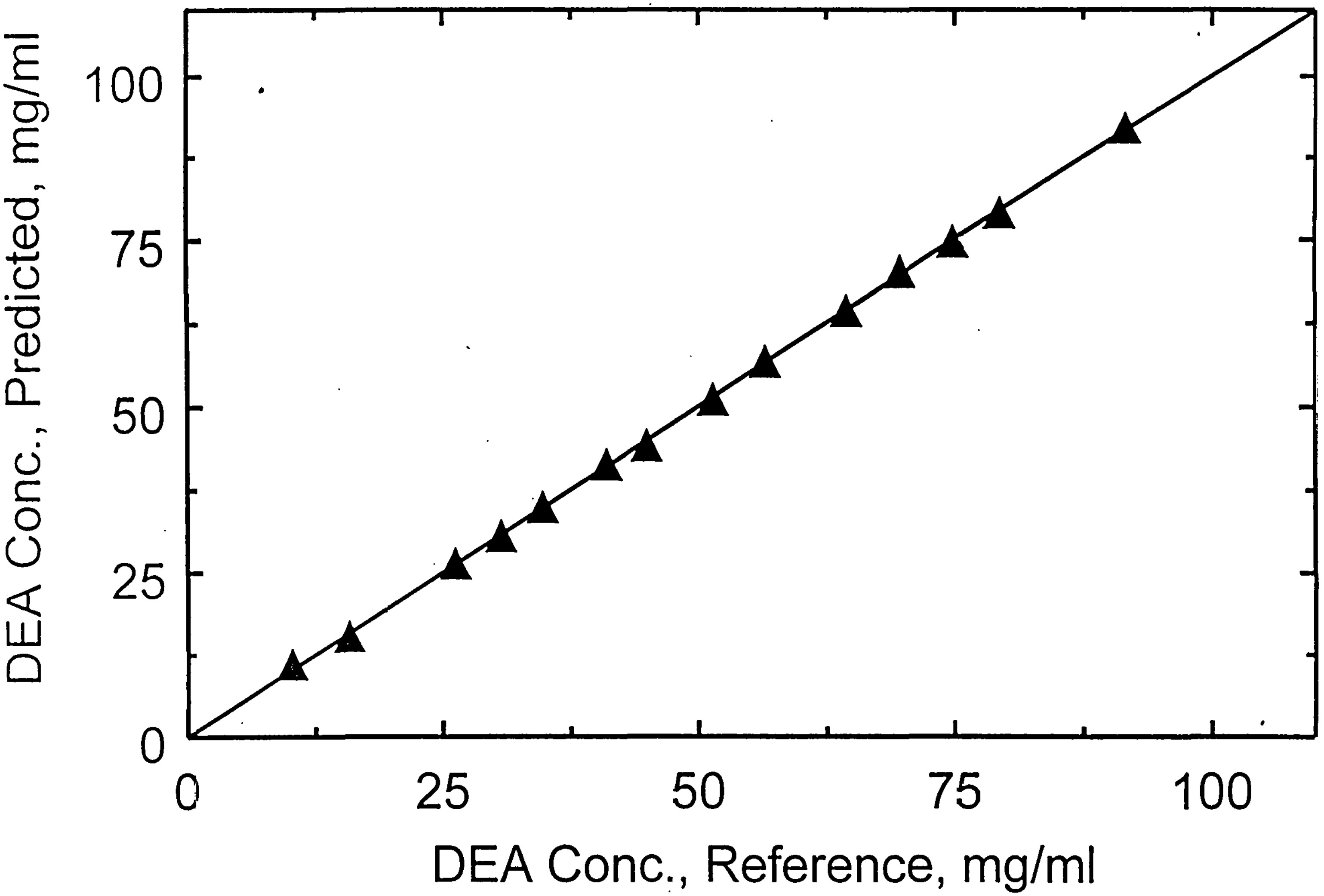


Fig. 5. (b) The cross-validated PLS calibration plot for the set of samples run on the FT instrument. The SEP is 0.41 mg/ml.

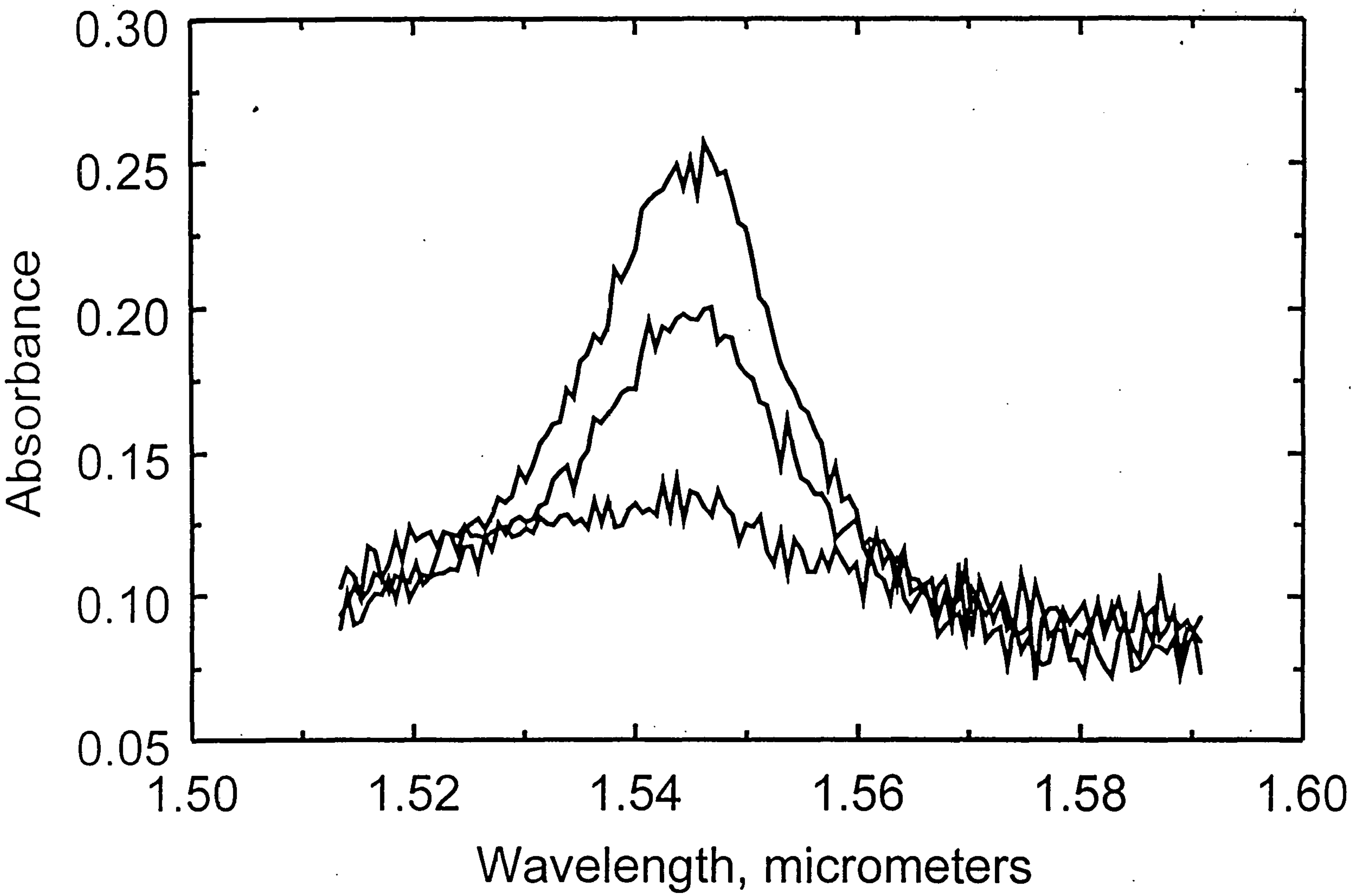


Fig. 6. (a) Representative spectra, acquired with the waveguide spectrometer. The samples are DEA in ethanol, path length 5 mm, with ethanol as the reference. The overtone of the NH stretch at 1.545 μm is evident.

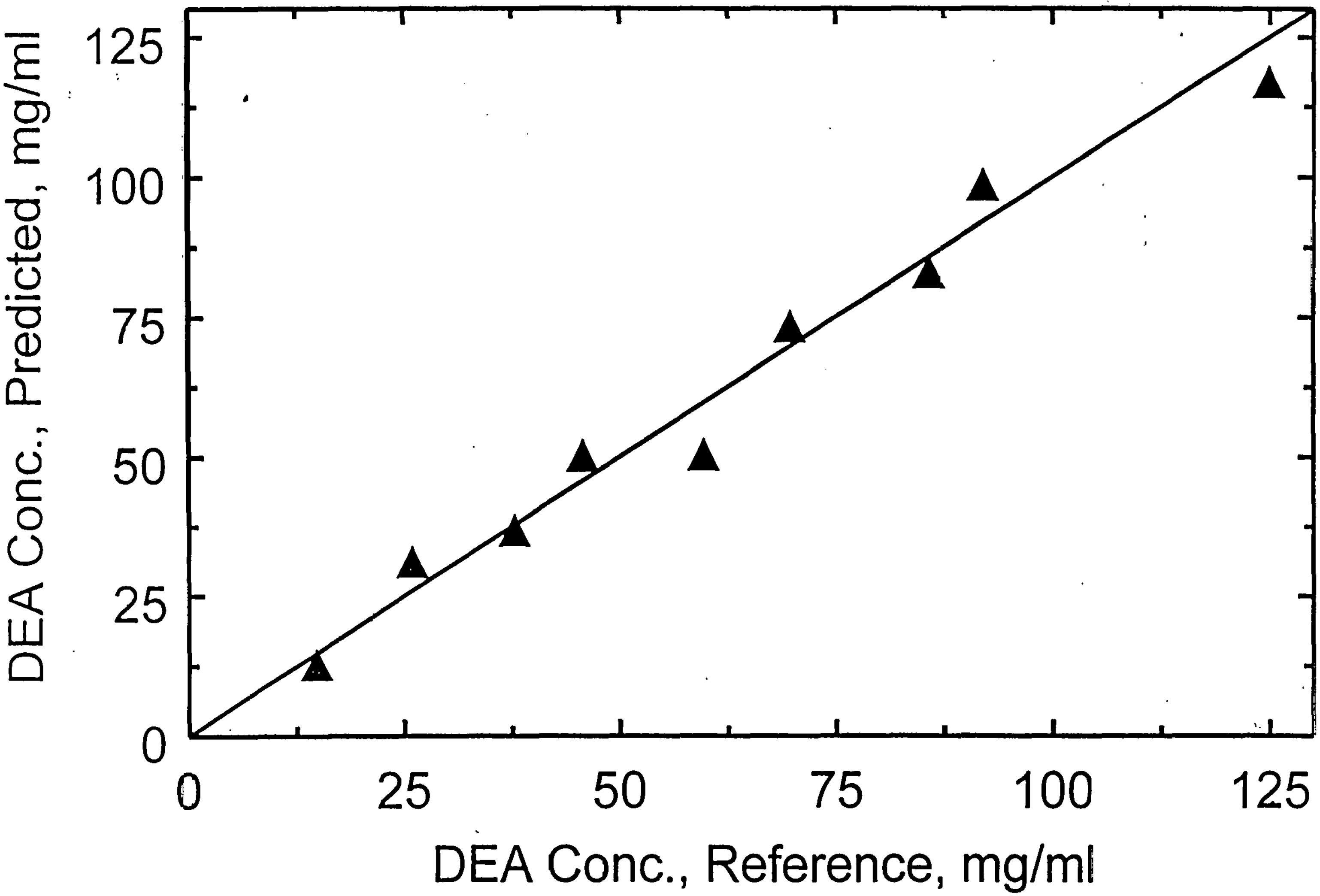


Fig. 6. (b) The cross-validated PLS calibration plot for the set of samples run on the waveguide spectrometer. The SEP is 5.4 mg/ml.

Disclosure of Technical Advance (TA)

"A Two Dimensional Array of Optical Interference Filters Produced by Lithographic Alterations of the Index of Refraction."

Sumanth Kaushik and Brian R. Stallard

Description

Optical interference filters appear in a number of forms. The class of filters with the highest transmission and narrowest linewidths are Fabry-Perot filters. These filters are formed by sandwiching a half-wave $\lambda_0/2$ (or some multiple thereof) spacer layer between two mirrors. The lowest absorptive losses are typically found in mirrors made from multi-layer dielectric stacks that are usually quarter wave in thickness ($\lambda_0/4$). The bandpass of these filters (we denote the center wavelength of the pass-band as λ_0) is determined by the thickness of the spacer layer and the bandwidth is determined both by the thickness of the spacer layer, as well as the reflectivities of the two mirrors.

The most common Fabry-Perot filters commercially available have pass-bands that are spatially uniform with variations as small as .1%. Filters with spatially varying pass-bands, however, are readily available. These filters are manufactured by proportionally varying the thickness of all the layers uniformly across the surface of the filters. These filters, called wedged filters, are either circular (circular variable filter - CVF) or linear (linear variable filter - LVF). In the former, the thicknesses of the multilayer films vary linearly with the angular position across a circular substrate; in the latter, they vary linearly with position across a rectangular substrate.

Although these filters are widely used, they have several limitations. First, with wedge filters, it is not possible to have both large dispersion (>> 10 nm/mm) as well as good spectral response. The dispersion is proportional to the wedge angle.; however a large wedge angle results in poor transmission characteristics (due to non-normal reflections from the angled mirror facet), as well as a loss in bandwidth (due to wavefront inhomogenity across the angled facets). Owing to these limitations, wedge filters are large (e.g. typical LVFs in the 300 nm - 700 nm spectral range are 60 mm in length) and have dispersion of about 5-10 nm/mm (CVFs are typically 25-50 mm in diameter and have a dispersion of 1 nm/degree). Significant improvements in these numbers is not expected ; therefore, ultra-miniaturization of devices (e.g. sub-millimeter) using LVF and CVF is not possible with existing technology.

Another restrictive feature of wedge filters is that the pass-band can change only in a continuous manner across the substrate. In many spectroscopic applications (such as the design of optical sensors), one is often interested in only a limited number of distinct, and often non-uniformly spaced spectral lines. However, with a spectrophotometer using a wedge filter, one cannot pick and choose the desired channel unless one has a well collimated (uniform) beam that encompasses all the desired channels. Since maintaining a good collimation along with acceptable light levels is difficult (especially if the channels are separated by 60 mm), it is often necessary to either move the spectrophotometer (e.g. a rotating CVF), or it is necessary to steer the beam. Neither of these strategies is well suited for rapid data collection, especially if the desired channels are physically separated by a large distance.

Another impediment to the miniaturization of wedge filters is that they are essentially one-dimensional structures. Typically, pass-bands can be varied along only one dimension (linear position in LVF and azimuthal angle in CVF). A two-dimensional wedge filter is difficult to fabricate and to our knowledge, none are available commercially.

We describe here a new technique to produce on a single substrate, a two dimensional array of optical interference filters in which the pass-band at any spatial location on the substrate is a variable, user supplied parameter. An operational diagram of a typical "filter array" is shown in Figure 1. This array is arranged in a pixel format similar to that of to a two-dimensional photodetector array (e.g. silicon CCD array). This is to suggest a possible application of such a filter array, in conjunction with a two-dimensional detector array, to the design of a compact spectrophotometer. However, unlike conventional miniature spectrophotometers that utilize LVFs and CVFs, each channel can be individually assigned a different pass-band wavelength over a relatively large spectral range ($\Delta\lambda/\lambda_0 \sim .4$). In applications where information at only a few wavelengths are desired, a spectrophotometer made using the filter array can be extremely miniature since the pixel size can be made as small as the dimensions of the detector pixel (< 100 μm).

Our technique is to form an interference filter by optically contacting two dielectric mirrors, where a series of one and two dimensional periodic steps is lithographically patterned on the top surface of one of the dielectric mirrors. A sketch of our design is shown in Figure 2. In the absence of any patterning, the juxtaposition of two identical dielectric mirrors leads to the formation of a natural Fabry-Perot cavity having a width of $\lambda_0/(2n_L)$. The presence of the two mirrors on the top and bottom of this cavity yields the prototypical optical interference filter that is well documented in the literature, with the pass-band centered at λ_0 and a bandwidth determined by the finesse. Note that in Fig. 2, both the multilayer mirrors end with a low-index layer. Since each of the low (high) index layer is quarter wave thick ($\lambda_0/(4n_L)$ for the low index layer and $\lambda_0/(4n_H)$ for the high index layer), the combination of the two mirrors yields the cavity with the width $\lambda_0/(2n_L)= 2\,\lambda_0/(4n_L)$. The dielectric mirror shown in Fig. 2 (with three high index Silicon layers) has a reflectivity of approximately 98%.

The pass-band is changed by etching a periodic pattern onto the top surface of one of the mirrors. The etched surface appears as a set of periodic grooves or steps in the outer quarter wave layer of one of the dilectric mirrors. In this quarter wave layer, the index of refraction is different from that of the bulk film and this difference can be controlled by varying the geometry of the pattern (e.g. changing the pitch and width of the grooves). Varying the index of refraction of the cavity region changes its optical thickness thereby changing the pass-band of the interference filter. For the one-dimensional pattern in Fig. 2, the effective index of refraction is approximated as:

$$n_{\parallel}^2 = n_{Air}^2 f + n_{Layer}^2 (\Lambda - f) \tag{1}$$

for polarization parallel to the groove direction and

$$\frac{1}{n_{\perp}^2} = \frac{1}{n_{Air}^2} f + \frac{1}{n_{Layer}^2} (\Lambda - f) \tag{2}$$

for polarization perpendicular to the groove direction where Λ is the pitch of the groove and f is the fill factor. This relation assumes that there are no diffracted orders other than the zeroth order (hence, such such structures are often called zeroth order gratings) and has been established experimentally to be reasonably accurate for $\Lambda \leq \lambda/2$ [1]. In Figure 3, we plot the transmission curve as a function of the fill factor for a groove pitch of (Λ=400 nm). This figure is obtained from a multi-layer thin films computer code. The layer thicknesses of the quarter wave layers are chosen to have the pass-band for zero fill (f=1) to be centered at 1550 nm. From Fig. 3, it is seen that by changing the fill factor, the pass-band of the filter changes.

From (1) and (2), it is seen that the etched layer has an index of refraction that is different for differing polarizations. This birefringence can be quite large and has been measured experimentally (for patterned SiO_2) to be as much as $\Delta n = n_{\perp} - n_{\parallel} = .06$ [1]. Therefore, the use of a grooved spacer layer such as that shown in Fig. 2, yields a narrowband polarization filter with high polarization extinction ratios ($< 10^{-3}$) as is seen in Figure 4. By introducing two-dimensional patterns (but still maintaining the sub-wavelength periodicity in all symmetry directions), such as that shown in Figure 5, the polarization response of the filter can be changed and the birefringence eliminated. This is particularly important for optical sensor work where polarization independent behaviour is desired.

Structures with groove periods smaller than the wavelength such as that shown in Fig. 2 have been investigated for the design of ultra-low reflectance anti-reflection coatings [2-5]. However, a systematic study of the physics and engineering aspects of zeroth-order gratings is at present lacking.

References

1. D. C. Flanders, "Submicrometer periodicity gratings as artificial anisotropic dielectrics," *Appl. Phys. Lett.*, **42,** 492 (1993).
2. R. C. Enger and S. K. Case, "Optical elements with ultrahigh spatial-frequency surface corrugations," *Appl. Optics,* **22,** 3220 (1983).
3. Y. Ono, Y. Kimura, Y. Ohta and N. Nishida, "Antireflection effect in ultrahigh spatial-frequency holographic relief gratings," *Appl. Optics,* **26,** 1142 (1987).
4. D. H. Raguin and G. M. Morris, "Analysis of antireflection-structured surfaces with continuous one-dimensional surface profiles," *Appl. Optics,* **32,** 2582 (1993).
5. N. F. Hartman and T. K. Gaylord, "Antireflection gold surface-relief gratings: experimental characteristics," *Appl. Optics,* **27,** 3738 (1988).

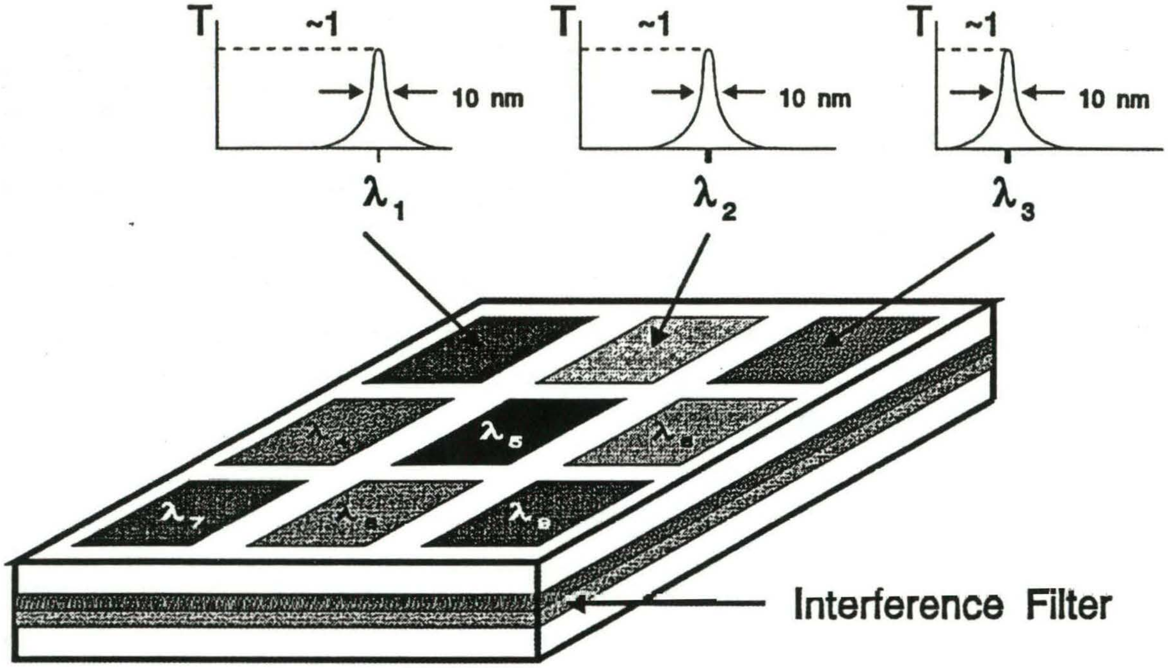


Figure 1: Operational diagram of a two-dimensional array filter.

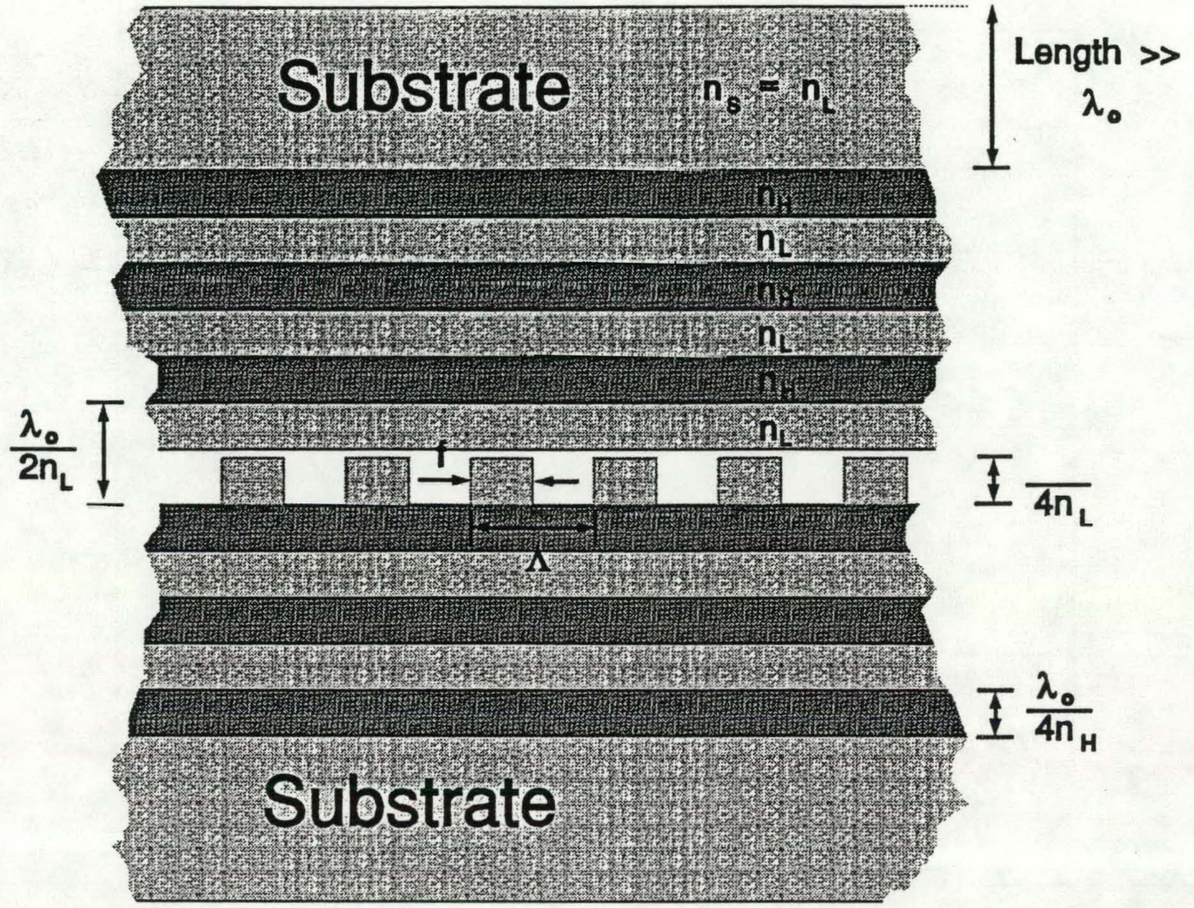


Figure 2: Schematic of a one-dimensional patterned interference filter.

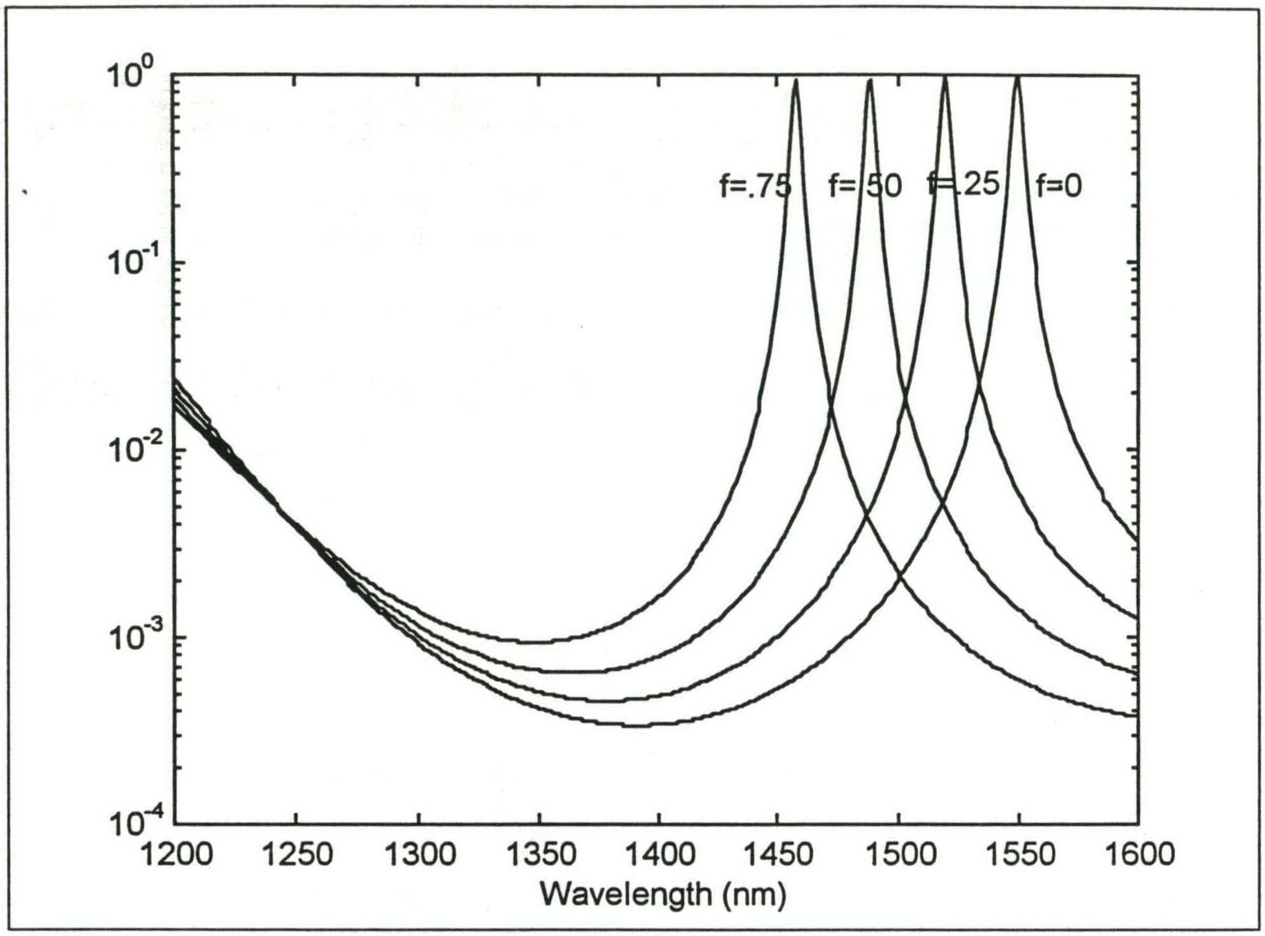


Figure 3: Transmission vs. Wavelength as a Function of *f* for Device in Fig. 2.

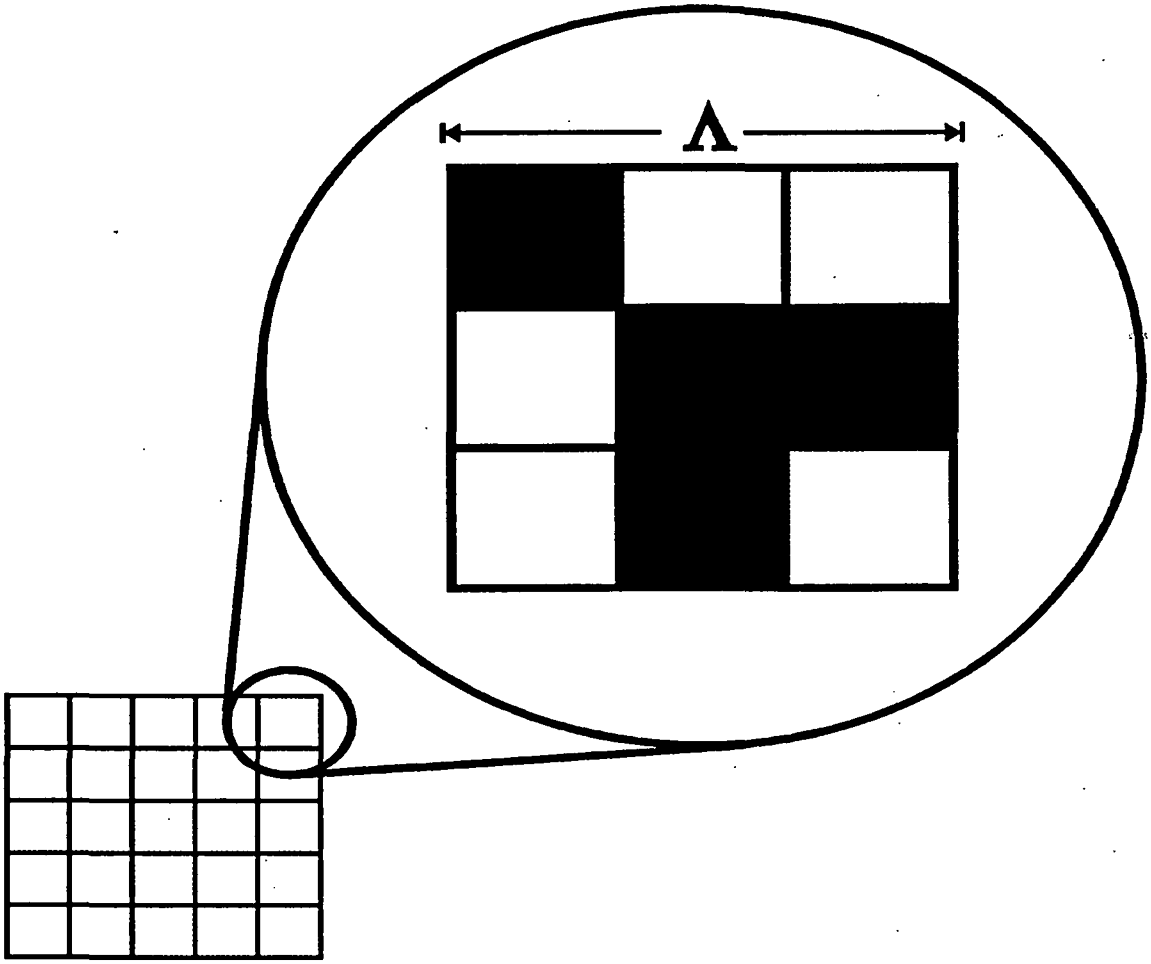


Figure 4: Example of a two-dimensional pattern. The pattern would appear on the top quarter-wave layer shown in Fig. 2.

Commercial Potential

We expect our optical interference filter to find a number of useful and potentially important industrial applications. We enumerate here some anticipated uses of our device

1. *Integrated Optical Sensors:* The optical filter when used in conjunction with a broadband semiconductor diode source and an array detector is a miniature spectrophotometer. Such a spectrophotometer can form the basis of a low-cost miniature optical sensor.

2. *Polarizing Interference Filters:* As described earlier, by appropriate patterning, the optical interference filter can be made polarization sensitive with extinction ratios better than than 10^{-3}. The filter described here is useful for applications requiring a minature low-cost narrowband filter (e.g. optical sensor and laser design).

Although the design of this product was not motivated by direct solicitation, we believe that there is a nascent market for this product and therefore, a commercial interest for the filter described here is expected. For example, presently Hamamatsu Photonics K.K. manufactures a commerical integrated spectrophotometer (S4111-S4114 series) in the 300 nm -960 nm spectral regime using a LVF and Si CCD array. In addition to companies such as Hamamatsu that specialize in photonics products, the interference filter array would also be of interest to custom detector manufacturers such as Santech Corp., EG&G Judson and Epitaxx Inc.. These companies would be motivated by potential applications to optical sensors.

An attractive feature of our design from the viewpoint of its commericialization is its potentially low manufacturing cost. Since the principal technical merit of our filter array is its small stize, we anticipate most applications using our design to capitalize on this particular aspect. The advantage of fabricating small structures is that a large number can be fabricated on an individual substrate. It is expected that over 200, 5x5 filter arrays (with a pixel size of 150 μm x 150 μm) can be fabricated on a single 2 inch substrate. For a typical 2 inch substrate (costing roughly $200), the nominal coating cost is roughly $1000 and an e-beam lithography cost is roughly $2000. Assuming that 200 arrays can be made on each substrate, the manufacturing cost per filter array, excluding packaging, can be well under $50. This cost can be further reduced if multiple substrates are coated per coating run (usually 10-15 two inch substrates can be coated on each run). This cost is to be compared with currently available filters such as LVF and CVF where natural economies of scale are lacking and the manufacturing cost per filter can be well over $100 (typically $300-$500).

Due its is potential uses and its relatively low manufacturing costs, we believe that the filter described in this TA has good commericialization potential. We suggest that a patent application be sought for this in United States. Since there are many photonics vendors in Japan and Europe who specialize in optical filters and detectors, it would be advisable to seek patents in Japan and a few of the European countries (e.g. Germany, England and France).

A Two Dimensional Array of Optical Interference Filters Produced by Lithographic Alterations of the Index of Refraction

S. Kaushik and B. R. Stallard

ABSTRACT

We describe a new concept for producing, on a single substrate, a two-dimensional array of optical interference filters where the pass-band of each element can be independently specified. The interference filter is formed by optically contacting two dielectric mirrors so that the top quarter-wave films of the two mirrors form a Fabry-Perot cavity having a half-wave thickness. In the new device, we propose to etch an array of sub-wavelength patterns into the top surface of one of the mirrors before forming the cavity. The patterns must have a pitch shorter than the operational wavelength in order to eliminate diffraction. By changing the index of refraction of the half-wave layer, or the optical thickness of the cavity, the patterning is used to shift the pass-band and form an array of interference filters. One approach to producing the array is to change the fill factor of the pattern. Once the filter array is produced it may be mated to a two-dimensional detector array to form a miniature spectrophotometer.

KEYWORDS: sub-wavelength structures, interference filter, spectrophotometer,

2. CURRENT TECHNOLOGY FOR FILTER ARRAYS

Optical interference filters appear in a number of forms. The class of filters with the highest transmission and narrowest linewidths are Fabry-Perot filters. These filters are formed by sandwiching a half-wave (or some multiple thereof) spacer layer between two mirrors. The lowest absorptive losses are found in mirrors made from multilayer dielectric stacks that are usually quarter-wave in thickness. The pass-band of these filters (i.e. the center wavelength, λ_0) is determined by the optical thickness of the spacer layer. The bandwidth is determined by both the optical thickness of the spacer layer and the reflectivities of the two mirrors. The most common Fabry-Perot filters commercially available have pass-bands that are spatially uniform with variations as small as 0.1%. Filters with spatially varying pass-bands are also commercially available and are manufactured by proportionally varying the thickness of all the layers uniformly across the surface of the filters. These filters, called wedge filters, are either circular (circular variable filter - CVF) or linear (linear variable filter - LVF). In the former, the thicknesses of the multilayer films vary linearly with the angular position on a circular substrate; in the latter, they vary linearly with position across a rectangular substrate.

Although wedge filters are widely used, they have several limitations. First, it is not possible to have both large dispersion (>> 10 nm/mm) and retain other desirable characteristics. The dispersion is proportional to the wedge angle; however, a large wedge angle results in poor transmission characteristics (due to non-normal reflections from the angled mirror facet), and an increase in spectral bandwidth (due to wavefront inhomogenity across the angled facets). Owing to these limitations, wedge filters are large (e.g. typical LVFs in the 300 nm - 700 nm spectral range are 60 mm in length) and have dispersion of about 5-10 nm/mm. CVFs are typically 25-50 mm in diameter and have a dispersion of 1 nm/degree. Significant improvements in these numbers are not expected; therefore, ultra-miniaturization of devices employing LVFs or CVFs is not possible with existing technology.

Another restrictive feature of wedge filters is that the pass-band can change only in a continuous manner across the substrate. In designing an optical sensors one may be interested only in a relatively small number of distinct, and non-uniformly spaced spectral bands. However, with a wedge filter one cannot pick and choose the desired channel unless one has a well collimated and uniform beam that encompasses all the desired channels. Since maintaining a good collimation along with acceptable light levels is difficult (especially if the channels are separated by 60 mm), it is necessary to either move the filter (e.g. a rotating CVF), or steer the beam. Neither of these strategies is well suited for extremely rapid data collection, especially if the desired channels are widely spaced throughout the spectrum.

Another impediment to the miniaturization of a spectrophotometer based on a wedge filter is that the filter is essentially a one-dimensional structure. Typically, pass-bands can be varied along only one dimension (linear position in LVF and azimuthal angle in CVF). A two-dimensional wedge filter is difficult to fabricate and to our knowledge, none is available commercially.

3. A NEW TECHNOLOGY

We describe here a new technique to produce on a single substrate, a two dimensional array of optical interference filters in which the pass-band at any spatial location on the substrate is a variable, user supplied parameter. A schematic diagram of the new type of filter array is shown in Fig. 1. This array is arranged in a pixel format similar to that of a two-dimensional photodetector array (e.g. silicon CCD array). This is to suggest that such a filter array, in conjunction with a two-dimensional detector array, can form a compact spectrophotometer. However, unlike conventional miniature spectrophotometers that utilize LVFs and CVFs, each channel can be individually assigned a different pass-band over a relatively large spectral range ($\Delta\lambda/\lambda_0 \sim 0.1$ to 0.35). In sensor applications, where information at only a few wavelengths is required, a spectrophotometer based on this type of filter array can be extremely miniature since the pixel size can be made as small as the dimensions of the detector pixel ($< 100\ \mu$m).

Our technique is to form an interference filter by optically contacting two dielectric mirrors, where a series of one and two dimensional periodic steps is lithographically patterned on the top surface of one of the dielectric mirrors. A sketch of our design is shown in Fig. 2. Each dielectric mirror (with three pairs of alternating layers of high index Si and low index SiO_2) has a reflectivity of approximately 98%. Note that both of the multilayer mirrors terminate with a low index layer. Since each of the layers is a quarter-wave thick, combining the two mirrors yields a cavity width, $\lambda_0/(2n_L)$, where n_L is the index of refractive of the low index material. The two mirrors on the top and bottom of this cavity form a prototypical optical interference filter that is well documented in the literature, with the pass-band centered at λ_0 and a bandwidth determined by the finesse.

The pass-band is changed by etching a periodic pattern onto the top surface of one of the mirrors. The patterned surface appears as a set of periodic grooves or steps in the outer quarter-wave layer of one of the dielectric mirrors. In this quarter-wave layer, the index of refraction is different from that of the bulk film and this difference can be controlled by varying the geometry of the pattern (e.g. changing the pitch and width of the grooves). Varying the index of refraction of the cavity region changes its optical thickness thereby changing the pass-band of the interference filter. For the one-dimensional pattern in Fig. 2, the effective index of refraction is approximated as:

$$n_{\parallel}^2 = n_{Air}^2 f + n_{Layer}^2 (\Lambda - f) \tag{1}$$

for polarization parallel to the groove direction, and

$$\frac{1}{n_{\perp}^2} = \frac{1}{n_{Air}^2} f + \frac{1}{n_{Layer}^2} (\Lambda - f) \tag{2}$$

for polarization perpendicular to the groove direction, where Λ is the pitch of the grooves and f is the dimension of the etched region (the fill factor, f, is defined as f/Λ, also, $0 \leq f \leq 1$, where f=0 indicates no etched region). These expressions assume that there are no diffracted orders other than the zeroth order (hence, such structures are often called zeroth order gratings). Eqs. (1) and (2) have been shown experimentally to be reasonably accurate for $\Lambda \leq \lambda/2$ [1]. In Fig. 3, we plot the transmission curve as a function of the fill factor for a groove pitch of 400 nm. This figure is obtained from a multilayer thin films computer code, assuming patterning on only one of the mirror surfaces. The layer thicknesses of the quarter wave layers are chosen to have the pass-band for f=0 centered at 1550 nm. From Fig. 3, it is seen that by changing the fill factor, the pass-band of the filter changes. The spectral range of the tuning can be readily shown to depend on the index of refraction of the cavity dielectric:

$$\frac{\Delta\lambda}{\lambda_0} = \frac{1}{4}\frac{(n-1)}{n} \ . \tag{3}$$

The range, $\Delta\lambda$, is 274 nm for Si, and 120 nm for SiO_2, when λ_0 is 1550 nm. This equation assumes that only half of the cavity is patterned. Twice the range is possible if one is willing to pattern the surface of both mirrors before forming the cavity.

From Eqs. (1) and (2), it is seen that the etched layer has an index of refraction that is different for differing polarizations. This birefringence can be quite large and has been measured experimentally (for a single layer of patterned SiO_2) to be as much as $\Delta n = n_{\perp} - n_{\parallel} = 0.06$ [1]. Therefore, the use of a grooved spacer layer such as that shown in Fig. 2, yields a narrow band polarization filter with polarization extinction ratios better than 10^3, as is seen in Fig. 4. By introducing two-dimensional patterns, but maintaining the sub-wavelength periodicity in all symmetry directions, such as that shown in Fig. 5, the birefringence of the filter can be eliminated. This is particularly important for optical sensor work where polarization independent behavior is normally desired.

Structures with groove periods smaller than the wavelength such as that shown in Fig. 2 have been investigated for anti-reflection coatings[2-5]. However, such work has not been extended, to our knowledge, to the area of interference filters, as suggested herein.

4. CONCLUSIONS AND COMMERCIALIZATION POTENTIAL

This new type of interference filter array overcomes the principal limitations of the wedge filter that were discussed in Section 2. In summary, the new technology has the following features: (1) individual array elements are spectrally flat, (2) pass-bands can be specified independently for each array element, (3) two dimensional arrays can be readily fabricated, and (4) miniaturization is readily achieved.

There are at least two uses for these devices. (1) *Integrated Optical Sensors.* The optical filter when used in conjunction with a broadband semiconductor diode source and an array detector is a miniature spectrophotometer. Such a spectrophotometer can form the basis of a low-cost miniature optical sensor. (2) *Polarizing Interference Filters.* As described earlier, by appropriate patterning, the optical interference filter can be made polarization sensitive with extinction ratios better than 10^3.

An attractive feature of our design from the viewpoint of its commercialization is its potentially low manufacturing cost. It is expected that over 200, 5x5 element arrays (with an element size of 150 μm x 150 μm) can be fabricated on a single 2 inch substrate. For a typical 2 inch substrate, costing \$200, the coating cost is about \$1000 and the e-beam lithography cost is about \$2000. Assuming that 200 arrays can be made on each substrate, the manufacturing cost per filter array, excluding dicing and mounting, is only \$16. This cost can be further reduced if multiple substrates are coated per coating run (usually 10-15 two inch substrates can be coated on each run). This cost is to be compared with currently available filters such as LVF and CVF where natural economies of scale are more limited and the retail cost is typically \$300-\$500.

5. ACKNOWLEDGMENTS

We have benefited from discussions with Mial Warren and Robert Smith at Sandia. Arnold Howard, Richard Corless, and others at Sandia are collaborating to fabricate these devices. The project is supported by the Laboratory Directed Research and Development Program. Sandia National Laboratories is operated by Lockheed Martin Corporation for the U. S. Department of Energy under contract DE-AC04-94AL8500.

6. REFERENCES

1. D. C. Flanders, "Submicrometer periodicity gratings as artificial anisotropic dielectrics," *Appl. Phys. Lett.,* **42,** 492 (1993).
2. R. C. Enger and S. K. Case, "Optical elements with ultrahigh spatial-frequency surface corrugations," *Appl. Optics,* **22,** 3220 (1983).
3. Y. Ono, Y. Kimura, Y. Ohta and N. Nishida, "Antireflection effect in ultrahigh spatial-frequency holographic relief gratings," *Appl. Optics,* **26,** 1142 (1987).
4. D. H. Raguin and G. M. Morris, "Analysis of antireflection-structured surfaces with continuous one-dimensional surface profiles," *Appl. Optics,* **32,** 2582 (1993).
5. N. F. Hartman and T. K. Gaylord, "Antireflection gold surface-relief gratings: experimental characteristics," *Appl. Optics,* 27, 3738 (1988).

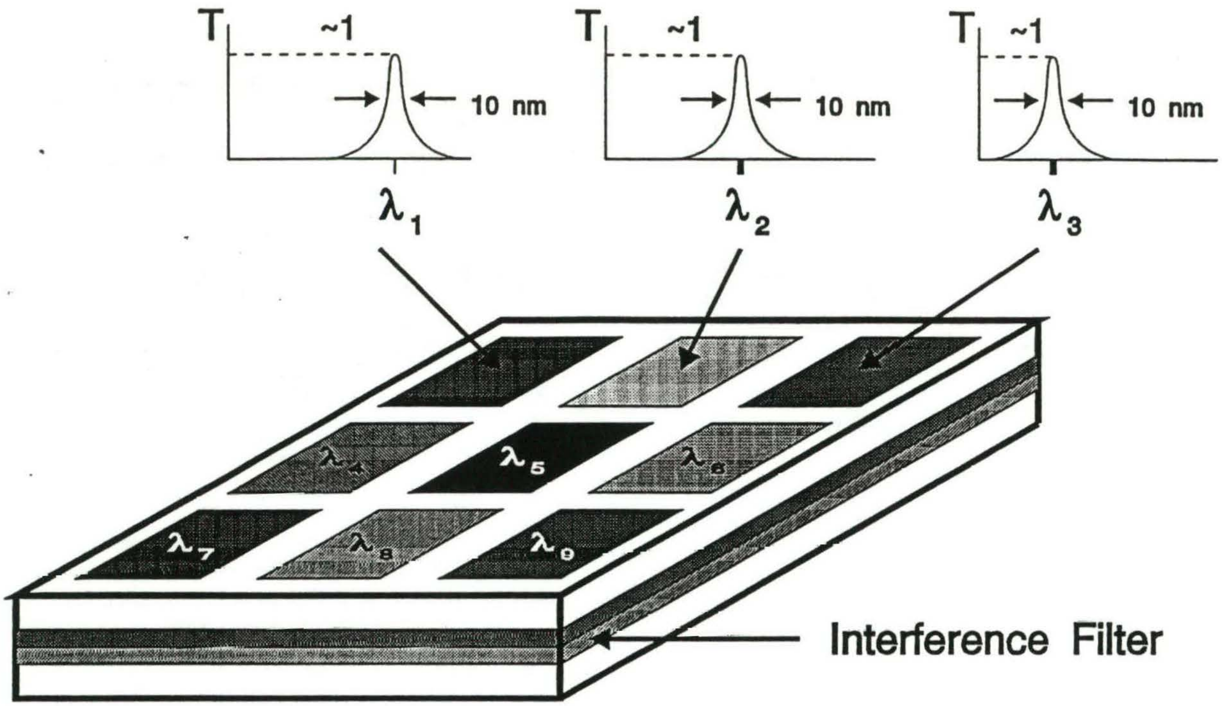


Figure 1. Schematic diagram of a two-dimensional filter array.

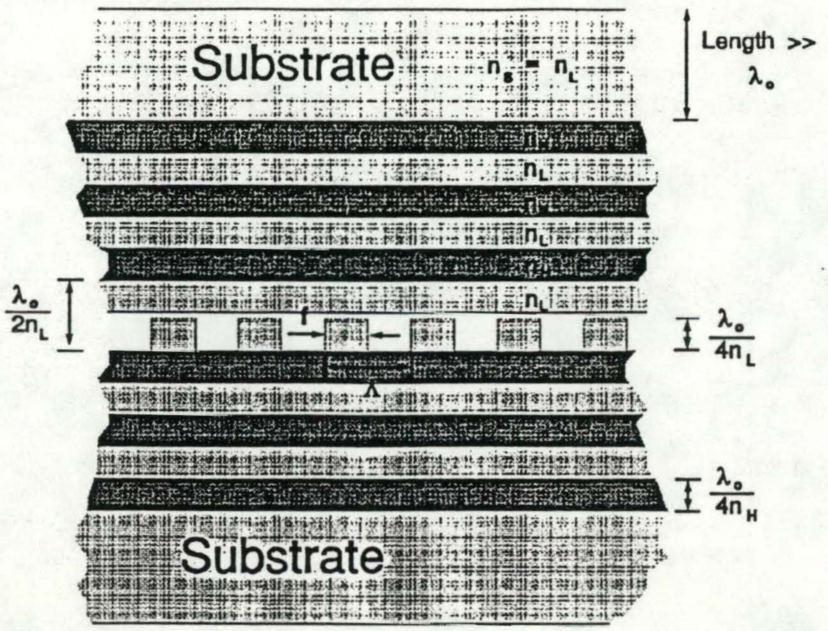


Figure 2. Cross-section of a Fabry-Perot filter that is tuned by sub-wavelength patterning. The high and low index materials, n_H and n_L, are assumed to be Si and SiO_2 for the calculations of Figs. 3 and 4.

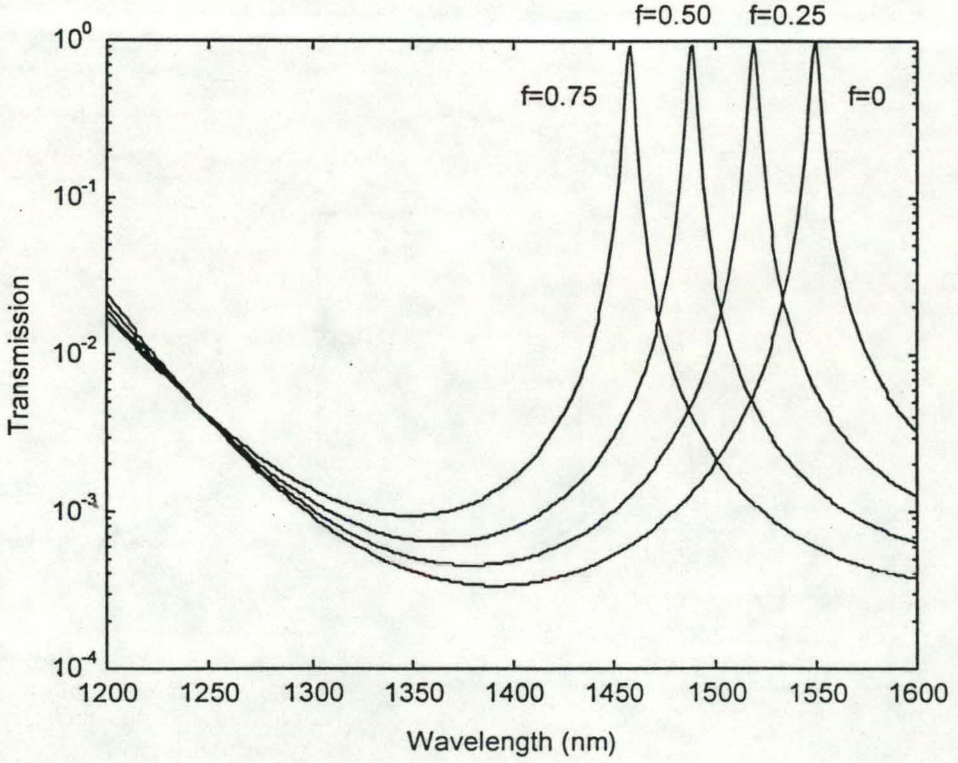


Figure 3. Calculated transmission vs. wavelength curves for the device in Fig. 2. Several values of the fill factor, f, from 0 to 0.75, are plotted. Shifts are double what is shown if patterning is on both mirror surfaces.

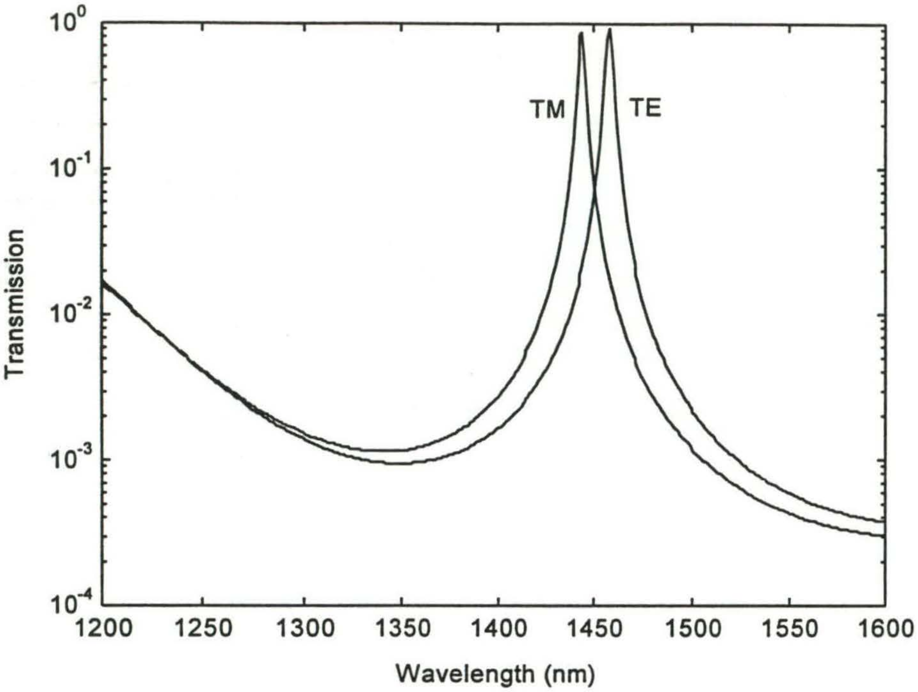


Figure4. Calculated transmission vs. wavelength curves for TE (parallel to grooves), and TM (perpendicular to grooves) polarization for the device in Fig. 2. The value of the fill factor is 0.75.

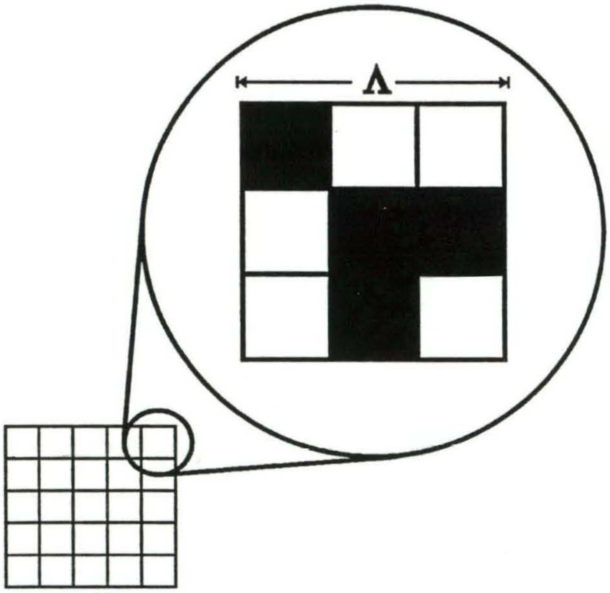


Figure5. Example of a two-dimensional pattern where the fill factor is varied by turning on or off the sub-pixels. The pitch, Λ, of the repeating pattern is chosen to be less than half the wavelength to avoid diffraction.

Fresnel Equations and Transmission Line Analogues for Diffraction Gratings

Sumanth Kaushik
Sandia National Laboratory, Albuquerque, NM 87185-1405

ABSTRACT

A simple and intuitive formalism is presented to describe diffraction in multi-layered periodic structures. We use the well known results from scalar analysis (wave propagation in homogeneous layered media) and show that they can be generalized rather readily to vector problems such as diffraction analysis. Specifically, we derive : (1) generalized Fresnel equations appropriate for reflection and transmission from an infinitely thick grating, (2) a generalized Airy formula for thin-film to describe reflection and transmission of light through a lamellar grating and (3) a matrix propagation method akin to that used for multi-layer thin film analysis.

The results developed here complement the recent work on R-matrix and S-matrix propagation algorithms that have been used in connection with modal and differential grating theories. These algorithms have proven to be numerically stable for calculating diffraction efficiencies from deep groove gratings. The formalism developed here expands upon the earlier literature by providing important details that are hitherto unavailable.

1 Introduction

Modal methods to calculate grating efficiencies have a voluminous literature [1]. Modal techniques are easy to understand conceptually, and easy to implement computationally. The underlying concept in these techniques is an expansion of the electric/magnetic fields in terms of an orthogonal basis set, solving for the expansion coefficients by matching boundary conditions and inverting the associated matrices [2, 3, 4]. Although simple to state and implement, many implementations of this technique suffer from numerical instabilities caused by arithmetic overflows arising from the presence of large exponentials in the Rayleigh expansion. This problem is especially severe for surface relief gratings and multilayer gratings where numerical errors propagated across layers lead to poor convergence [6].

The focus of this paper are "pure" modal methods (to contrast from related techniques such as coupled-mode or coupled-wave theories [1]) as applied to multi-layer and surface relief gratings. These classes of gratings have received considerable attention lately and efficient numerical methods have been developed recently to compute grating peformance [4, 6].

The purpose of this paper is not to add yet another modal technique to the rather large set of extant methods. Instead, the aim here is to distill the recent results in a form where: (1) the underlying physics is obvious and (2) the numerics for even the most sophisticated grating computation can be written down by inspection.

Propagation of light in a multi-layer thin film stack is describable in terms of cascaded 2×2 scattering/transfer and propagation matrices. The matrix elements are the reflection and transmission coefficients obtained from Fresnel equations. This method is standard, well understood and widely used in the optics community. The principal aim of this paper is to show that existing methods used in analysis of wave propagation in 1D stratified media are readily generalizable for inhomogeneous (but periodic) media. In the subsequent discussion, the inhomogeneous medium is assumed periodic. The results here do not apply directly to media with random or aperiodic spatial variations in the complex permittivity and susceptibility.

The principal results in this paper are: (1) the development of "vector" Fresnel equations for electromagnetic waves diffracting from an infinitely thick lamellar grating, (2) the development of "vector" Airy-formula for light propagating through a lamellar grating and (3) a generalized matrix propagation method for multi-layer and surface relief gratings that is similar in spirit to multi-layer optical thin film theory.

In our formulation, typical scalar quantities such as reflection and transmission coefficients found in standard 1D analysis, appear as infinite dimensional matrices. The rows of these matrices represent the different scattered orders and the columns correspond to the coupling between the different orders (i.e. the same mode coupling described in coupled-wave analysis). Table 1, for example, summarizes the generalized Fresnel reflection formulas or light reflecting from a periodic surface shown in Fig. 1. Using these generalized reflection and transmission coefficients,

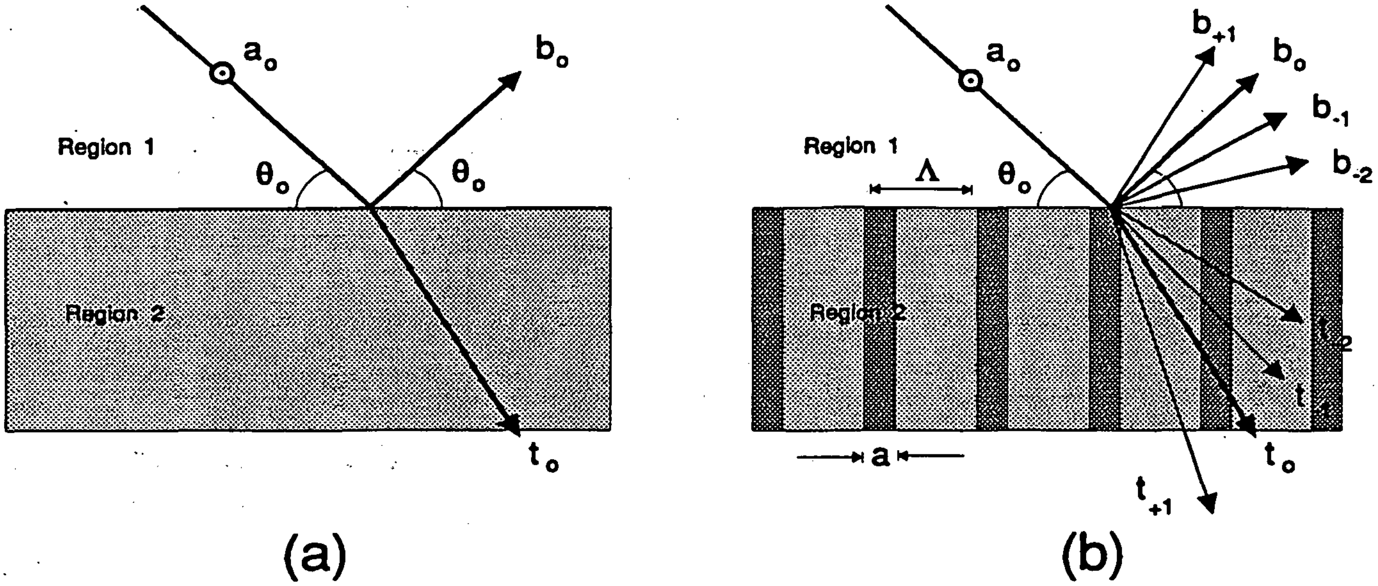


Figure 1: Reflection and transmission of light at the interface of: (a) an infinitely thick homogeneous layer with complex permittivities ϵ_2, μ_2; and (b) an infinitely thick lamellar grating with complex permittivities $\epsilon_2(x), \mu_2(x)$. Here, a_0 is the incident amplitude of the relevant field ($a_0 = E_{y0}$ for TE-polarization, and $a_0 = H_{y0}$ for TM-polarization). See text for further explanation of notation.

	Homogeneous Interface	Inhomogeneous Interface
Reflection	$\Gamma = \dfrac{\sqrt{\frac{\sigma_2}{\sigma_1}}k_{1z} - \sqrt{\frac{\sigma_1}{\sigma_2}}k_{2z}}{\sqrt{\frac{\sigma_2}{\sigma_1}}k_{1z} + \sqrt{\frac{\sigma_1}{\sigma_2}}k_{2z}}$	$\mathbf{\Gamma} = \dfrac{1}{\mathbf{\Sigma}_1^{-1}\mathbf{K}_{1z} + \mathbf{K}_{2z}\mathbf{\Sigma}_2^{-1}} \cdot \left[\mathbf{\Sigma}_1^{-1}\mathbf{K}_{1z} - \mathbf{K}_{2z}\mathbf{\Sigma}_2^{-1}\right]$
Transmission	$T = \dfrac{2k_{1z}}{\sqrt{\frac{\sigma_2}{\sigma_1}}k_{1z} + \sqrt{\frac{\sigma_1}{\sigma_2}}k_{2z}}$	$\mathbf{T} = \dfrac{2}{\mathbf{K}_{1z}\mathbf{\Sigma}_2 + \mathbf{\Sigma}_1\mathbf{K}_{2z}}\mathbf{K}_{1z}$
Power Flow	$S = \underbrace{k_{1z}\lvert a\rvert^2}_{\text{incident}} - \underbrace{k_{1z}\Gamma^2\lvert a\rvert^2}_{\text{reflected}}$	$S = \underbrace{\mathbf{a}^\dagger \cdot (\mathbf{K}_{1z} + \mathbf{K}_{1z}^*) \cdot \mathbf{a}}_{\text{incident}} - \underbrace{\mathbf{a}^\dagger \cdot \mathbf{\Gamma}^\dagger(\mathbf{K}_{1z} + \mathbf{K}_{1z}^*)\mathbf{\Gamma} \cdot \mathbf{a}}_{\text{reflected}}$

Table 1: Comparison of the generalized Fresnel formulas for light reflecting from an inifinitely thick lamellar grating (see Fig. 1b) with the "classical" Fresnel formulas for light reflecting from a homogeneous interface (see Fig. 1a). Bold face implies matrix and vector quantities and † represents Hermitian conjugate. See text for detailed description.

we show how an appropriate transfer matrix can be constructed for analyzing multi-layer structures. These will be discussed in greater detail in the subsequent sections.

The results in this paper bear a strong resemblence to the R- and S-matrix algorithms that have appeared recently [7, 6, 8, 9]; however the results are *not* identical. In contrast to earlier work, this paper approaches the diffraction problem from the viewpoint of physics. Unlike earlier work, mathematical features of the problem such as convergence, stability etc., are of secondary importance; although, as we will see later, the algorithms presented here do offer a few advantages over previous methods. We seek here instead, to draw an intimate connection between the grating problem and homogeneous multi-layer theory. It is hoped that the added physical intuition will facilitate further development in the mathematical treatment of diffraction.

The paper is organized as follows. In Sec. 2, we derive the reflection and transmission coefficients at an interface between two inhomogeneous interfaces and show that it is formally equivalent to the equations obtained between two homogeneous layers. In Sec. 3, these results are used to derive a set of generalized Fabry-Perot equations for a lamellar grating. In Sec. 4, an algorithm for multi-layer grating analysis is developed. In Sec. 5, we discuss some of the immediate consequences of our work and compare it with earlier work in this field.

2 Generalized Fresnel Equations for Periodic Stratified Media

We begin the analysis by considering an incident electromagnetic field on an infinitely thick grating as shown in Figure 1. The fields in Regions 1 and 2 can be expressed in terms of a modal expansion:

$$\begin{aligned} \sum_{l} \phi_l(x)(a_l + b_l) &= \sum_{q} \psi_q(x) t_q, \\ \sum_{l} \frac{1}{\sigma_1(x)} \phi_l(x)(\mathbf{k}_{1z})_l (a_l - b_l) &= \sum_{k} \frac{1}{\sigma_2(x)} \psi_q(x)(\mathbf{k}_{2z})_q t_q, \end{aligned} \tag{1}$$

where a_l, b_l and t_q are the mode amplitudes of the incident, reflected and transmitted fields respectively. The mode functions $\phi_l(x)$ and $\psi_k(x)$ are the eigenfunctions of the Maxwell's equations in Regions 1 and 2, with eigenvalues $(\mathbf{k}_{1z})_l$ and $(\mathbf{k}_{2z})_2$ respectively. These are defined in Appendix A. The functions $\sigma_1(x)$ and $\sigma_2(x)$ appear in Maxwell's equations and are

$$\begin{aligned} \sigma_i(x) &\rightarrow \mu_i(x) \quad \text{for TE Polarization,} \\ \sigma_i(x) &\rightarrow \epsilon_i(x) \quad \text{for TM Polarization.} \end{aligned}$$

The above mode expansion is standard and the reader is referred to the literature for further discussion [2, 3, 4].

To proceed further, we use well known properties of the mode functions $\phi_l(x)$ and $\psi_k(x)$ (see Appendix A) and define the four matrices $\mathbf{\Sigma_1}$, $\mathbf{\Sigma_2}$, $\mathbf{K}_{1z}$ and $\mathbf{K}_{2z}$ whose matrices elements are:

$$\begin{aligned} \left(\mathbf{\Sigma_1}^{-1}\right)_{ql} &= \langle\psi_q|\tfrac{1}{\sigma_1}|\phi_l\rangle \sim \sqrt{\frac{\sigma_2}{\sigma_1}}, \\ \left(\mathbf{\Sigma_2}^{-1}\right)_{ql} &= \langle\psi_q|\tfrac{1}{\sigma_2}|\phi_l\rangle \sim \sqrt{\frac{\sigma_1}{\sigma_2}}, \\ (\mathbf{K}_{1z})_{ll'} &= (\mathbf{k}_{1z})_l \delta_{ll'} \quad , \\ (\mathbf{K}_{2z})_{qq'} &= (\mathbf{k}_{2z})_q \delta_{qq'} \quad . \end{aligned} \tag{2}$$

We have used the bra-ket notation used in quantum mechanics where, for example, $\langle\psi|x\rangle \equiv \psi^{\dagger}(x)$ is the adjoint function (see Appendix A). In (2), we have also indicated the value taken by the matrix elements of the Σ matrices in the limit of a homogeneous media (for which the Σ matrices become diagonal). These matrices, along with the analytical inverses in Appendix A are, to the author's knowledge are new, and is one of the important contributions of this paper.

Using these matrices, (1) can be rewritten as

$$\begin{aligned} \Sigma_2^{-1} \cdot (\mathbf{a} + \mathbf{b}) &= \mathbf{t}, \\ \Sigma_1^{-1} \mathbf{K}_{1z} \cdot (\mathbf{a} - \mathbf{b}) &= \mathbf{K}_{2z} \cdot \mathbf{t}, \end{aligned} \tag{3}$$

where **a**, **b** and **t** are vectors whose elements are the mode amplitudes a_l, b_l and t_q respectively. Note that the vectors and matrices above are infinite dimensional since there are infinite number of eigenmodes that can be excited by the grating. The $\mathbf{\Sigma}_i$ matrices are complex and have non-zero off-diagonal matrix elements if and only if Region 2 is inhomogeneous, and are responsible for the coupling of the different scattering modes to the incident wave. The $\mathbf{K}_{iz}$ matrices are diagonal, but have complex numbers. The above matrix equation can be solved for **b** and **t** in terms of **a** to yield:

$$
\begin{aligned}
\mathbf{b} &= \left\{ \frac{1}{\mathbf{\Sigma}_1^{-1}\mathbf{K}_{1z} + \mathbf{K}_{2z}\mathbf{\Sigma}_2^{-1}} \cdot \left[\mathbf{\Sigma}_1^{-1}\mathbf{K}_{1z} - \mathbf{K}_{2z}\mathbf{\Sigma}_2^{-1}\right] \right\} \cdot a, \\
&\equiv \mathbf{\Gamma} \cdot \mathbf{a} \\
\mathbf{t} &= \left\{ \mathbf{\Sigma}_2^{-1} \frac{2}{\mathbf{\Sigma}_1^{-1}\mathbf{K}_{1z} + \mathbf{K}_{2z}\mathbf{\Sigma}_2^{-1}} \mathbf{\Sigma}_1^{-1}\mathbf{K}_{1z} \right\} \cdot \mathbf{a} \\
&\equiv \mathbf{T} \cdot \mathbf{a}.
\end{aligned}
\tag{4}
$$

With the aid of equations (3) and (4), we now derive an equation for power flow. This is calculated using Poynting's theorem, which is written

$$
S_{1z} = \frac{1}{2\Lambda} \int_0^{\Lambda} dx (\mathbf{E} \times \mathbf{H}^* + \mathbf{E}^* \times \mathbf{H}),
\tag{5}
$$

where we have followed standard procedure and averaged the power flow across the grating period Λ (see [3]). Substituting (1) and (3) into (5) and using the orthogonality relations summarized in Appendix A, we get

$$
\begin{aligned}
\mathbf{S}_{1z} &= \mathbf{a}^{\dagger} \cdot (\mathbf{K}_{1z} + \mathbf{K}_{1z}^{*}) \cdot \mathbf{a} - \mathbf{a}^{\dagger} \cdot \mathbf{\Gamma}^{\dagger} (\mathbf{K}_{1z} + \mathbf{K}_{1z}^{*}) \mathbf{\Gamma} \cdot \mathbf{a}, \\
\mathbf{S}_{2z} &= \mathbf{a}^{\dagger} \cdot \mathbf{T}^{\dagger} (\mathbf{K}_{2z} + \mathbf{K}_{2z}^{*}) \mathbf{T} \cdot \mathbf{a}.
\end{aligned}
\tag{6}
$$

In (6), in order to simplify notation, we have absorbed all numerical and physical constants into the field amplitude vector **a**.

We indentify the first term in $\mathbf{S}_{1z}$ with the power incident and the second with the power reflected. The transmitted power $\mathbf{S}_{1z}$ given above is strictly valid only for dielectric gratings where continuity of power flow demands that S be constant over all space (i.e. $\mathbf{S}_{1z} = \mathbf{S}_{2z} = Constant$). For conductive/lossy gratings, the orthogonality relations cannot be used to simplify the integration in (5) and $\mathbf{S}_{2z}$ varies with z.

However, this is not a serious issue from the viewpoint of computation. For infinitely thick conductive gratings, only $\mathbf{S}_{2z}(z = 0) = \mathbf{S}_{1z}$ is important as it represents the power absorbed/dissipated by the lossy grating. For gratings with finite thickness, one can calculate (5) in the free-space region (Regions 1 and 3 shown in Figure 3 in Sec. 3) where the orthogonality relations are recovered and where (6) is again applicable. Thus, the power absorbed by the grating would simply be

$$
\mathrm{Power}_{abs} = \mathbf{S}_{1z} - \mathbf{S}_{3z}
$$

Therefore, all measurable quantities can be computed without calculating $\mathbf{S}_{2z}$. However, in order to quantity truncation errors, an explicit evaluation of $\mathbf{S}_{2z}$ has been found to be useful [3]. This issue is beyond the scope of this paper and the reader is referred to the literature for a deeper discussion.

In Table 1, the principal results of this section are summarized and compared with the "classical" Fresnel formulae appropriate for light impinging on a homogeneous interface. The equations are the same form for both the homogeneous as well as the inhomogeneous problem. The principal difference is that for the latter, the reflection and transmission coefficients are matrices whereas for the homogeneous problem, these are scalar quantities. This merely reflects that fact that only specular reflection and transmission is possible from a homogeneous surface. It is easy to show that in the limit that the grating period becomes infinite (i.e. homogeneous), the matrices become diagonal and the reflectivity and transmissivity obtained by the matrix formulation are identical to the scalar formulae.

3 Generalized Airy-Formula for Lamellar Gratings

In this section, we will apply the results of Sec. 2 to a lamellar grating with finite thickness. For light propagating through a homogeneous layer (with uniform complex permittivity), the reflection and transmission coefficients are given by the well known Airy thin-film formula which are described in a number of texts (for example, see [11]). For a slab shown in Fig. 2a (where, for simplicity, Region 1 and 3 are the same media), with reflectivities r_1, r_2 and transmissivities t_1, t_2 (where the indices refer to the media from which the beam is incident), the reflected and transmitted amplitudes are given by [11]:

$$\begin{aligned} \Gamma &= r_1 + \frac{t_1 t_2 r_2 \exp(-2j\phi)}{1 - r_2^2 \exp(-2j\phi)}, \\ t &= \frac{t_1 t_2 \exp(-2j\phi)}{1 - r_2^2 \exp(-2j\phi)}. \end{aligned} \tag{7}$$

To derive similar expressions for a lamellar grating such as that shown in Fig. 2b, we write the equations for the reflected and transmitted fields at each of the interface and solve a set of linear matrix equations. To derive the results, we use Fig. 3 that is often used in "classical" Fabry-Perot analysis. Here, the quantities $\mathbf{a}_i$ and $\mathbf{b}_i$ $(i = 1, 2)$ are vectors whose elements are the mode amplitudes of the incoming and outgoing fields at each of the two interfaces.

The overall transmission and reflection coefficients will be expressed in terms of the transmission and reflection at each of the two interfaces. Table 1 provides the coefficients for the beam incident from Region 1. However, in addition to these coefficients, it is also necessary to compute the coefficients when the beam is incident from Region 2. This corresponds to t_2 and r_2 in (7). These coefficients can be calculated by setting up the complementary problem,

$$\begin{aligned} \sum_q \psi_q(x)(a'_q + b'_q) &= \sum_l \phi_l(x) t'_l \\ \sum_q \frac{1}{\sigma_2(x)} \psi_q(x)(\mathbf{k}_{2z})_q (a'_l - b'_l) &= \sum_l \frac{1}{\sigma_1(x)} \psi_l(x)(\mathbf{k}_{1z})_l t'_l, \end{aligned} \tag{8}$$

and solving for the vectors $\mathbf{b}'$ and $\mathbf{t}'$ in terms of $\mathbf{a}'$, we arrive at the following expressions:

$$\begin{aligned} \mathbf{\Gamma}_2 &= -\mathbf{\Sigma}_2^{-1} \mathbf{\Gamma}_1 \mathbf{\Sigma}_2 \\ \mathbf{T}_2 &= \mathbf{\Sigma}_2 \mathbf{T}_1 \mathbf{K}_{1z}^{-1} \mathbf{\Sigma}_1 \mathbf{K}_{2z}. \end{aligned} \tag{9}$$

Here $\mathbf{\Gamma}_1$ and $\mathbf{T}_1$ are the same coefficients defined in Sec. 2. In the limit of a homogeneous slab, the relations reduce to $\mathbf{\Gamma}_2 = -\mathbf{\Gamma}_1$ and $\mathbf{T}_2 = \mathbf{T}_1 \mathbf{K}_{2z} / \mathbf{K}_{1z}$. This corresponds exactly to the scalar result. In short, the reciprocal coefficients used in scalar theory generalize rather directly to the vector problem. These relations, to the knowledge of this author, is new and represents an important contribution of this paper.

Using $\mathbf{\Gamma}_2$ and $\mathbf{T}_2$ defined in (9) and $\mathbf{\Gamma}_1$ and $\mathbf{T}_1$ defined in Sec. 2, the equations that relate $\mathbf{a}_i$ and $\mathbf{b}_i$ at the interfaces are:

$$\begin{aligned} \mathbf{b}_1 &= \mathbf{\Gamma}_1 \mathbf{a}_1 + \mathbf{T}_2 \mathbf{a}_2 & \qquad \mathbf{b}_3 &= \mathbf{\Gamma}_2 \mathbf{a}_3 + \mathbf{T}_1 \mathbf{a}_4 \\ \mathbf{b}_2 &= \mathbf{T}_1 \mathbf{a}_1 + \mathbf{\Gamma}_2 \mathbf{a}_2 & \qquad \mathbf{b}_4 &= \mathbf{T}_2 \mathbf{a}_3 + \mathbf{\Gamma}_1 \mathbf{a}_4. \\ \\ \mathbf{a}_2 &= \exp(-j\mathbf{K}_{2z} d) \mathbf{b}_3 & \qquad \mathbf{a}_3 &= \exp(-j\mathbf{K}_{2z} d) \mathbf{b}_2 \\ &\equiv \mathbf{P}_0 \mathbf{b}_3 & \qquad &\equiv \mathbf{P}_0 \mathbf{b}_2 \end{aligned} \tag{10}$$

The equations for $\mathbf{a}_2$ and $\mathbf{a}_3$ relate the field amplitudes at one end of the slab to the field amplitudes at the opposite end through a simple propagation matrix. We note that $\mathbf{a}_i$ and $\mathbf{b}_i$ are infinite dimensional vectors, which in any practical calculations are to be truncated. However, in all our analysis, no assumption of truncation will be made and all results will be exact.

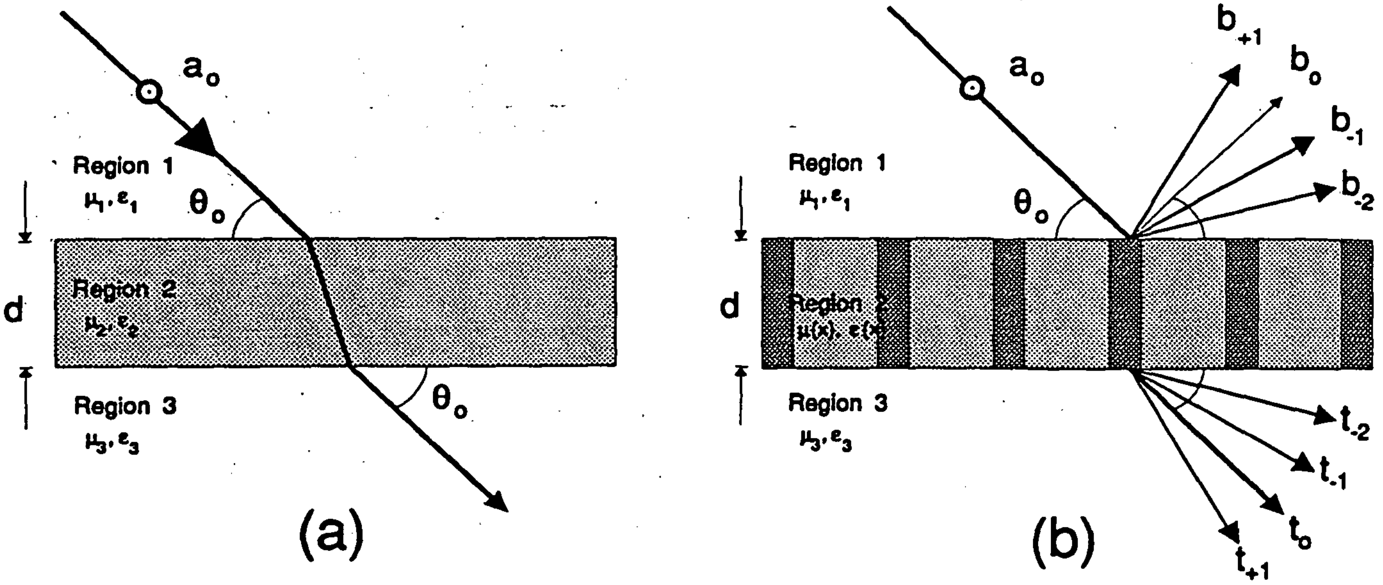


Figure 2: Reflection and transmission of light at the interface of: (a) an homogeneous layer with complex permittivities ϵ_2, μ_2 and a thickness d; and (b) a lamellar grating with complex permittivities $\epsilon_2(x), \mu_2(x)$ and thickness d. Here, a_0 is the incident amplitude of the relevant field ($a_0 = E_{y0}$ for TE-polarization, and $a_0 = H_{y0}$ for TM-polarization). See text for further explanation of notation.

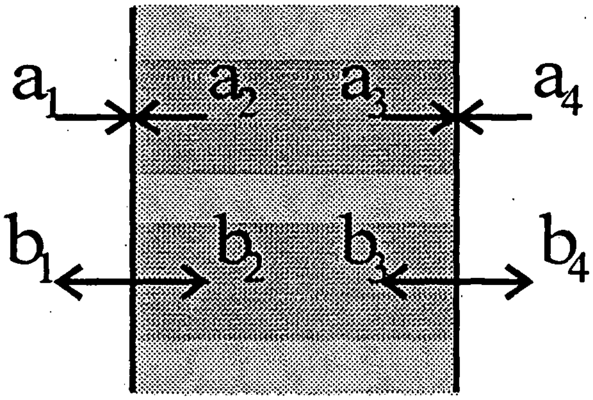


Figure 3: Incoming and outgoing (vector) fields at the interfaces of a lamellar grating. Here, $\mathbf{a}_i$ and $\mathbf{b}_i$ are vectors whose elements are the mode amplitudes of the incoming and outgoing fields respectively. These vectors are related via matrix equations detailed in the text.

The solution we seek is to express $\mathbf{b}_1$ and $\mathbf{b}_4$ in terms of $\mathbf{a}_1$ and $\mathbf{a}_4$. Setting $\mathbf{a}_4 = 0$, gives the overall reflection and transmision coefficients for a wave propagating from left to right. Equation (10) can be solved to yield the following expressions for $\mathbf{b}_1$ and $\mathbf{b}_4$:

$$\begin{aligned} \mathbf{b}_1 &= \left\{ \Gamma_1 + \mathbf{T}_2\mathbf{P}_0 \frac{1}{1 - \Gamma_2\mathbf{P}_0\Gamma_2\mathbf{P}_0} \Gamma_2\mathbf{P}_0\mathbf{T}_1 \right\} \mathbf{a}_1 \\ &\equiv \tilde{\Gamma}\mathbf{a}_1 \\ \mathbf{b}_4 &= \left\{ \mathbf{T}_2\mathbf{P}_0 \frac{1}{1 - \Gamma_2\mathbf{P}_0\Gamma_2\mathbf{P}_0} \mathbf{T}_1 \right\} \mathbf{a}_1 \\ &\equiv \tilde{\mathbf{T}}\mathbf{a}_1. \end{aligned} \tag{11}$$

We see from (11) that the overall reflection and transmission coefficients $\tilde{\Gamma}, \tilde{\mathbf{T}}$ for the lamellar grating are identical in form to that of the homogeneous slab given in (7).

The simple structure of (11) suggests immediately the underlying physics. The term $(1 - (\Gamma_2\mathbf{P}_o)^2)$ is merely the geometric sum obtained by the summation of the infinite series

$$\frac{1}{1 - (\Gamma_2\mathbf{P}_o)^2} = \sum_{n=0}^{\infty} (\Gamma_2\mathbf{P}_o)^{2n}.$$

This series accounts for the inifinite multiple reflections of both incident beam and *all of the diffracted components.* It should be noted that the Σ matrices do not commute and consequently, the order of the matrix is extremely important. However, the order is physically intutive and therefore can be written by inspection (as we will do in the next section for multi-layer gratings where all the three regions are different).

In Table 2, we have summarized the principal results of this section. The similarities between the scalar (homogeneous slab) and vector problem (lamellar grating) are quite apparent. We now apply these results to formulate a theory of multi-layer gratings.

4 Multi-layer and Surface Relief Diffraction Gratings

In the previous section, we developed a modal theory for a lamellar grating. Here, we extend the results for cascaded gratings. Such an analysis is appropriate for surface relief gratings where the gratings can be approximated as a series of stacked lamellar gratings [5, 4]. This is illustrated in Fig. 4 where a surface relief is divided vertically into a set of N slabs where each slab is assumed to be a lamellar grating. The diffraction efficiencies are computed by developing a suitable algorithm to cascade the (vector) reflection and transmission coefficients from each of the layers. This approach has proven quite satisfactory and the reader is referred to the literature for a resume on this method [5, 4]. Here, we describe a physically intutive, but exact method to cascade multi-layer lamellar gratings.

In scalar multi-layer thin film theory, a transfer matrix is defined at each interface to relate the fields from one side of the interface to the other. An overall transfer matrix is then calculated by multiplying the individual transfer matrices, from which the overall reflectance and transmittance is extracted. The transfer matrix approach is however not well suited for grating analysis and leads to numerical instabilities. This is well documented in the literature.

An alternative approach that has been recently developed is the R- and S-matrix propagation algorithm [6, 8, 9]. In these algorithm, one propagates the impedance/admittance of the layers, rather than the field itself. The impedance/admittance is a better behaved quantity which appears to have none of the numerical instability problems associated with the transfer matrix method. Although the S-matrix algorithm as applied to grating problems is relatively new, its use in scalar theory is quite well established (though the name has not been widely used in the optics community). For example, equation (7) is one of the simplest applications of the algorithm.

For applying the S-matrix approach to the vector problem, we refer to Figures 3 and 4. We assume the surface relief grating is divided into M (not necessarily equal) layers. Each layer is modelled as a lamellar grating. At each interface i, we define a pair of reflection and transmission matrices $\gamma_i^{\pm}, \mathbf{t}_i^{\pm}$ as described in Sec. 2. Here, we use the

	Homogeneous Slab	Lamellar Grating
Reflection	$\tilde{\Gamma} = r_1 + \frac{t_1 t_2 r_2 P_0^2}{1 - r_1^2 P_0^2}$	$\tilde{\mathbf{\Gamma}} = \mathbf{\Gamma}_1 + \mathbf{T}_2 \mathbf{P}_0 \frac{1}{1 - \mathbf{\Gamma}_2 \mathbf{P}_0 \mathbf{\Gamma}_2 \mathbf{P}_0} \mathbf{\Gamma}_2 \mathbf{P}_0 \mathbf{T}_1$
Transmission	$\tilde{T} = \frac{t_1 t_2 P_0}{1 - r_2^2 P_0^2}$	$\tilde{\mathbf{T}} = \mathbf{T}_2 \mathbf{P}_0 \frac{1}{1 - \mathbf{\Gamma}_2 \mathbf{P}_0 \mathbf{\Gamma}_2 \mathbf{P}_0} \mathbf{T}_1$
Stokes (reciprocity) Relations	$r_2 = -r_1$ $t_2 = t_1 k_{2z} / k_{1z}$	$\mathbf{\Gamma}_2 = -\mathbf{\Sigma}_2^{-1} \mathbf{\Gamma}_1 \mathbf{\Sigma}_2$ $\mathbf{T}_2 = \mathbf{\Sigma}_2 \mathbf{T}_1 \mathbf{K}_{1z}^{-1} \mathbf{\Sigma}_1 \mathbf{K}_{2z}$

Table 2: Comparison of the generalized Airy relations for light reflecting from a lamellar grating (Fig. 2b) with "classical" formulae for light reflecting from a homogeneous slab (see Fig. 2a). Bold faces implies matrix and vector quantities. The matrix propagation is given by $\mathbf{P}_0 = \exp(-j\mathbf{K}_{2z}d)$ and the scalar propagation by $P_0 = \exp(-jk_{2z}d)$. See text for detailed description

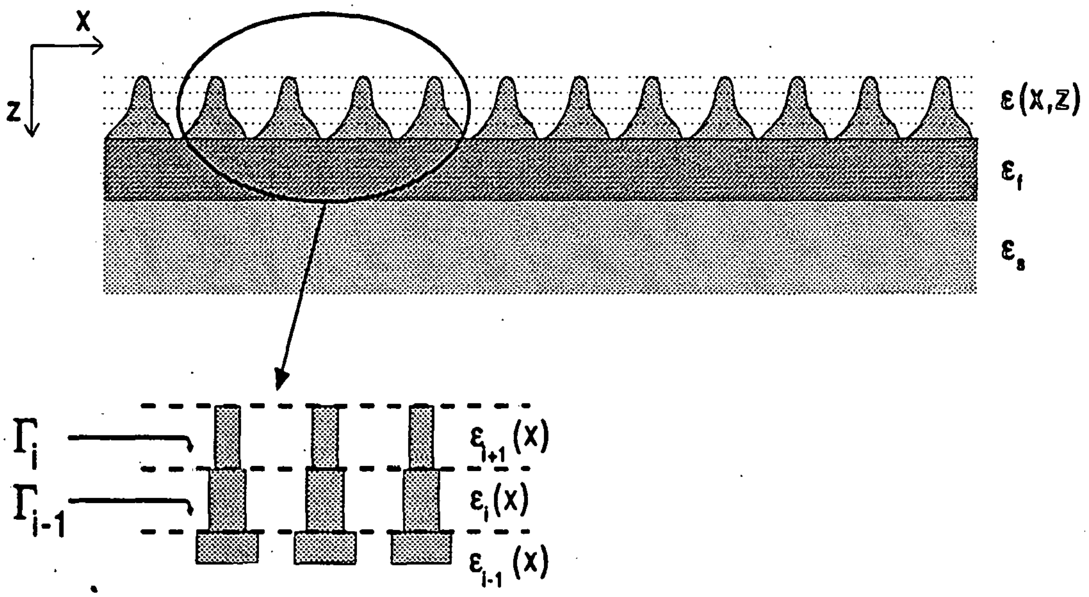


Figure 4: Diffraction from a surface grating is modelled as diffraction from a stack of lamellar gratings. The surface relief grating is divided into M layers with each being modelled as a lamellar grating (i.e. $\epsilon(x) \rightarrow \{\epsilon_1(x), \epsilon_2(x) \cdot \epsilon_M(x)\}$. Diffraction efficiencies for the entire structure are computed by recursively calculating the reflection matrix $\mathbf{\Gamma}_i$ using the reflection matrix from the adjacent layer $\mathbf{\Gamma}_{i-1}$. The recursive algorithm for this procedure is detailed in the text.

notation ± to refer to the direction of the wave propagation. The reflection and transmission matrices as seen just above the interface, $\bar{\mathbf{\Gamma}}_i^+$ and $\bar{\mathbf{T}}_i^+$, are expressed in terms of $\gamma_i^\pm, \mathbf{t}_i^\pm$ and the reflection matrix $\bar{\mathbf{\Gamma}}_{i-1}^+$ seen at the next sub-layer $i-1$. The matrix $\bar{\mathbf{\Gamma}}_i$ is then combined with the $\gamma_{i+1}^\pm, \mathbf{t}_{i+1}^\pm$ matrices to yield $\bar{\mathbf{\Gamma}}_{i+1}^+$ and $\bar{\mathbf{T}}_{i+1}^+$. This procedure is repeated till $i = M$, at which point $\bar{\mathbf{\Gamma}}_M^+$ and $\bar{\mathbf{T}}_M^+$ correspond to the overall reflection and transmission of the structure.

The matrices $\bar{\mathbf{\Gamma}}_i^+$ and $\bar{\mathbf{T}}_i^+$ can be calculated from $\gamma_i^\pm, \mathbf{t}_i^\pm$ and $\bar{\mathbf{\Gamma}}_{i-1}^\pm$ by setting up the following equations that are very similar to (10).

$$\begin{aligned}
\mathbf{b}_1 &= \gamma_i^+ \mathbf{a}_1 + \mathbf{t}_i^- \mathbf{a}_2. & \mathbf{b}_3 &= \bar{\mathbf{\Gamma}}_{i-1}^+ \mathbf{a}_3 \\
\mathbf{b}_2 &= \mathbf{t}_i^+ \mathbf{a}_1 + \gamma_i^- \mathbf{a}_2 & \mathbf{b}_4 &= \bar{\mathbf{T}}_{i-1}^+ \mathbf{a}_3 \\
\mathbf{a}_2 &= \exp(-j\mathbf{K}_{2z,i} d)\mathbf{b}_3 & \mathbf{a}_3 &= \exp(-j\mathbf{K}_{2z,i} d)\mathbf{b}_2 \\
&\equiv \mathbf{P}_i \mathbf{b}_3 & &\equiv \mathbf{P}_i \mathbf{b}_2.
\end{aligned} \tag{12}$$

The overall reflectivity when viewed from the ith layer is given by expressing $\mathbf{b}_1 = \bar{\mathbf{\Gamma}}_i^+ \mathbf{a}_1$ and $\mathbf{b}_4 = \bar{\mathbf{T}}_i^+ \mathbf{a}_1$. It is straighforward to solve the above matrix equations to yield

$$\begin{aligned}
\bar{\mathbf{\Gamma}}_i^+ &= \gamma_i^+ + \mathbf{t}_i^- \mathbf{P}_i \bar{\mathbf{\Gamma}}_{i-1}^+ \mathbf{P}_i \frac{1}{1 - \gamma_i^- \mathbf{P}_i \bar{\mathbf{\Gamma}}_{i-1}^+ \mathbf{P}_i} \mathbf{t}_i^+ \\
\bar{\mathbf{T}}_i^+ &= \bar{\mathbf{T}}_{i-1}^+ \mathbf{P}_i \frac{1}{1 - \gamma_i^- \mathbf{P}_i \bar{\mathbf{\Gamma}}_{i-1}^+ \mathbf{P}_i} \mathbf{t}_i^+.
\end{aligned} \tag{13}$$

As one would expect, these equations are identical in form to the scalar Fabry-Perot equations. It should be pointed out that whereas (13) is similar *in spirit* to the equations found in Ref. [8], they are not identical. It is apparent from comparing (13) with equation (12) of Ref. [8] that the latter requires several additional matrix operations. A more detailed discussion of this subject will be forthcoming in future publications.

5 Discussion

The recursive algorithm (13) described in the previous section is numerically stable under almost any conditions. It can only fail under rare circumstances when the denominator in (13) becomes singular. Such scenarios are rare and correspond to the excitation of local resonances (e.g. surface plasmons, etc). Additionally, the equations are applicable for any incident polarization. Although the derivations above have assumed non-conical mounting, extension of our algorithm to conical mounting is also straightforward and will be forthcoming in future publications.

The recursive algorithm given by (13) is similar to the algorithm outlined in Ref. [8]. Although Li refers to the algorithm as an "R-matrix" algorithm, it is infact a misnomer and would be more appropriately called S-matrix [10]. In Ref. [8], no direct comparison is made to the R-matrix algorithm in Ref. [6]. In principal, both algorithms are exact treatments of the diffraction problem (in the limit of infinite stratification layers). As we will see below, both algorithms are based on the same physics; however, they differ in their numerical properties. It is therefore useful to examine these differences in greater detail.

The choice of basis set used in this paper is a plane wave basis set. The choice of a plane wave basis set allows us to express the propagation matrix $\mathbf{P}_i$ simply as a diagonal matrix. The use of other basis sets (e.g. sinusoidal basis set used in Ref. [6]), although related to the plane wave set through a similarity transform, does not allow for a simple identification with the scalar results.

But beyond pedagogy, the use of alternate basis sets can lead to disaster if one is not careful about the implementation. For example, the algorithm presented in Ref. [6], can potentially breakdown when layer thicknesses become zero or are half-integral multiples of the internal propagation wave numbers. The reason for this can be understood by considering the physics of the problem.

The R-matrix algorithm is successful because it is based on propagating impedances/admittances across the interfaces. Conceptually, this is akin to the calculation of input impedances of a multi-segment transmission line with multiple loads. Therefore, consider a simple transmission line with characteristic impedance Z_0 that is terminated by a load impedance Z_L. The impedance at some point to the left of the load, for example, $z = -l$, is given as [13]

$$Z(z=-l) = Z_0 \frac{(Z_L - Z_0)\exp(jkl) + (Z_L - Z_0)\exp(-jkl)}{Z_L + Z_0)\exp(jkl) + (Z_L + Z_0)\exp(-jkl)} = Z_0 \frac{Z_L + jZ_0 tan(kl)}{Z_0 + jZ_L tan(kl)}. \tag{14}$$

However, notice that whereas the first equation behaves sensibly with all values of l, the second equation, due to the present of the $\tan(x)$ function, exhibits unpredictable behavior for values where the tangent function becomes infinite. Of course, since the numerator and denominator approach infinity at the same rate, the ratio is a well defined quantity for all l. However when the $\tan(x)$ terms appear in matrices, the ratio between the numerator and denominator, due to numerical and truncation errors, can lead to potentially unpredicatable behaviour near the singularity.

In Ref. [6], the matrix-elements of the so-called "sector matrices" (counterpart to the $\gamma_i^{\pm}$ and $\mathbf{t}_t^{\pm}$, are defined in terms of $\cot(x)$ and $\csc(x)$ functions. These functions arise naturally from the sinusoidal basis set used in the analysis. When the arguments of the functions become integral multiples of 2π (e.g. thickness of grating layer much smaller than the wavelength), the sector-matrices may become ill-conditioned leading to poor convergence of the algorithm. In fact, the algorithm produces non-sensical results when the thickness of a layer is zero.

To contrast this, equation the S-matrix treatment (i.e. eqn. (13)) has no singular functions in its definition, either in the numerator, or the denominator. In fact, it can be shown after some tedious algebraic manipulation that the (13) can be cast in the form similar to the first equation in (14). Therefore, the transmission and reflection coefficients behave sensibly for all values of layer thickness.

One final point that should be noted is that (13) appears to be rather well suited for approximations. For example, if the term $\gamma_i^- \mathbf{P}_i \bar{\Gamma}_{i-1}^+ \mathbf{P}_i$ is assumed to be small, then one can expand the denominator by series, keeping only the first term. This approximation leads to the solution proposed by Pai and Awada [5] and also corresponds to the Bremmer series expansion [6, 12].

6 Conclusion

We have shown in this paper an equivalence between scalar and vector theories of reflection and transmission. Specifically, we have derived : (1) a set of generalized Fresnel equations appropriate for reflection and transmission from an infinitely thick grating, (2) a generalized Airy formula for describing reflection and transmission of light through a lamellar grating and (3) a matrix propagation method akin to matrix methods used for multi-layer thin film analysis.

The formalism developed here is directly applicable for calculating diffraction efficiencies of surface relief gratings. The recently developed S-matrix propagation algorithm appears to be closely related to results presented here. This paper exapnds on the earlier work by providing important inversion formulas and Stokes relations that were lacking in the earlier work. The recursive algorithm presented in this paper is similar to earlier work, but appears to be more computationally efficient.

A direct numerical comparison of our algorithm with earlier algorithms was not within the scope of this effort. It will be forthcoming in future publications.

7 Acknowledgements

The author acknowledges helpful discussions with Drs. Lifeng Li (at U. Arizona) and Brian Stallard (at SNL). The author also thanks Zach Wade for assisting in the figures. The project is supported by the Laboratory Directed Research and Development Program. Sandia National Laboratories is operated by Lockheed Martin Corporation for the U.S. Department of Energy under contract DE-AC04-94AL8500.

A Appendix

The following is a brief summary of the important properties of the eigenfunction expansion used in this paper. A more complete discussion can be found in the literature [2, 3, 4]. Maxwell's equations, after appropriate separation of variables, can be written as

$$E(x,z)=E(x)\exp(-jk_z z),$$

where k_z is the eigenvalue of the following equation

$$\left[\sigma(x)\frac{d}{dx}\frac{1}{\sigma(x)}\frac{d}{dx}+k_0^2\epsilon(x)\right]F_l(x)=k_z^2F_l(x), \tag{15}$$

where $\sigma(x)$ and $\epsilon(x)$ are periodic in x (i.e. $\sigma(x+\Lambda)=\sigma(x), \epsilon(x+\Lambda)=\epsilon(x)$). Here $\sigma(x)$ is defined to be

$$\begin{array}{lcll}\sigma_i(x) & \rightarrow & \mu_i(x) & \text{for TE Polarization,}\\ \sigma_i(x) & \rightarrow & \epsilon_i(x) & \text{for TM Polarization.}\end{array}$$

where $\epsilon(x)$ and $\mu(x)$ are the spatially varying (complex) permittivities of the layer.

In order to express aribtrary fields in terms of the modal fields above, it is also necessary to solve the adjoint problem:

$$\left[\bar{\sigma}(x)\frac{d}{dx}\frac{1}{\bar{\sigma}(x)}\frac{d}{dx}+k_0^2\bar{\epsilon}(x)\right]F_l^+(x)=(k_z^+)^2F_l^+(x), \tag{16}$$

where $\bar{\sigma}$ indicates complex conjugate and $+$ is used to signify adjoint.

It can be shown that the eigenfunctions $\{F_l\}$ and $\{F_l^\dagger\}$ are bi-orthonormal and complete (see [2, 3, 4]). Bi-orthonormality is expressed in terms of the inner product

$$\langle F_l|\frac{1}{\sigma(x)}|F_{l'}\rangle \equiv \int_0^\Lambda dx\frac{1}{\sigma(x)}\bar{F}_l^\dagger(x)F_{l'}(x)=\delta_{ll'},$$

where $\delta_{ll'}$ is the Kronecker delta. Completeness implies that a function $h(x)$ satisfying the periodic boundary conditions may be expanded as

$$h(x)=\sum_{l=-\infty}^{\infty}\langle F_l|\frac{1}{\sigma(x)}|h\rangle F_l(x),$$

which, borrowing notations from quantum mechanics, can be written as

$$\mathbf{1}\equiv\sum_{l=-\infty}^{\infty}\frac{1}{\sigma(x)}|F_l\rangle\langle F_l|,$$

where $\mathbf{1}$ is the identity operator.

Using these orthogonality properties, we establish the following matrix identities for the Σ matrices defined in Sec. 2.

$$\begin{aligned}(\boldsymbol{\Sigma}_1)_{lq} &= \langle\phi_l|\frac{1}{\sigma_2}|\psi_q\rangle,\\ (\boldsymbol{\Sigma}_2)_{lq} &= \langle\phi_l|\frac{1}{\sigma_1}|\psi_q\rangle.\end{aligned}$$

These were derived by demanding $\boldsymbol{\Sigma}_1{}^{-1}\boldsymbol{\Sigma}_1=\mathbf{1}$ and $\boldsymbol{\Sigma}_2{}^{-1}\boldsymbol{\Sigma}_2=\mathbf{1}$.

References

[1] T. K. Gaylord and M. G. Moharam, "Analysis and Applications of Optical Diffraction by Gratings," *Proc. of IEEE*, pp.894-935 (1985).

[2] L.C Botten, M.. S. Craig, R. C. McPhedran, J. L. Adams and J.R. Andrewartha, "The dielectric lamellar diffraction grating," *Optica Acta*, **28**, pp. 413-428 (1981).

[3] L.C Botten, M.. S. Craig, R. C. McPhedran, J. L. Adams and J.R. Andrewartha, "Finitely conducting lamellar diffraction grating," *Optica Acta*, **28**, pp. 1087-1102 (1981).

[4] L. Li, "A modal analysis of lamellar diffraction gratings in conical mountings," *J. Mod. Optics*, **40**, pp. 553-573 (1993).

[5] D. M. Pai and K. A. Awada, "Analysis of dielectric gratings of arbitrary profiles and thicknesses," *J. Opt. Soc. Am. A*, **8**, pp. 755-762 (1991).

[6] L. Li, "Multilayer modal method for diffraction gratings of arbitrary profile, depth, and permittivity," *J. Opt. Soc. Am. A*, **10**, pp. 2581-2591 (1993).

[7] L. F. DeSandre and J. M. Elson, "Extinction-theorem analysis of diffraction anomalies in overcoated gratings." *J. Opt. Soc. Am. A*, **8**, pp. 763-777 (1991).

[8] L. Li, "Bremmer series, R-matrix propagation algorithm, and numerical modeling of diffraction gratings," *J. Opt. Soc. Am. A*, **11**, pp. 2829-2836 (1994).

[9] F. Montiel and M. Neviere, "Differential theory of gratings: extension to deep gratings of arbitrary profile and permittivity through the R-matrix propagation algorithm," *J. Opt. Soc. Am. A*, **11**, pp. 3241-3250 (1994).

[10] L. Li, Personal Communication

[11] W.H. Steel, *Interferometry*, Cambridge Univ. Press, (1983).

[12] H. Bremmer, "The WKB approximation as the first term of a geometric series," *Commun. Pure Appl. Math*, **4**, pp. 105-115, (1951).

[13] L. C. Shen and J. A. Kong, *Applied Electromagnetism*, PWS Publishers, (1983).

Disclosure of Technical Advance (TA)

"An Optical Probe with sub-Rayleigh Lithographic Alterations of the Index of Refraction."

Sumanth Kaushik

Description

The resolution R of a far-field optical instrument is determined by the wavelength λ_0 and the numerical aperture of the light gathering optics. The resolution also depends on the nature of the light source. Theoretical formulae for the resolution is found in standard optics texts. For illumination from a coherent light source $R=\lambda_0/(2\text{ N.A.})$ (Abbe limit). The resolution for illumination with an incoherent source is $R=.61\lambda_0/\text{N.A.}$ (Rayleigh limit). For an optical micrsosocope, the highest obtainable resolution at optical wavlengths (e.g. an oil immersion microscope with λ_0 = 500 nm, and N.A=1.45) is 210 nm (incoherent source) and 180 nm (coherent source).

By resolution, we mean the smallest resolvable distance between two lines of equal thickness. In the formulae given above, it is assumed these lines are part of grating with equal line and space. Thus, to say a microscope has a resolution of 200 nm, means that using the microscope, one can tell the difference between a homogeneous surface and a grating with a 400 nm pitch. Gratings with pitches smaller than 400 nm will appear as an homogeneous surface.

The resolution of optical instruments as given by the Abbe and Rayleigh limits is fundamental. The only improvements to resolution are shorter wavelength sources and higher numerical aperture optics. Improvements in light sources is certainly to be expected (e.g. deep UV and x-ray sources), however, its use in commerical instruments is not expected any time in the near future. Improvements in numerical aperture is limited by available materials and signficant improvements here is also not expected any time in the near future. Therefore, for practical purposes, the resolution of conventional optical instruments appears to have reached a fundamental limit.

Further increases in resolution of *optical systems*, however, is critical for a number of industrial and research applications. Non-optical systems such as electron micrscopes, provide much higher resolution than optical instruments and are readily available; however, optical diagnostics are: (1) operable under ambient conditions, (2) non-invasive and (3) cheaper to operate than these alternative methods. Two applications where an optical high resolution instrument is critical are semiconductor manufacturing and optical data storage.

In the next two to three years, the semiconductor industry will bring into production, devices with feature sizes of .35 μm. Present day optical microscopes do not provide an adequate resolution for the characterization of devices with such small feature sizes. In semiconductor manufacturing applications, non-invasive, *in situ* diagnostics are critical and existing non-optical instruments such as electron microscopes are a inconvenient and expensive solution

The optical data storage industry is faced with a similar technological hurdle. Higher data storage density demands higher resolution optics. Even with anticipated advances in available laser sources (e.g. blue-green lasers) and beam shaping optics (e.g. super-resolving systems), storage densities much larger than 2.5 - 3 Gbits/ inch^2 cannot be expected in the next decade using existing data storage techniques (presently only .5 Gbits/ inch^2 is a commercially viable technology).

A promising strategy to improve the resolution of optical instruments is to use the information obtained in near-field. The Abbe and Rayleigh limits are fundamental limits only in the far-field (distances larger than tens to hundreds of wavelengths from the object). In the near-field, the resolution is potentially unlimited, constrained only by the intensity of the source and signal-to-noise on the detectors.

Near-field optical systems has demonstrated a resolution of 50 nm (λ/12). . Laboratory near-field systems have demonstrated storage densities in the excess of 45 Gbits/inch^2. In fact, a near field microscope with resolution of 50 nm is available commerically -- providing high resolution at moderate cost.

However, several technical hurdles have prevented optical near-field techniques from gaining widespread usage. These are: (1) close working range (< 50 nm), (2) tips that are both expensive, difficult to fabricate and break often, and (3) poor signal collection efficiency leading to slow data collection rates. Efficient solutions to these problems are critical if near-field based instruments are to play a significant role in the industrial workplace.

Described here is a new near-field probe that is based on Bragg grating structures and optical waveguides. The probe has a comparative advantage over existing near-field tips in that it can operate at a distance of 150-200 nm with excellent signal to noise and is readily amenable to fast data collection. Preliminary calculations show that our design can provide resolutions of 100-200 nm (at an illumination wavelength of λ_0=1500 nm) with superior signal-to-noise.

The design described here is shown in Fig. 1. The near-field tip is a circular Bragg grating patterned on a thin layer deposited on a surface. Similar to Fabry-Perot resonators formed by linear Bragg gratings, the circular gratings form a two-dimensional resonator where the resonant frequency is determined by the width of the central zone and the pitch of the gratings. An incident beam (which can be either out-of-plane, or in-plane) can couple signficant energy only if the frequency of the incident light is at or near the resonant frequency Ω_0. Energy that is coupled into the resonator is measured by detectors placed at the end of the waveguides that collect the in-plane light emanating from the resonator.

Near-field optical measurement is done by bringing the central zone of the circular grating in close proximity (< 150-200 nm) to the object being interrogated. The object, possessing an index of refraction different from that of air, changes the effective index of the central zone. A change in the index of refraction implies a change in the cavity resonant frequency $\Omega_0+\Delta$. The physics is schematically illustrated in Fig. 2. Provided that the change in frequency is larger than the bandwidth of the resonance, the amount of external light energy (at frequency Ω_0) that can couple into the cavity decreases. This decrease is measured by the in-plane detectors at the end of the output waveguides. The probe described here is termed as a circular distributed bragg resonator rear-field probe (CDBR) since it is closely related to the physics of distributed Bragg reflectors (DBR) which has been well studied in the literature.

A schematic of a typical measurement is illustrated in Fig. 3. In Fig. 3, we have assumed an out-of-plane illumination, but an in-plane source is equally feasible. The wavelength of the source is chosen to be the resonant frequency of the resonator. The circular resonator is placed above the surface to be interrogated. The surface is characterized by randomly placed "pits." As the resonator is scanned above the surface, the measured signal in the in-plane detectors increases and decreases depending on whether a pit lies directly below the central zone of the resonator. As the pit moves away from the central zone, the signal increases to a maximum value. As a pit approaches the central zone, the signa monotonically decreases. Thus, the presence or absence of the pit is determined by the signal on the detectors.

The term "pits" is appropriate for optical data storage applications, but for other applications, it merely means some surface feature that is distinct from its neighbors. The principal requirement for a successful detection/measurement is a sufficient contrast in the index of refraction between the "pit" and the surrounding region.

Compared with current near-field probes, the principal advantages of the CDBR near-field probe described here are: (1) can be used at much greater distance, (2) superior contrast ratio and signal collecting efficiency, (3) owing to its "flat architecture," it is well suited to exploit "fly" technology used in magnetic head designs, (4) readily amenable for parallel data collection. By using several CDBR probes near each other, it is possible to have a near-field system which does not require precise tip or height control. This eliminates the need for servo-control and can enhance data collection rates significantly.

As mentioned above, presently near-field optical techniques are not promising owing to: (1) poor signal to noise, (2) poor data rates, and (3) complicated control machinery. The four advantages described above for the CDBR probe, are critical if near-field techniques are to play an important role in industrial diagonostics and for optical data storage.

In Figs. 4 -6, results of some hypothetical surfaces are presented based on computational results. In Fig. 4, random bit are assumed to be patterned on to a surface. The bit pattern consists of a pit being formed at each of the locations marked **1**. The pit is assumed to be circular with a diameter of 200 nm and a depth of 100 nm, and formed on a silicon surface. The resonator is assumed to be 100 nm from the surface. Thus, a bit sequence **1-0-1** represents two adjacent pits separated by 200 nm.

Figure 5 plots a typical output from the in-plane detectors as the resonator is scanned over the surface. Assuming the incident frequency is fixed at the resonant frequency Ω_0, the resonator only couples appreciable energy when the pit is directly below the central zone (representing an intensity maximum). All other time, owing the presence of the surface (silicon has a very high index of refraction), the resonator does not couple appreciable energy (hence the minimum). As it can be seen, the contour plots reveal clearly the presence of the pits. The contrast ratio is better than 100:1.

In Figure 6, the same surface is scanned in a slightly different way. Rather than monitoring the detector outputs, the laser frequency is tuned instead. By measuring the detuning of the laser frequency, the same information can be gleaned. This configuration could perhaps be useful for microscopy applications where the frequency information could provide added information to the surface composition.

As can be seen, the resolution of this probe is well less than 200 nm and with excellent contrast. Although the illustrated resolution is only 200 nm, the performance of this probe is significantly better than conventional far-field optics having similar resolution. The principal improvement is in the contrast ratio. The contrast is nearly two orders of magnitude better than conventional optics. In fact, for a contrast ratio of 2 (the typical number used to obtain the Rayleigh and Abbe limits), the resolution is better than 75 nm.

To see this, we have plotted in Figure 7, the intensity at the in-plane detectors as a pit is moved away from the origin. Here, the origin is when the center of the pit coincides with the center of the circular resonator. This corresponds to the maximum signal. As the pit moves away from the center, the region directly under the central zone is replaced with higher index material (in this case, silicon with n=3.42) and consequently, the resonator detunes and the signal drops.

The resonator is characterized in terms of a *Q* or Quality-Factor. The *Q* is related to the finesse of a Fabry-Perot cavity and is related to the pitch of the grating, the number of grooves in the grating, and the width of the central zone. An explicit computation of the cavity *Q* has not been done for the circular resonator geometry owing to a lack of available theoretical formulae. We have chosen values for *Q* that we believe are readily manufacturable. These estimates are based on what is readily available for linear Fabry-Perot structures. In Fig. 6, for a *Q*=200, we have contrast of 50. For a *Q*=500, this contrast improves to 300. In typical optical data storage system, a contrast ratio of about 50-60% is sufficient for detection. This indicates that the probe has apotential resolution of well below 100 nm.

A resolution less than 100 nm would represent a six-ten fold increase in resolution that can be available from optical systems (even with anticipated advances in shorter wavelength sources). Additionally it would mean that some of the more sophisticated data modulation schemes (e.g. RLL) can be used to get higher data density. Many of the sophisticated data encoding schemes cannot be applied on existing systems owing to insufficient contrast and resolution.

It is important to note that the results presented here assume a probe laser centered at about 1500 nm. Typical optical storage systems use lases at 780 nm, and recently 630 nm. The relatively long wavelegth used in the present analysis was motivated by: (1) easing fabrication tolerances, and (2) favorable optical properties of Silicon. The grating pitch is roughly $\lambda_0/2$; consequently, longer the wavelength, the easier it is to fabricate the gratings. Silicon has both a high index of refraction and negligible absorption at 1500 nm. Owing to the ready availibility of Si fabrication facilities at Sandia National Labs, it appears to be a good material for fabricating the resonator. However, designing the resonator for shorter wavelength lasers can lead to still higher resolution than quoted here.

The ideas described here are original and have not appeared elsewhere. However, the use of circular and linear resonators in other applications is not new. Circular resonators meeting fabrication tolerances of the sort needed for this application have been fabricated by others and is documented in the literature. Attached to the technical advance are some important references in the field. However, the use of distributed Bragg resonators for near-field optics, to the best knowledge of the author, has not appeared in the published literature.

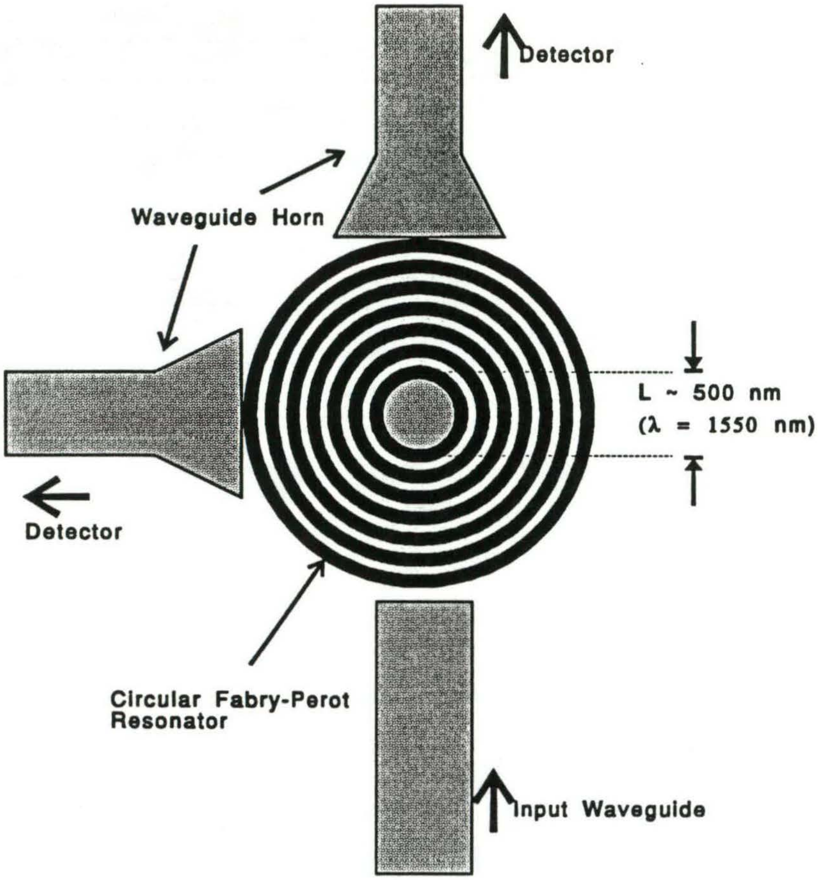


Figure 1: Schematic of the Near-Field Probe

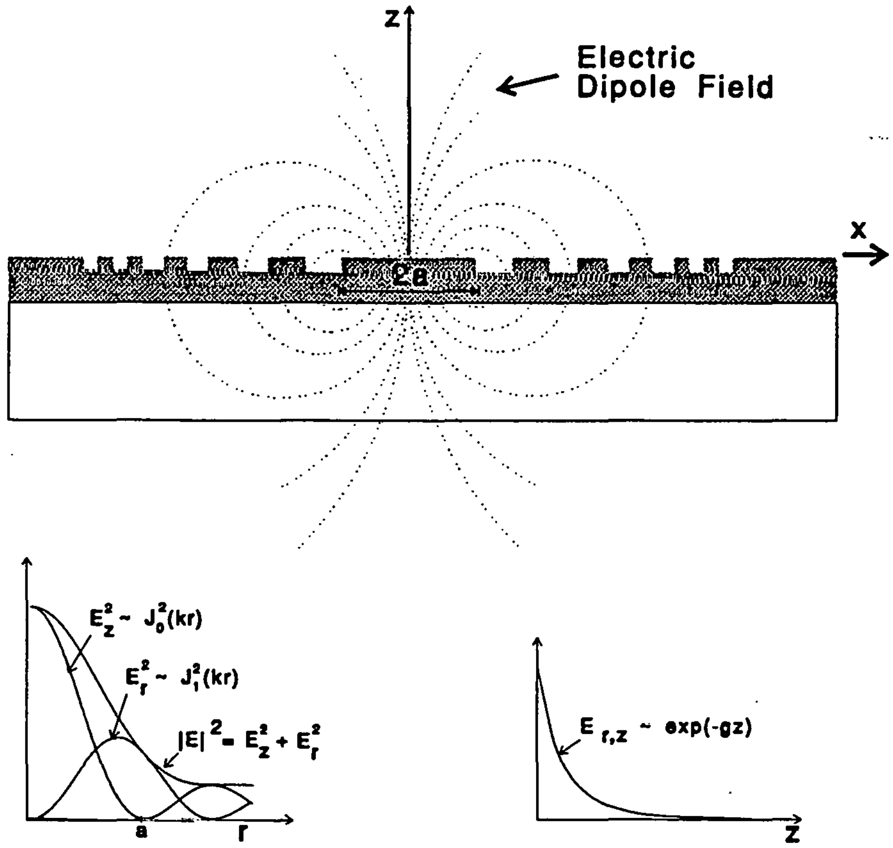


Figure 2: Physics of CDBR Near-Field Probe

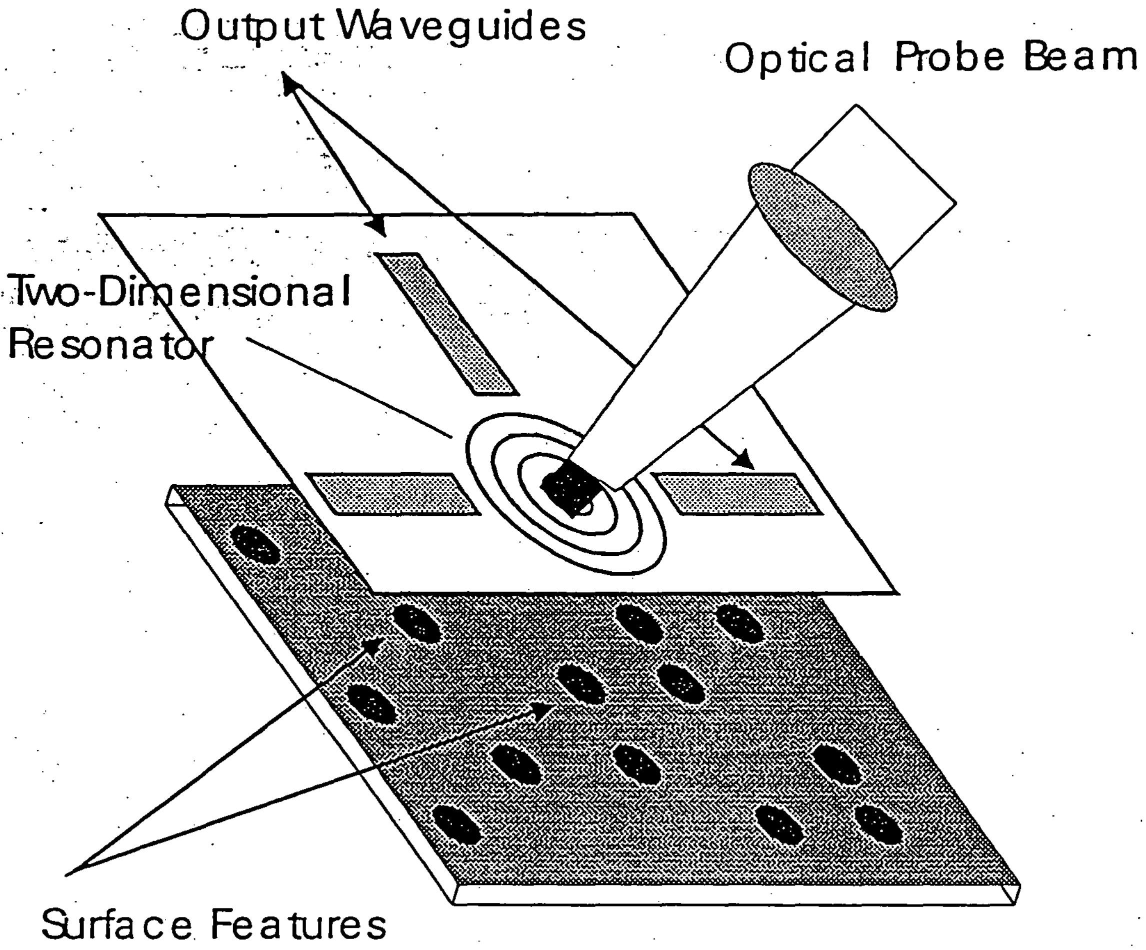


Figure 3:Schematic of a typical CDBR Application

```
FF =

     0     0     1     0     1     0     1     0     0
     0     0     0     0     0     0     0     0     0
     0     0     0     0     0     0     0     0     1
     0     0     0     0     0     0     0     0     0
     0     0     0     0     1     0     0     0     0
     0     0     0     0     0     0     0     0     0
     1     0     0     0     0     0     0     0     0
     0     0     0     0     0     0     0     0     0
     0     0     0     0     1     0     0     0     0
```

Figure 4: A random "pit" pattern. Each 1 corresponds to a pit of diameter 200 nm. A zero implies an unperturbed surface.

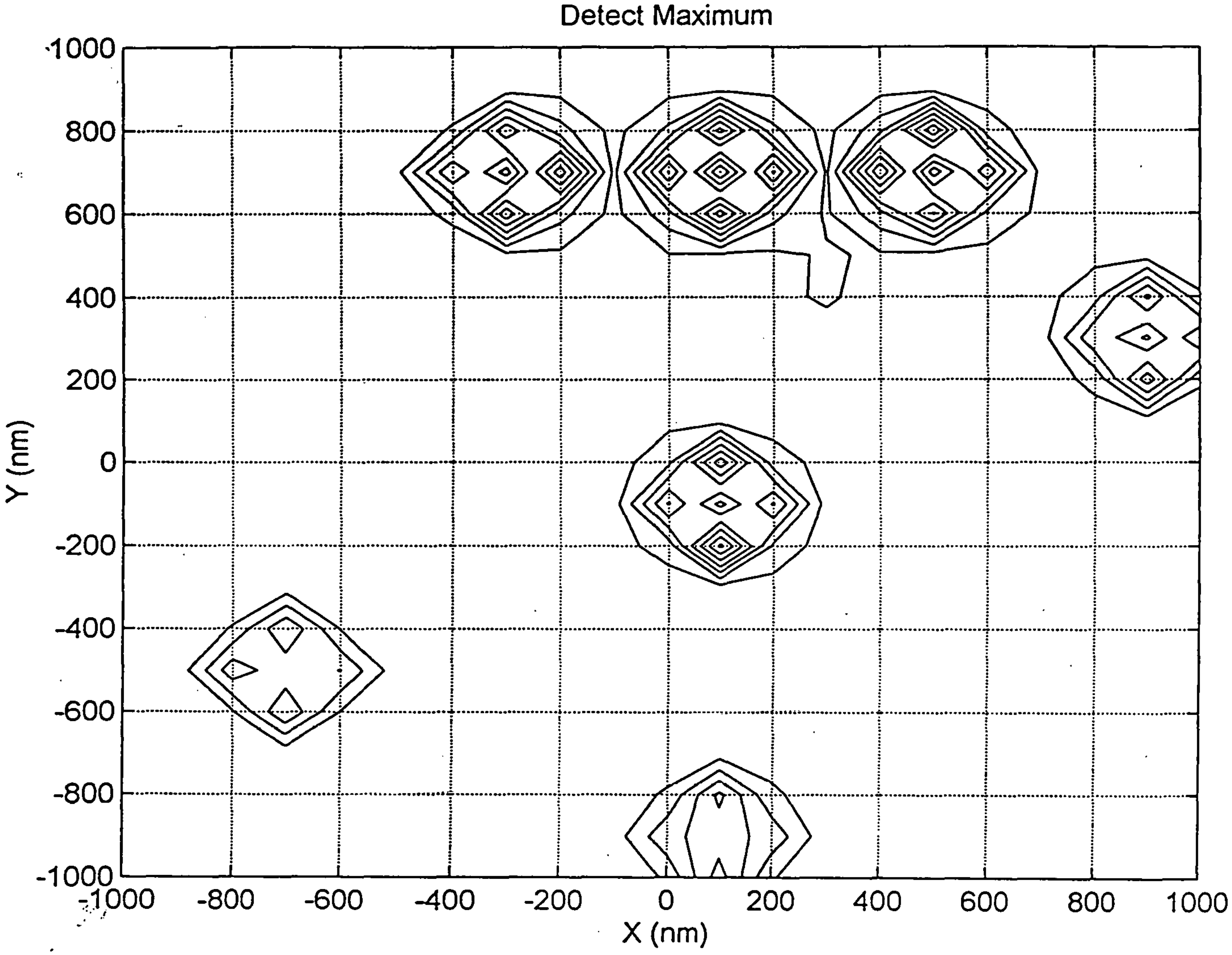


Figure 5: Contour plot of the detected intensity at the in-plane detectors. See text for details.

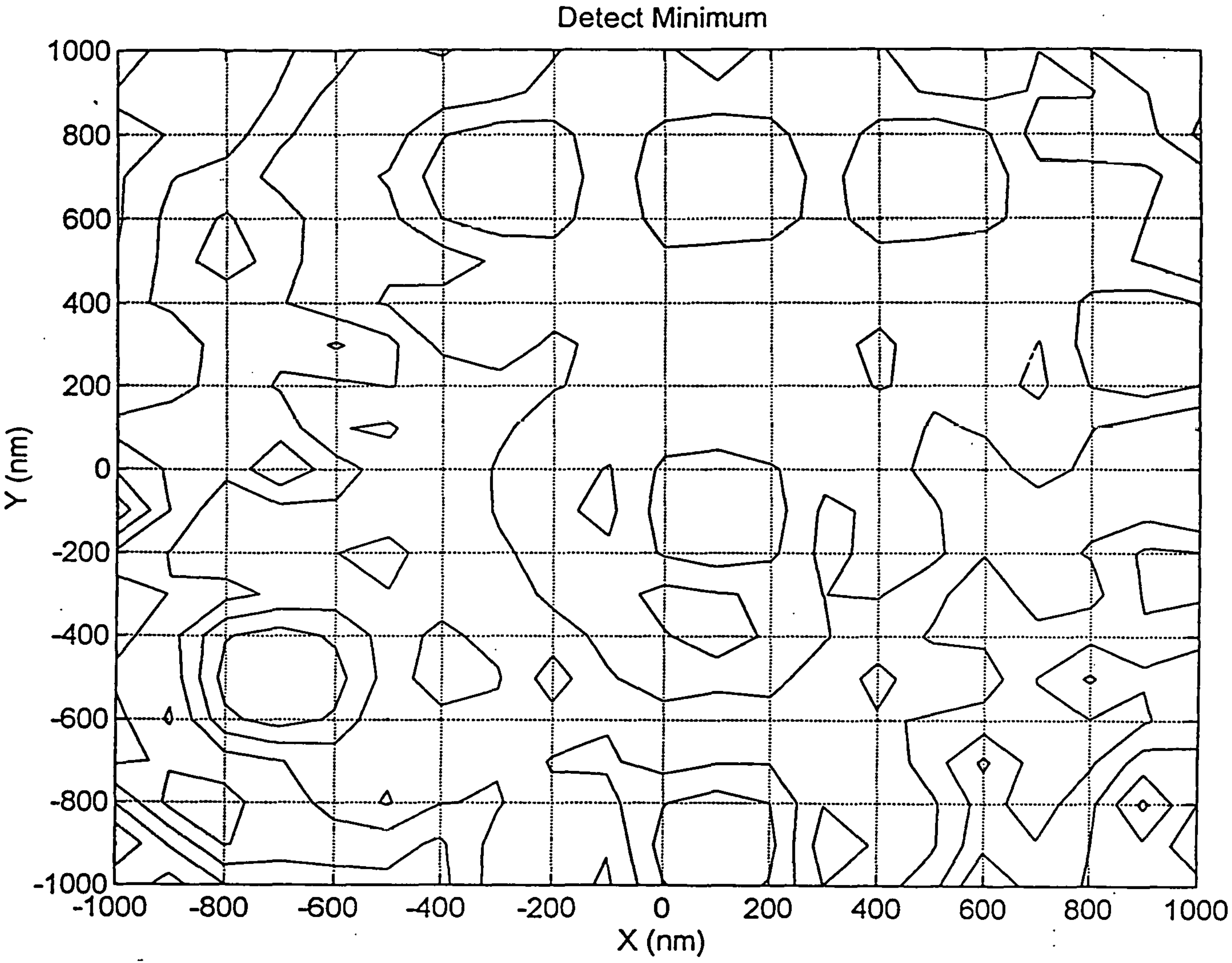


Figure 6: Contour plot of frequency shift in terms of fraction of resonant frequency.

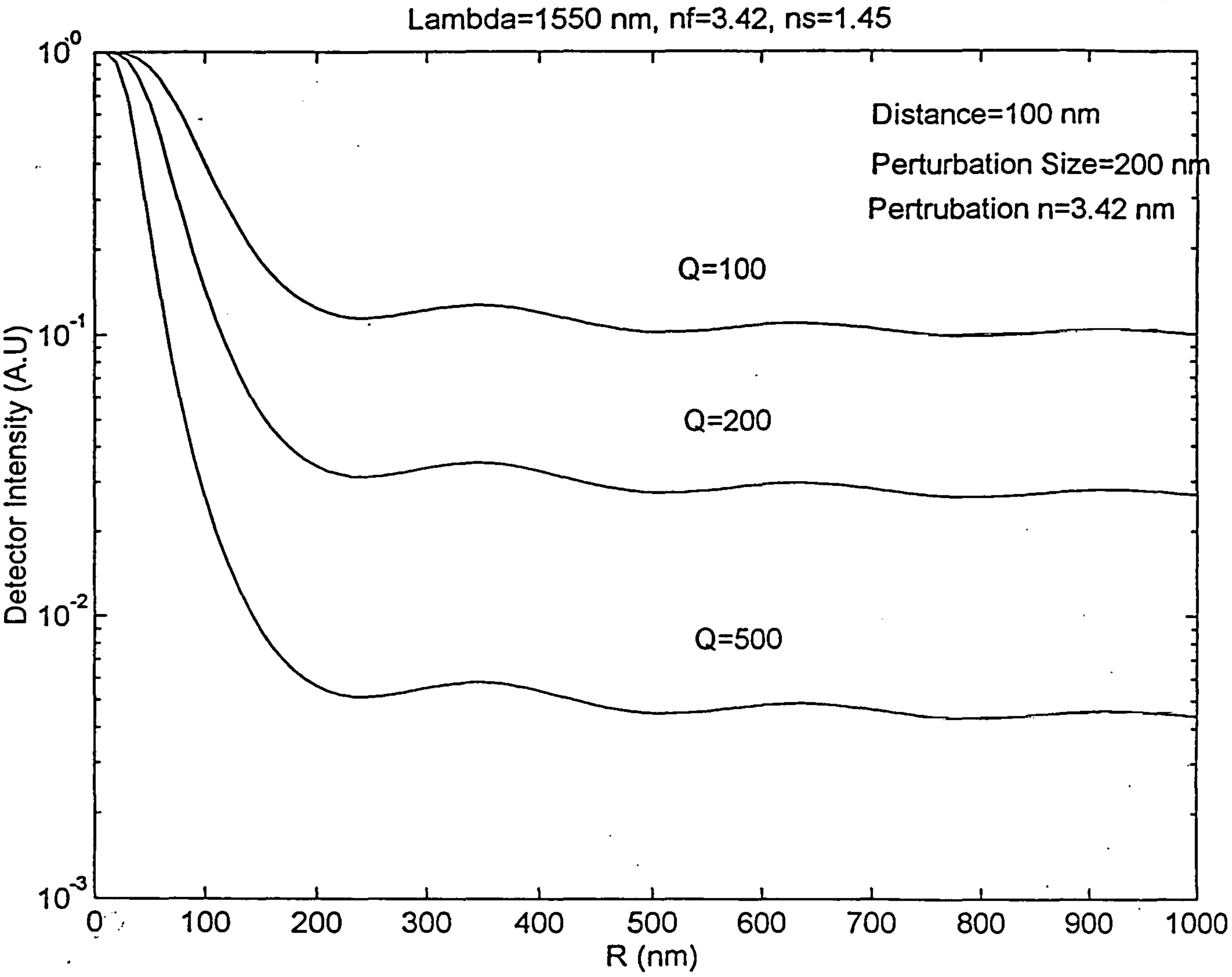


Figure 7: Intensity at in-plane detector as a function of pit position. R=0 implies pit and central zone of detector are concentric. See text for details

Commercial Potential

We anticipate our near-field probe to have significant commercial potential. As was mentioned earlier, the the semiconductor and optical data storage industry would benefit significantly from a a near-field optical instrument that is rugged, with high resolution and which has data collection rates.

We anticipate the following applications of the probe:

1. *High Desntiy Optical Read Head:* The CDBR near-field probe is well suited for use in an optical data storage instrument.
2. *Critical Dimension Measurement Probe:* With the sub-200 nm resolution of the probe, the probe is well suited for critical dimension measurmnts in the semiconductor industry.

Although the design of this product was not motivated by direct solicitation, we believe that there is a nascent market for this product and therefore, a commercial interest for the probe described here is expected. A number of companies are actively involved in high resolution microscopy. Presently, we are aware of one company manufacturing optical near-field instruments. Topometrix Inc. manufactures a near-field optical microscope based on the designs and research at AT&T Bell Laboratories. Their instrument, which roughly sells for $150,000 is capable of 50 nm resolution. However, it faces the limitations described earlier (low data rates, poor reliability, etc.) . This instrument has been purchased by a number of semiconductor companies (e.g. Motorala, Intel) high resolution optical diagnostics for their sub-micron development.

The application of near-field techniques to optical data storage is also a topic of commercial interest. Most notably, IBM has been looking at solid immersion lens as a possible avenue for high-density data storage. Solid immersion lens are described in one of the papers that is attached to this technical advance. It is clear near-field techniques will play a significant role in the future developments in optical data storage.

The CDBR probe described here has a number of advantages over competitive ideas in terms of manufacturing. It can be fabricated via lithography (unlike other techniques such as solid immersion lens and present near-field probes which require machining). Consequently, it can be produced in bulk quantities. This is critical if these probes are to be used in optical data storage applications.

Distribution:

1	MS	0603	G. R. Hadley, 1312
1		0603	I. J. Fritz, 1312
1		0603	R. Corless, 1312-1
1		0603	A. J. Howard, 1322
1		0603	G.A. Vawter, 1322
1		0603	J. R. Wendt, 1322
1		0343	J. A. Borders, 1823
20		1411	B. R. Stallard, 1823
1		1411	S. Kaushik, 1823
1		9018	Central Technical Files, 8523-2
5		0899	Technical Library, 4414
1		0619	Print Media, 12615
2		0100	Document Processing, 7613-2 For DOE/OSTI

Org.	Bldg.	Name	Rec'd by	Org.	Bldg.	Name	Rec'd by

Sandia National Laboratories