# Holographic Information Systems

**Thesis by**
**George Panotopoulos**

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy


California Institute of Technology

Pasadena, California

2003

(Defended July, 2002)

*To Panos and Evie, my parents.*

# Acknowledgments

I would like to start, as the unwritten rules of acknowledgment writing dictate, by thanking Professor Demetri Psaltis, who held the position of my academic advisor during my five years at Caltech. The opportunity he gave me to work among his group of extremely gifted students, as well as the oral and written presentation skills I acquired through my interaction with him, are greatly appreciated. I would also like to thank Prof. Dr. Karsten Buse. I first met Karsten while he was a postdoc in the Psaltis group, and I was a first year graduate student. Since then we had many opportunities to collaborate, culminating to my joining his group at the University of Bonn for two months, which was among the most enjoyable and productive periods I had while working on my thesis. During this time, I grew to appreciate the strict quality standards he applies to his scientific work, and I was inspired by his unwavering devotion to the well being of his students. While at Caltech I had the opportunity to interact with several other faculty members. Professors Perona, Abu-Mostafa, Tanguay, and Koch were kind enough to share their knowledge and insight, helping me gain better understanding of their respective fields, and contributing significantly to my scientific growth. For that I thank them. Last, but not least, I would like to express my gratitude to Professor Hercules Avramopoulos, who supervised my undergraduate thesis at NTUA, and first introduced me to experimental optics. He has since been a good friend and mentor, whose good advice I can always rely on.

The best aspect of graduate study at Caltech is the opportunity to interact with a number of exceptional students. I would like to thank George Barbastathis for sharing his experience in the NTUA to Caltech transition. Xin An, for helping me make my first steps in a free-space optics lab. Michael Levene, for dispensing valuable advice I was too naive to appreciate at the time, as well as Ernest Chuang and Greg Billock. Ali Adibi for helping me revive Greco-Persian animosity after 2500 years, Greg Steckman for restraining my socialist deviations, Shaw Wang, Wenhai Liu, and my longtime office mate Chris Moser, for teaching me how fast people can

change. Zhiwen Liu for his friendship, and for dispensing his extraordinary intuition so unassumingly. For my contemporaries, many thanks to Jose Mumbru, for his friendship and dry humor that helped us both through the OPGA project, and Yunping Yang, for looking after my figure. From the next generations of students, I would like to thank Hung-Te Hsieh for our excellent collaboration on the strong gratings project, for his careful review of the relevant chapters of this thesis, and of course for his remarkable cooking skills. Martin Centurion for his friendship and indulgence of my use of Spanish, Zhenyu Li, and Hua Long. I would also like to thank for their friendship a number of people who were members of the Psaltis group at various points in time: Benjamin Schupp, Irena Maravic, Todd Meyrath, Dirk Walther, Ingo Nee, Vijay Gupta, George Maltezos and Manos Fitrakis. Regarding the other members of the Psaltis group, many thanks to Yayun Liu for being the most fervent fan of my Hollywood career, and for supporting my delusions of grandeur, and to Lucinda Acosta, for taking care of administrative matters, but more importantly, for the motherly good care she takes of students. Special thanks to all the members of the Buse group in Bonn, for their help and friendship during my stay, and especially to Marc Luennemann, for being a model host and a very dear friend.

On a more personal level I would like to thank all my friends for their love and support, with a special mention to my NTUA classmates, who help me feel as if I never left every time I go back home. Special thanks to George and Polly for recreating a family feeling and co-hosting a number of parties in the Garibaldi street house, and to my other housemates, Faidonas and Alexandra. I would also like to thank Carla for her love, and for being there for me, during this very dark and frustrating period of my life.

Finally, I would like to express my gratitude to my family, and especially my parents, Panos and Evie, for their continuous love and support, and for providing me with an education far more important than the one I received through formal training. It is their devotion and hard work that laid the strong foundations upon which I have been able to build.

# Abstract

The goal of this work is to investigate the use of holographic techniques for information processing and transmission systems. Until recently information has been processed and transmitted mainly electronically. With the advent of optical fiber communications the monopoly of electronics has receded in the telecommunications field, but the domain of information processing is still dominated by electronic processors.

This thesis follows a top-down approach to the design of processors that integrate both electronic and optical components. It begins with the design considerations of a compact, rapidly reconfigurable opto-electronic processor, which possesses an optical bus in addition to the traditional electronic bus. The optical bus takes advantage of the massive parallelism that is afforded by optics and can be coupled to a holographic digital memory, allowing rapid reconfiguration of the device. The capability of rapid reconfiguration gives rise to a new computational paradigm, where the reprogramming of the device can become part of the computation. We suggest additional applications of this processor, namely as a smart reading head for large scale holographic disk memories. Finally we present novel algorithms that were developed specifically to take advantage of the additional capabilities of our processor.

The next section is concerned with the wavelength and angular tuning of strong volume holograms, both in the reflection and 90-degree geometries. Since photons have no charge, we need to rely on their wave properties to manipulate them, both for long-range transmission, such as telecommunications, and short-range transmission, such as on chip interconnects. In this section we investigate how volume holograms can be used to selectively redirect information bearing light beams.

The final part of this thesis is concerned with material issues. Holographic recording of strong volume gratings is one of the most commonly used approaches, and photorefractive materials have a strong bearing on the overall performance of the final system. Two properties of iron doped lithium niobate are investigated, namely

the dependence of absorption on temperature and the quadratic electro-optic coefficient. The former is crucial for the commonly used technique of thermal fixing, and the latter can become significant should we choose to use applied continuous fields to tune our gratings.

# Contents

# List of Figures

# List of Tables

# 1 Introduction

## 1.1 Holography for information systems

In this thesis we investigate the use of holography in the context of information processing. Information can be encoded on light beams, by modifying their temporal or spatial properties. Holograms can be used to manipulate the properties of those light beams, effectively processing the information they bear. In this section we will discuss, in the most general terms, what kind of processing can be implemented using holographic elements.

Plane waves are the eigenmodes of free space [1-1]. Sinusoidal gratings can be used to couple two such elementary modes [1-2]. It is also possible to superimpose multiple gratings in the same volume, thus enabling more complicated wavefronts to be coupled to the input wavefront. In most holographic memory systems, a

single reference plane wave is coupled to multiple gratings in order to reconstruct the highly structured information page [1-3]. These memories can be interfaced to optoelectronic circuits, thus enhancing key aspects of their performance [1-4,1-5].

Due to their large number of periods, holographic gratings exhibit significant selectivity to angular and wavelength detuning. It is, therefore, possible for strong volume holograms to act as filters that isolate and redirect specific spatial and spectral information. Such gratings can be used to implement tunable interconnects for a chip, act as wavelength filters [1-6,1-7] or demultiplexers [1-8] for Wavelength Division Multiplexing (WDM) fiber communications systems, or be used to perform spatially and spectrally selective imaging [1-9,1-10].

In this thesis we will focus on the investigation of two of these possible uses. The first is the use of holograms as memory elements, and their combination with optoelectronic circuits. The second is the use of strong holographic volume gratings as filtering elements for WDM optical communications. In the next sections of this chapter we present an outline of the thesis.

## 1.2 Optically programmable gate arrays

We will begin in Chapter 2 by detailing the design, implementation, and testing of a compact, rapidly reconfigurable, optoelectronic processor. We will go through initial design considerations and briefly mention all the constituent components, spending more time on the testing and integration of the silicon circuit that combines photodetectors with logic circuits. We will then detail experiments to implement and test the integration of these devices, culminating in the demonstration of holographic reconfiguration of our processor.

In Chapter 3, we will generalize our architecture, and show how a real-time video processing algorithm can be mapped to it. Then, we will present two algorithms specifically designed to take advantage of this device. The first one is a Neural Network based approach to the classification of cursive digits, that presents a new computational paradigm by introducing the reconfiguration of the device as part of the computation. The second one is an algorithm that allows OPGA's to be used

as smart head that are able to retrieve classes of objects from large, non indexed holographic databases.

## 1.3 Strong volume gratings

As mentioned in Section 1.1, strong volume gratings can be used to implement tunable interconnects as well as filters of WDM fiber communication systems. In Chapter 4 we will investigate the behavior of strong gratings, in the reflection geometry. We will start with a general treatment of waves in two-dimensional periodic media. Then we will focus on the reflection geometry, and investigate some of the finer aspects of the coupled mode theory, with applications to polarization dependence, angular, wavelength, and thermal detuning. Finishing with the reflection geometry we present experimental results confirming our theories.

We will then proceed to the 90-degree geometry, in Chapter 5. We will extend the coupled mode formalism to two dimensions, and present theoretical and experimental results regarding diffraction efficiency. Next, we will discuss wavelength detuning, and develop a numerical solution for the coupled-mode equations in that case. We will conclude the chapter with a numerical calculation of the impulse response of strong volume gratings in the 90-degree geometry.

## 1.4 Materials

In Chapter 4 and Chapter 5, we discuss the use of strong volume gratings as filters. In this context the lifetime of the recorded gratings becomes an issue. One of the most promising methods to increase the lifetime of holographically recorded gratings is thermal fixing, which involves recording at elevated temperatures. In Chapter 6 we will discuss the effects of temperature on the absorption spectrum of iron doped lithium niobate ($LiNbO_3$). First we will give an outline of the absorption mechanisms. Then we will present the experimental methods we used and the results obtained. Finally we will propose a theoretical model that explains our experimental observations, and discuss its implications in terms of possible applications and extensions to different dopants.

Our ability to record gratings in photorefractive materials is due to the electro-optic effect, which allows electric fields induced by charge redistribution to be manifested as a modulation of the refractive index. Though the presence of grating is due, for all practical purposes, to the linear electro-optic effect, it is of interest to investigate the magnitude of the quadratic electro-optic effect. In Chapter 7 we will begin by reviewing an experimental setup used to measure the linear electro-optic coefficients. We will then proceed to describe the limitations that prevent the use of such a setup for the measurement of the quadratic electro-optic coefficients, highlighting the challenges presented. Finally, we will present the design of a measurement system capable of measuring the quadratic electro-optic coefficients, even in the presence of much stronger linear electro-optic coefficients, as well as results for the case of manganese-doped lithium niobate ($LiNbO_3$:Mn).

## 1.5 References

[1-1] A. Yariv and P. Yeh, *Optical Waves in Crystals*. 1984, New York: John Wiley & Sons.

[1-2] H. Kogelnik, *Coupled Wave Theory for Thick Hologram Gratings.* Bell System Technical Journal, 1969. **48**(9): pp. 2909-2947.

[1-3] H. J. Coufal, G. Sincerbox, and D. Psaltis, eds. *Topics in Applied Physics: Holographic Memories*. 1999, Springer: Berlin, New York, Tokyo.

[1-4] J. Mumbru, G. Zhou, S. Ay, X. An, G. Panotopoulos, F. Mok, and D. Psaltis. *Optically reconfigurable processors*. in *Euro-American Workshop on Optoelectronic Information Processing*. 1999: SPIE.

[1-5] J. Mumbru, G. Panotopoulos, D. Psaltis, X. An, F. Mok, S. Ay, S. Barna, and E. Fossum. *Optically programmable gate array.* in *Proc. SPIE of Optics in Computing*. 2000.

[1-6] V. Leyva, G. A. Rakuljic, and B. O'Conner, *Narrow bandwidth volume holographic optical filter operating at the Kr transition at 1547.82 nm.* Applied Physics Letters, 1994. **65**(9): pp. 1079-1081.

[1-7] J. Hukriede, I. Nee, D. Kip, and E. Kratzig, *Thermally fixed reflection gratings for infrared light in $LiNbO_3$:Ti:Fe channel waveguides.* Optics Letters, 1998. **23**(17): pp. 1405-1407.

[1-8] S. Breer, H. Vogt, I. Nee, and K. Buse, *Low-crosstalk WDM by Bragg diffraction from thermally fixed reflection holograms in lithium niobate.* Electronics Letters, 1998. **34**(25): pp. 2419-2421.

[1-9] G. Barbastathis, M. Balberg., and D.J. Brady, *Confocal microscopy with a volume holographic filter.* Optics Letters, 1999. **24**(12): pp. 811-813.

[1-10] W. H. Liu, D. Psaltis, and G. Barbastathis, *Real-time spectral imaging in three spatial dimensions.* Optics Letters, 2002. **27**(10): pp. 854-856.

# 2 Optically programmable gate array

## 2.1 Introduction

Information processing and transmission has been dominated, until recently, by the use of electronic encoding. This is due to the fact that electrons, thanks to their charge, are fairly easy to manipulate, and since they constitute part of matter, it is straightforward to achieve interaction between a device and an electronic stream of information. In the last few years, optical and optoelectronic systems have been widely adopted in the realm of telecommunications, a trend mainly due to their higher bandwidth. Still, there is no similar penetration of use of optical techniques in local information processing. This is due to photons having no charge, thus forcing us to manipulate them by taking advantage of their wave properties. Neverthe-

less, it is considerably harder to force light to interact with matter than it is to force electrons to do so.

On the other hand, again due to the lack of charge, photons have the very advantageous property of not interacting with each other. This inherently reduces the crosstalk of optical communications, and allows unprecedented degrees of parallelism in on-chip interconnects, hardly conceivable in the case of purely electronic circuits. Furthermore, the decrease of device size in VLSI circuits is not followed by a similar decrease in interconnect dimensions. As devices shrink and operating frequencies increase the need for more efficient, less real-estate-consuming interconnection solutions will become pressing. We believe that optical interconnects will provide an elegant solution to this problem [2-1].

In this chapter, we propose the use of photonics to implement data transfer, on a chip scale. More specifically, we discuss the design, fabrication, and characterization of a compact, rapidly reconfigurable optoelectronic processor [2-3, 2-4], which relies on the interfacing of a silicon optoelectronic circuit with a holographic optical memory [2-5, 2-6]. The device we propose is an extension of a Field Programmable Gate Array (FPGA) [2-7], a widely used class of reconfigurable digital circuits. The main idea behind FPGA's is that any Boolean function can be implemented by proper combination of elementary Boolean functions (e.g., all Boolean functions can be implemented using NAND gates). FPGA circuits contain two components:

- Reconfigurable look-up tables (LUT's).
- Reconfigurable interconnects.

The LUT's implement elementary Boolean functions. By configuring the look-up tables and combining their outputs using interconnects, FPGA's can implement any logic function. The versatility, ease of programming, and computing power of FPGA's has propelled them from limited use in development and prototyping, which was their initial target application, to widespread use in many commercial electronic products. In addition, FPGA's, occupying the middle ground between the processing power of Application Specific Integrated Circuits (ASIC's) and the ease

of programming of microprocessors, are widely used in demanding Digital Signal Processing (DSP) applications [2-8, 2-9].

The enhancement of FPGA's with optical interfaces has been discussed in literature, mainly in terms of high-bandwidth I/O [2-10] and interchip interconnects [2-11]. The device we propose, namely the Optically Programmable Gate Array (OGPA), addresses one of the main limitations of FPGA's: Their long reconfiguration time. FPGA's have only one set of inputs/outputs (I/O), they communicate electronically through the pins of the chip. Since most of these I/O resources are used for the transfer of data that is processed by the device, very few I/O ports are available for its reconfiguration. As a result the reconfiguration of an FPGA can last several milliseconds, and in practice FPGA's are programmed before the data processing begins and do not alter their functionality during operation.

By contrast, the OPGA has an additional set of inputs, namely a set of photodetectors on the surface of the chip. These inputs are linked to a holographic memory that contains the reconfiguration data necessary to program the logic functionality of the device. The inherent 2-D aspect of holographic memories, combined with fast access, allows OPGA's to be reprogrammed in a matter of microseconds, thus providing a new computational paradigm, where the reconfiguration of the device can be part of the computation.

In this chapter we discuss the design and implementation of the OGPA device. We present the elements of the device and some of the trade-offs involved. Then we focus on the chip that combines the reconfigurable logic and detectors, discuss the methods used to characterize it and the corresponding results. Finally, we introduce a demo setup that showcases the successful reconfiguration of an electronic circuit, using a holographic memory.

## 2.2 Elements of the OPGA

The OPGA is composed of the following basic elements (see Figure 2-1):

- A silicon integrated circuit combining reconfigurable logic and photodetectors.

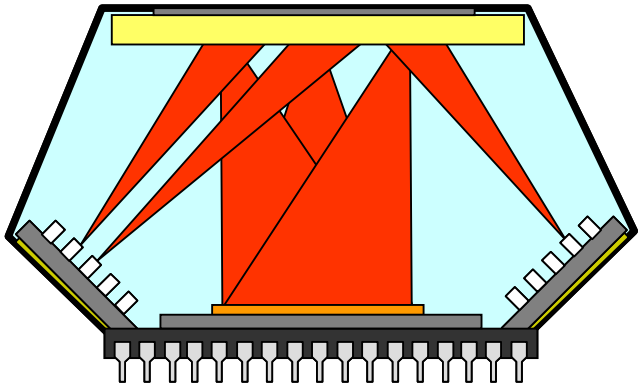


**Figure 2-1.** Compact OPGA module. The addressing devices access different pages stored in the holographic memory element. The information contained in these pages is imaged on the surface of the chip. There photodetectors collect the information, which is then used to reconfigure the programmable logic that is also implemented on the chip.

- A holographic memory element, where the reconfiguration templates are stored.
- An addressing device, used to access information in the holographic memory, and transfer it to the silicon circuit.

The processing element of the OGPA is located on the silicon circuit, and is implemented by configuring properly the logic blocks and interconnects. The information required to configure the processing element is stored in the holographic memory element. When we want to change the functionality of the processing element we use an addressing device to select a specific reconfiguration data page from the memory, and transfer it on the surface of the chip. There, the photodetectors receive this information, which is then routed to the reconfigurable elements. The information is used to change the parameters of these elements, leading to the implementation of a new functionality.

The holographic memory element can be made using several different materials, such as polymers (Aprilis [2-12], DuPont [2-13]), or photorefractive crystals ($LiNbO_3$:Fe), and the information can be recorded using a number of different multiplexing techniques, such as reflection geometry [2-14], transmission geometry, $90^o$ geometry [2-15], shift multiplexing [2-16], and peristrophic multiplexing [2-17]. In selecting the holographic material and multiplexing technique, we need to address the competing needs for memory capacity and efficient reconstruction of the stored information [2-18].

The addressing devices are crucial in determining the minimum reconfiguration time, in the sense that they should enable us to access rapidly different memory pages, and have enough power, so that the information that is reconstructed on the surface of the chip can be integrated by the photodetectors in a short time. Devices that fulfill these requirements are arrays of Vertical Cavity Surface Emitting Lasers (VCSEL's) [2-19], and Micro-electro-mechanical (MEM's) arrays combined with laser diodes.

## 2.3 The OGPA chip

The OPGA chip consists of two main elements: a photodetector array and reconfigurable logic. Since the reconfiguration data is to be transferred from the memory to the photodetectors, and then to be used to reconfigure the LUT's and interconnects, it would make sense to place the reconfigurable logic elements and the photodetectors used to reconfigure them in close proximity. Such an optically

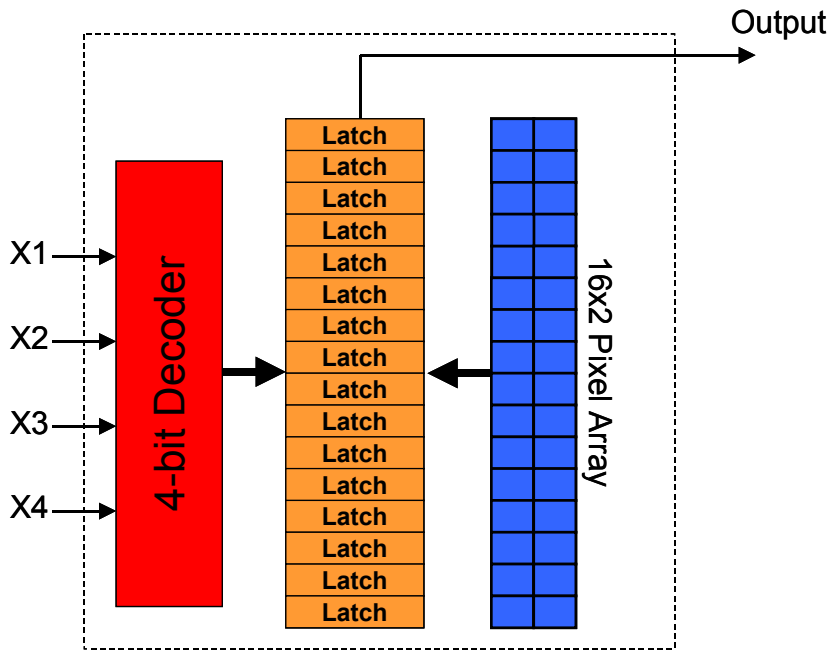



**Figure 2-2.** Optically Reconfigurable Look-Up Table. The information received by the photodetector array is stored in the latches. The input defines through a decoder which latch value will be output, thus implementing the LUT.

reconfigurable LUT is shown in Figure 2-2. Optical information collected by the Active Pixel Sensor (APS) array is transferred to an array of 16 latches. A four-bit

decoder receives an input and presents the contents of the corresponding latch at the output. By selecting the values stored in the latches we can implement any 4-bit input, 1-bit output Boolean function.

Combining the photodetectors and reconfigurable logic elements has obvious advantages, in terms of minimizing the wiring required for configuration-data transfer. On the other hand, such a design results in bigger pixels with smaller fill-factor, putting a strain on the power requirements of the laser source, and the diffraction efficiency requirements of the holographic memory.

We designed and fabricated a prototype OPGA chip in collaboration with Photobit Inc. The chip was fabricated using a 0.35 micron, CMOS technology. It is shown in Figure 2-3, where the two main elements, the photodetector array and the reconfigurable logic, are easily distinguishable. In this prototype chip we chose to imple-

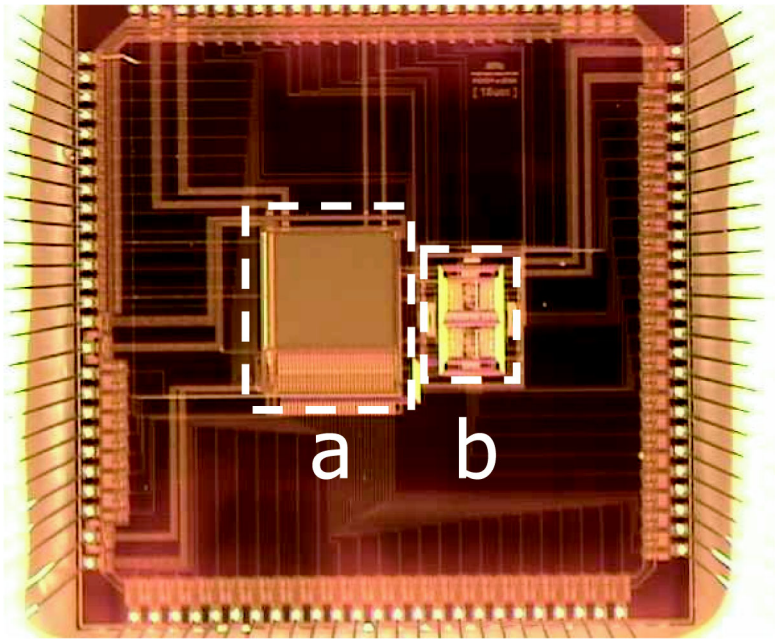



**Figure 2-3.** Photograph of the OPGA chip. (a) 64 × 64 Active Pixel Sensor Array. (b) Reconfigurable logic.

ment the photodetector array and the logic at different locations, mainly in order to facilitate testing. The trade-offs between the two possible designs should be con-

sidered in future implementations. In the following sections we describe in detail these two elements.

### 2.3.1 The APS array

The photodetector array was implemented as a 64 × 64 active pixel sensor differential array [2-20, 2-21]. Each pixel of this array is an active pixel, the difference

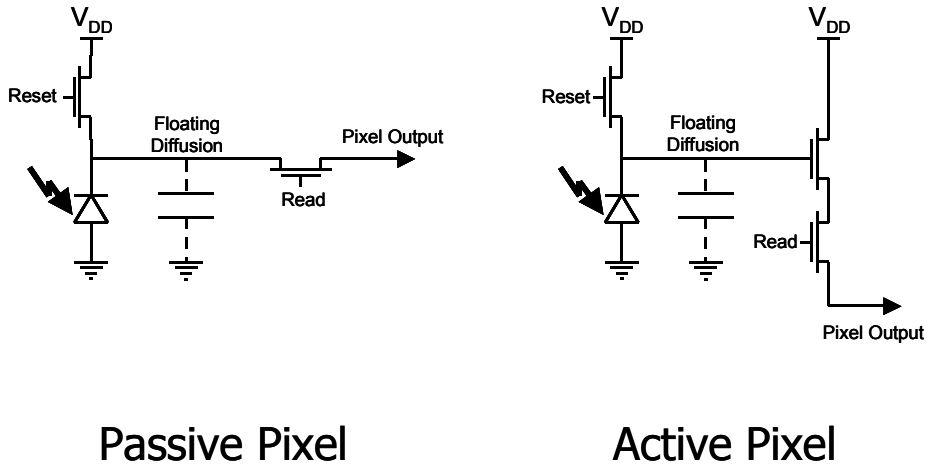



**Figure 2-4.** Schematic diagrams of Passive and Active Pixels. For the Passive Pixel the charge of the floating diffusion is output directly. For the Active Pixel it is amplified in-situ before being output.

between a passive pixel and an active pixel being that in an active pixel the charge generated by the photodiode is actively amplified before being red-out (see Figure 2-4). As a result, active pixels achieve significantly better signal-to-noise (SNR) ratios than their passive counterparts.

The array also employs differential encoding, so that every pair of pixels encodes only one bit, i.e., a pixel combination of dark-bright corresponds to 1 and a combination of bright-dark corresponds to 0. This encoding is a rudimentary error correcting code, used to counter the fact that the intensity across large reconstructed holographic data pages is not constant. The pixel combinations dark-dark and bright-bright are not permitted, and result to a random flickering value and a zero value respectively. Examples illustrating the concept of differential encoding are shown in Figure 2-5. The differential encoding is implemented by driving the output voltages of the two pixels of a pixel pair through a comparator, as shown in

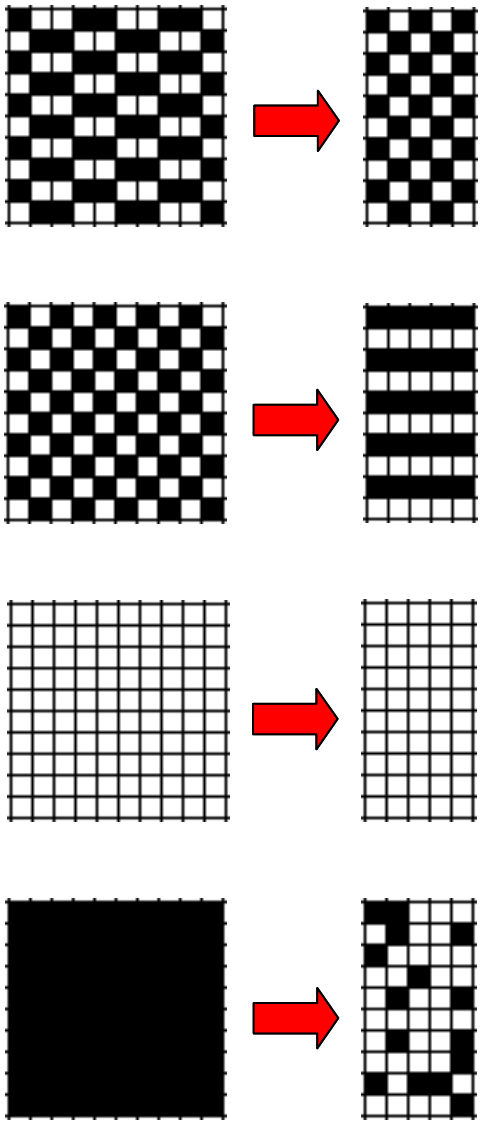


**Figure 2-5.** Examples of differential encoding. The left column corresponds to the data projected on the APS array, with white squares corresponding to illuminated pixels and black ones to dark pixels. The right column corresponds to the data that is read out, with black corresponding to 1 and white corresponding to 0. The last two examples illustrate the response of the array to nondifferential inputs.

Figure 2-6. The nonpermitted states result from the fact that the corresponding inputs drive the comparator out of its normal operation region. The different behavior for different nonpermitted states stems from the asymmetric design of the comparator.

The APS array can be read row-wise, and rows can be accessed in random order. When a row is being read out, the corresponding pixel values are compared pair-

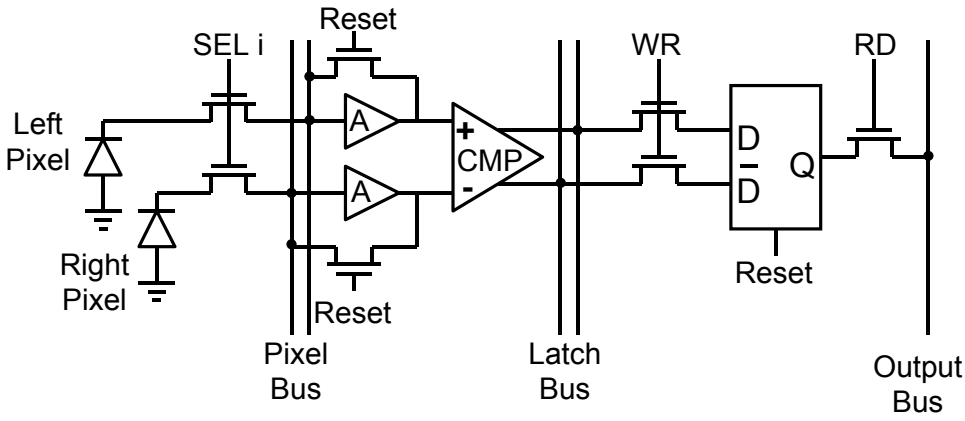



**Figure 2-6.** Schematic diagram of the differential APS pair. The two pixels are amplified independently, and then driven through a comparator. The outcome of the comparison is stored in a latch, and accessed during readout.

wise, and the resulting outcomes are stored at the output latches of the array. Due to differential encoding, the 64 × 64 APS array produces 64 lines of 32 bits.

## 2.3.2 The reconfigurable logic

The reconfigurable logic, illustrated in Figure 2-7, consists of the following:

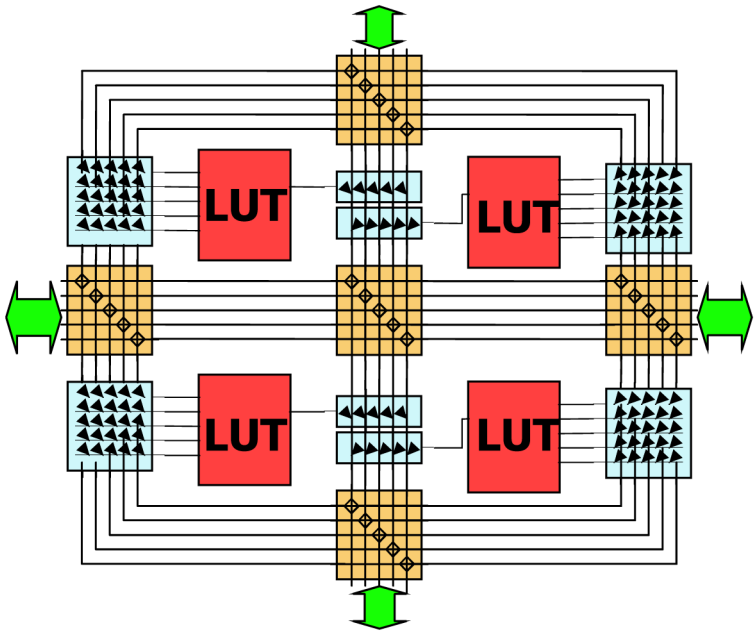



**Figure 2-7.** Schematic diagram of the reconfigurable logic. It is composed from four 5-bit input, 1-bit output LUT's, five interconnection matrixes and four I/O ports.

- Four 5-bit input, 1-bit output LUT's.
- Five interconnection matrixes.
- Four I/O ports.
- A five-line bus is used to transmit data to and from the different elements.

Each input of each LUT can be connected to any bus line, and the same holds for its output. Therefore the reconfiguration of a LUT block requires $2^5 = 32$ bits for the LUT itself, 5 × 5 bits for the reconfiguration of the input connections, and 5 bits for the reconfiguration of the output connections, bringing the total to 62 bits. As a result two lines of the APS array are necessary to reconfigure any LUT block.

Each of the five interconnection matrixes connects four segments of the 5-line bus. The corresponding lines of the four segments can be connected in 6 different ways, therefore 5×6 bits are required for the reconfiguration of an interconnection matrix, corresponding to a single line of the APS array.

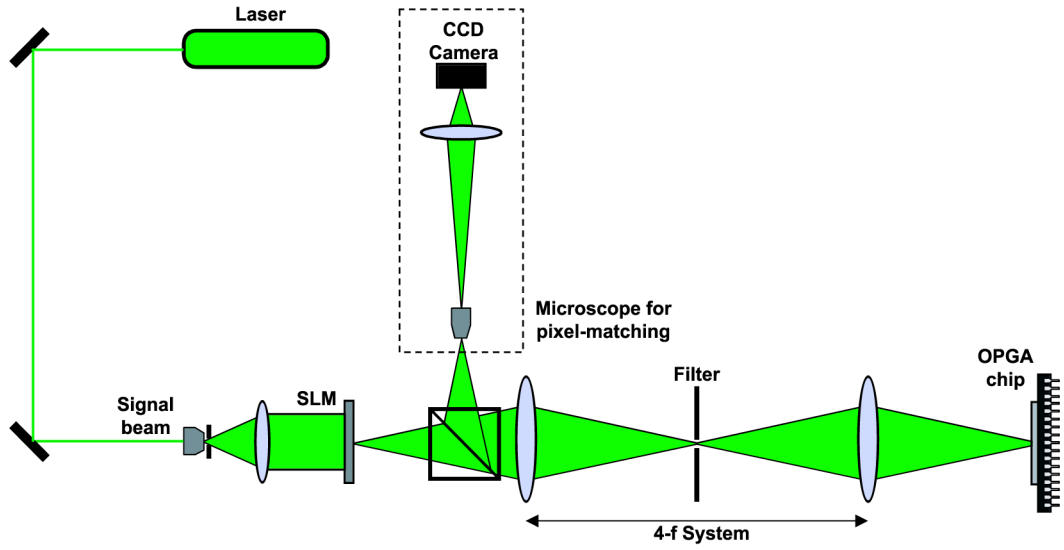
The I/O ports are implemented using tri-state buffers, so that they can be set to operate in input-only, output-only, or input-output mode.

The reconfiguration of the entire logic requires 4×2 + 5×1 = 13 lines of the APS array. Since each data page consists of 64 lines, it contains enough information for four complete reconfigurations of the logic.

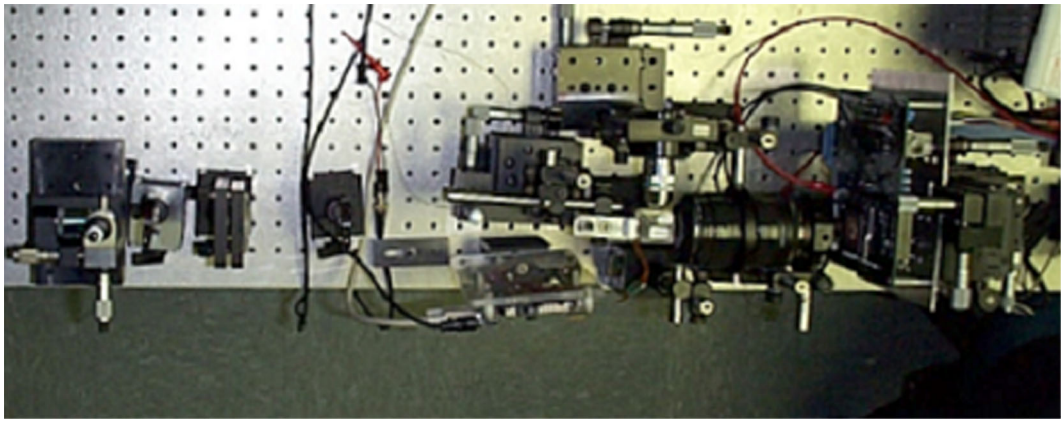
### 2.3.3 Performance characterization of the APS array

The first test we performed on the chip was to characterize the APS array. The chip is mounted on a board that generates the necessary bias voltages and timing signals. It also has an interface to the parallel port of a PC, through which we are able to collect and visualize the information displayed on the APS array. In order to test the APS array we need to pixel-match it to a Spatial Light Modulator (SLM), a liquid crystal device commonly used in overhead digital projectors. The pixel pitch of the APS array was designed to be 15 μm, equal to the pixel pitch of the Kopin 320 CyberDisplay SLM. There are several well-known techniques used to pixel-match SLM's to detectors. They rely on the display of known geometrical patterns on the SLM, and subsequent adjustment of the detector's position, based on the readout.

In the case of the APS array, the use of such a technique was rendered impossible due to the differential encoding used. Due to that, the output of the array flickers randomly when it is not pixel matched, and only stabilizes when we are very close to pixel match. In order to circumvent this hurdle we designed the optical setup shown in Figure 2-8.




(a)



(b)

**Figure 2-8.** Pixel matching experimental setup. (a) Schematic diagram. (b) Detail of the experimental setup.

In this setup the light of a HeNe laser is collimated and spatially modulated using the SLM. It is then imaged on the surface of the chip using a 4-f system, formed by

two low-aberration Nikon lenses. The chip and board are mounted on a combination of translation and rotation stages allowing five degrees of freedom (three translational and two rotational). The surface of the chip acts as a partial reflector, so the illuminating pattern is reflected back through the 4-f system. A beam-splitter, placed between the SLM and the first lens, splits the reflected light, part of which is imaged on a CCD camera through a microscope objective. A pair of LED's illuminates the surface of the chip independently, so that the image of the chip is superimposed to the image of the illuminating pattern on the CCD. Using the rotation and translation controls of the chip mount we are able to align the APS pixels and the illuminating pattern, thus achieving pixel-matching between the two, as shown in Figure 2-9.

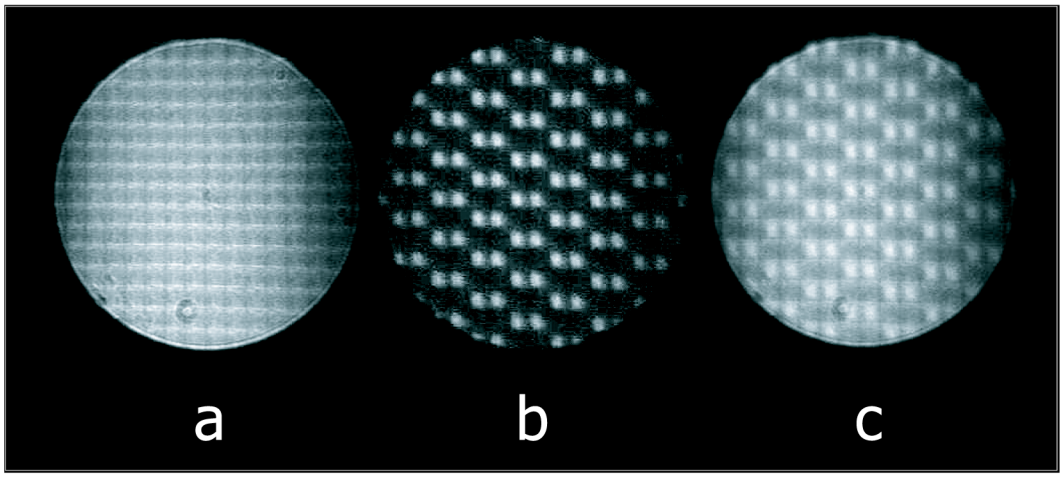
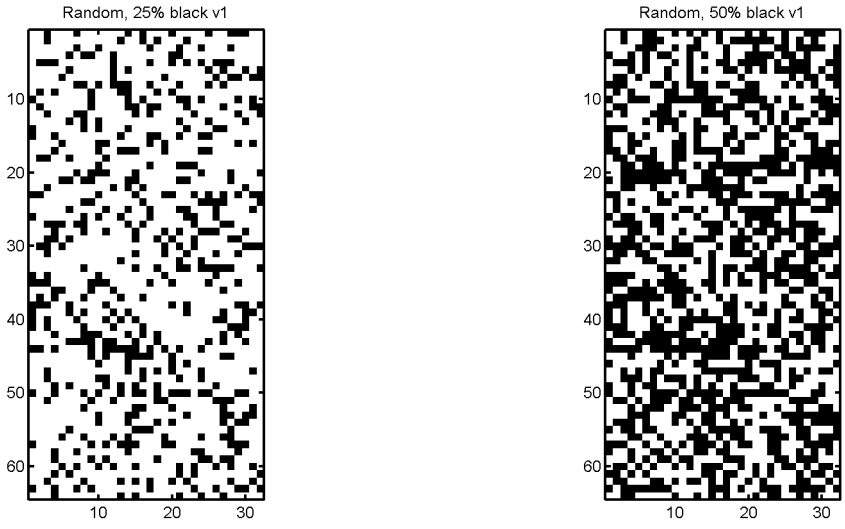


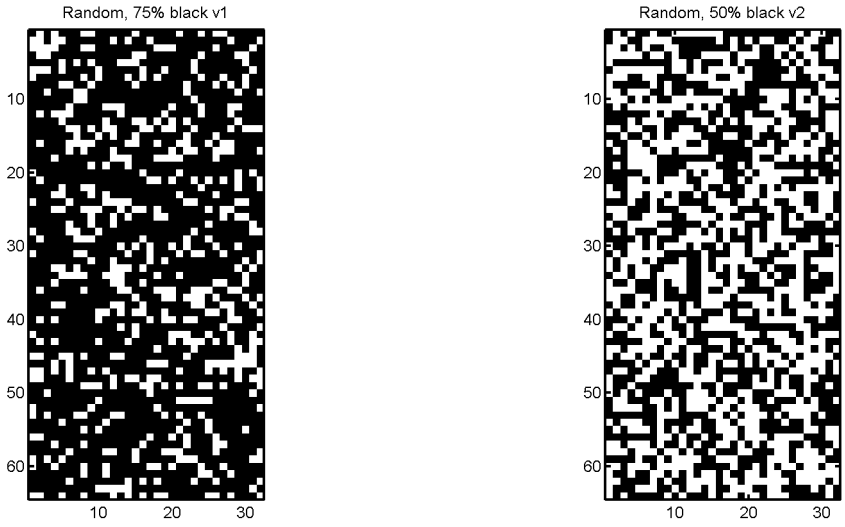

**Figure 2-9.** Images used for pixel matching. (a) Image of pixels. (b) Image of illumination pattern. (c) Superposition of pixels and illumination pattern.

The next step is to vary the intensity of known illuminating patterns using a variable attenuator, and measure the number of errors on the APS output. We used both random and structured patterns, shown in Figure 2-10. More information about the patterns can be found in Table 2-1. The probability of error $P_e$ is plotted vs. intensity (arbitrary units) in Figure 2-11, for three different random patterns. The three patterns are generated using the same distribution of random numbers by changing a threshold. Very low intensities lead essentially to the dark-dark combination

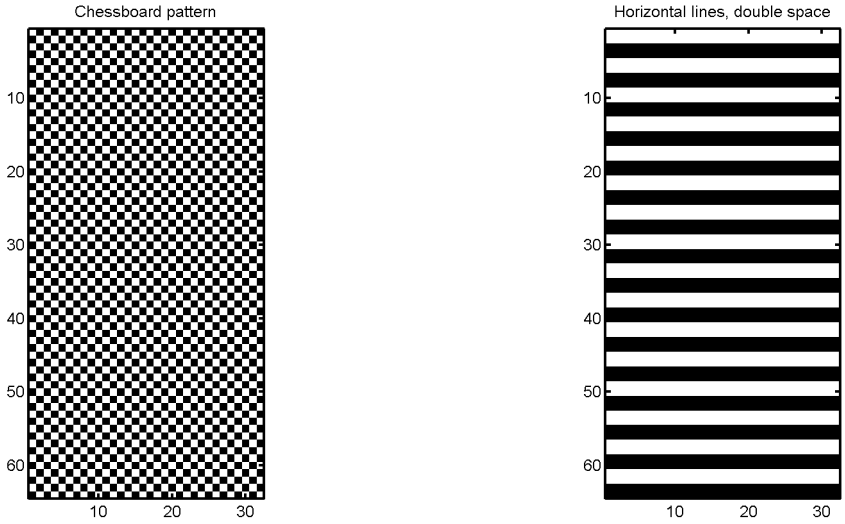



**(a)** **(b)**

**Figure 2-10.** Patterns used to test the APS array.




**(c)** **(d)**




**(e)** **(f)**

**Figure 2-10.** Patterns used to test the APS array.

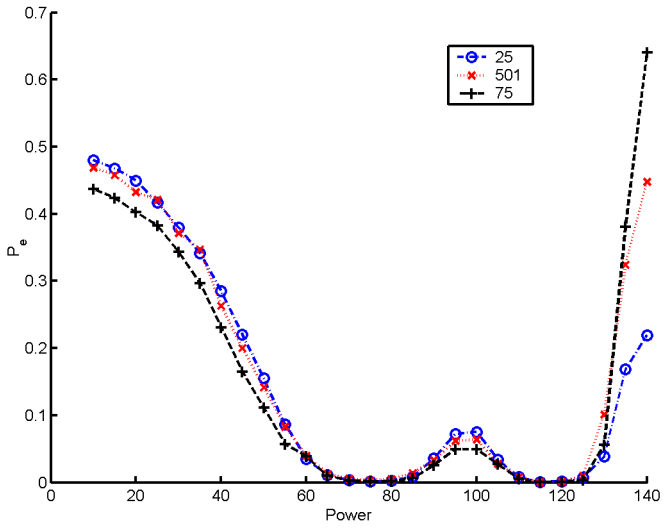



**Figure 2-11.** Probability of readout error $P_e$ vs. illuminating intensity (arbitrary units) for three random patterns.

being displayed on the APS array, and as a result we get quasi-random flickering, leading to a $P_e$ of about 50% for all patterns. At the other end of the spectrum,

| Pattern | Description | % Black pixels | Figure |
|---|---|---|---|
| 25 | Random pattern | 25 | Figure 2-10 (a) |
| 501 | Random pattern | 50 | Figure 2-10 (b) |
| 75 | Random pattern | 75 | Figure 2-10 (c) |
| 503 | Random pattern | 50 | Figure 2-10 (d) |
| chs | Chessboard pattern | 50 | Figure 2-10 (e) |
| hlds | Alternating 2-pixel witch black and white lines | 50 | Figure 2-10 (f) |

**Table 2-1.** Patterns used to test the APS array.

when the displayed pattern intensity is very high, we have the bright-bright pixel combination, which leads to an all-white output (white is equivalent to 0). As a result, the $P_e$ for each pattern is close to its percentage of black (or 1) pixels. In the middle of these two extremes we have a region where the $P_e$ drops and even reaches zero. This region is interrupted by a bump, whose height seems to be inde-

pendent of the percentage of dark pixels in the pattern. The actual intensity per pixel corresponding to the minimum BER (75 arbitrary units) is 0.37 pW.

In Figure 2-12, we compare the $P_e$ for two different random patterns, both having

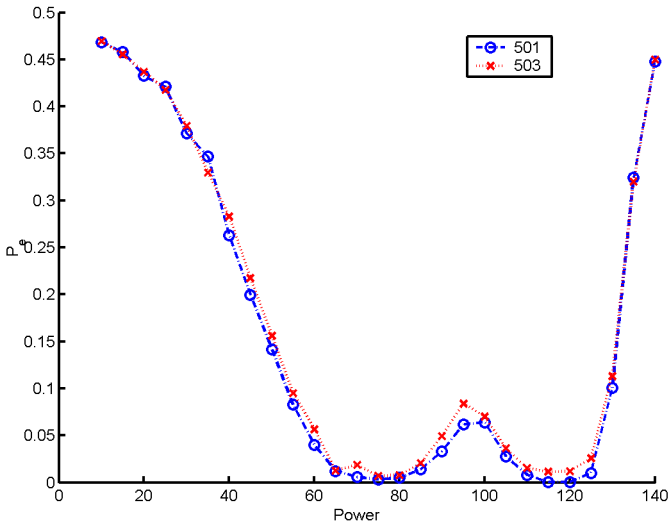



**Figure 2-12.** Probability of readout error $P_e$ vs. illuminating intensity (arbitrary units) for two random patterns, both having 50% black pixels.

50% black pixels. Their behavior is in good agreement. In order to gain some insight as to the origin of the bump we compare the $P_e$ of a random pattern with 50% black pixels, a chessboard pattern, and a pattern consisting of two-pixel thick horizontal lines. Note that all patterns have 50% black pixels. The results, shown in Figure 2-13, indicate that the frequency of change from black to white and vice-versa along the vertical direction is related to the $P_e$ in the middle intensity range. Indeed, the horizontal lines pattern has the lowest such frequency, followed by the random pattern, and then the chessboard pattern, which has the highest such frequency possible. This behavior is an indication that the errors in the middle intensity range are due to limitations of the latches, which are unable to switch states rapidly enough.

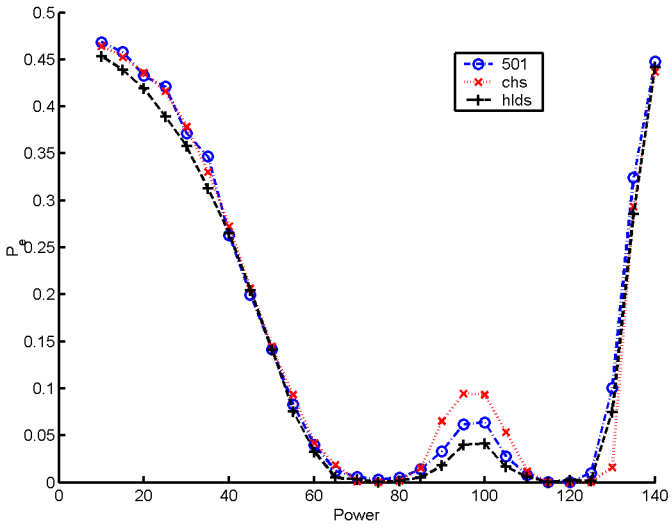



**Figure 2-13.** Probability of readout error $P_e$ vs. illuminating intensity (arbitrary units) for one random and two structured patterns, all three having 50% black pixels.

In conclusion, the latches are a target area for future improvements of the APS array. Nevertheless, the current design allows error-free readout of the information displayed on the APS in a specific intensity range. This enabled us to incorporate the chip to a holographic memory system, in order to demonstrate the reconfiguration of an electronic circuit, using a holographic memory, as described in the next section.

## 2.4 The OPGA demo

In this section we discuss the implementation of a system allowing the reconfiguration of the OPGA chip using data stored in a holographic memory. The main advantage of this design over the one using an SLM as the reconfiguration template source, is the considerably lower switching time of holographic memories. The schematic diagram of the setup is shown in Figure 2-14. It is very similar to the setup used to test the APS array with a few modifications and additions, namely a piece of DuPont photopolymer was mounted on a rotational stage, and placed between the SLM and the beam-splitter. The configuration data was recorded on the DuPont photopolymer [2-12] using peristrophic multiplexing [2-17]. The choice

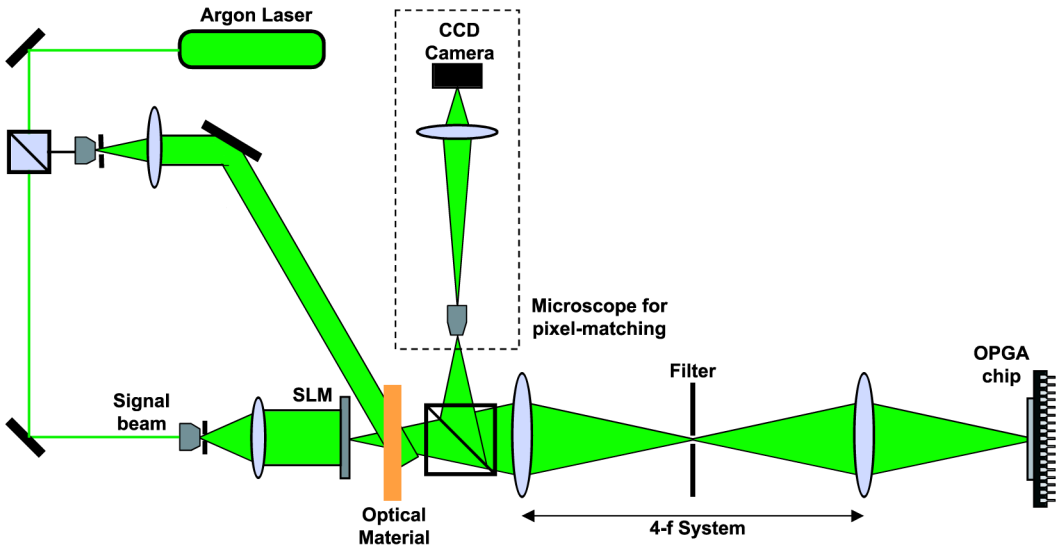
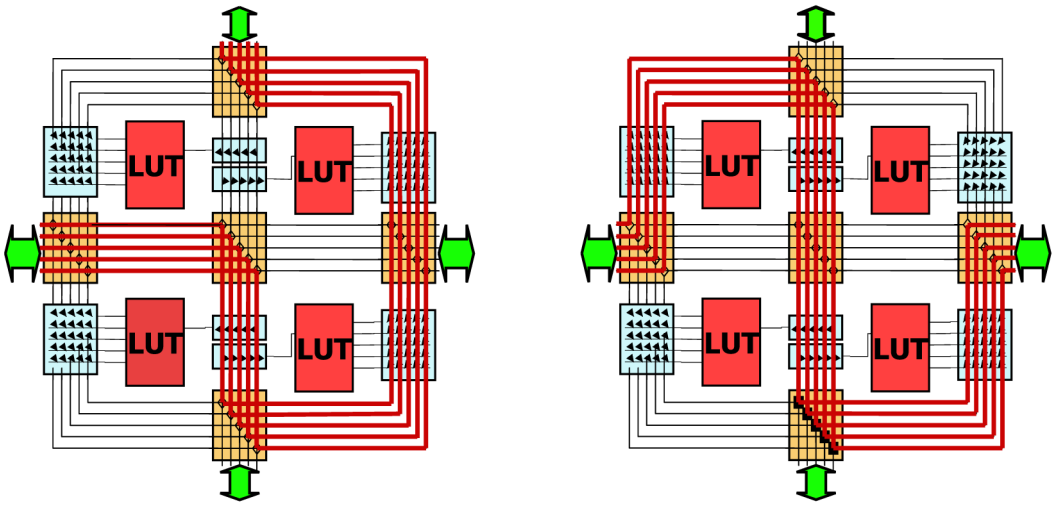



**Figure 2-14.** Schematic diagram of OGPA demo setup.

of this material dictated the change of the laser source, which in this case was an Ar-ion laser, producing cw light at wavelength $\lambda$=514 nm. Part of the raw beam of the laser was tapped-off, collimated, and used as the reference beam for the recording of the reconfiguration information.

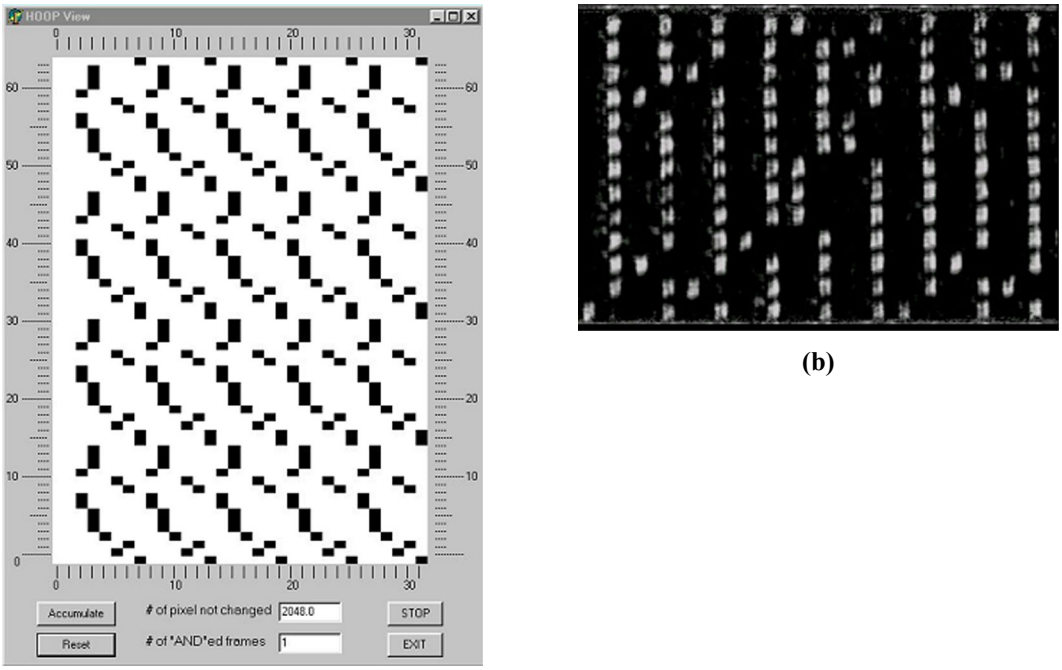
We recorded two different data pages, each of them containing two reconfiguration templates. The first reconfiguration, contained in the first data page, receives an input at the left I/O port of the logic block and routes it to the upper I/O port. The second reconfiguration of the first data page receives an input again from the left I/O port and routes it to the right I/O port. The functionality of the two reconfigurations is illustrated in Figure 2-15. We reconstructed the first data page. The corresponding APS readout is shown in Figure 2-16, along with a detail of the actual hologram.

We can program the electronic board that provides the bias and timing signals to the OPGA to input a counter to the left I/O port. Then, by monitoring the output of the upper and right ports with an oscilloscope (see Figure 2-17), we can confirm the successful configuration of the device. The time it takes for the counter output to switch from the upper to the right I/O is an upper bound for the reconfiguration




(a) (b)

**Figure 2-15.** Functionality of the reconfiguration templates contained in the first data page. Thick lines denote the connections implemented in this reconfiguration.




(a) (b)

**Figure 2-16.** (a) APS readout of the first data page. (b) Detail of the hologram containing the first data page.

time $T_{rcf}$, and is measured to be 127 μs. The fact that we only perform a partial reconfiguration of the device does not affect this estimate. Since the limiting effect

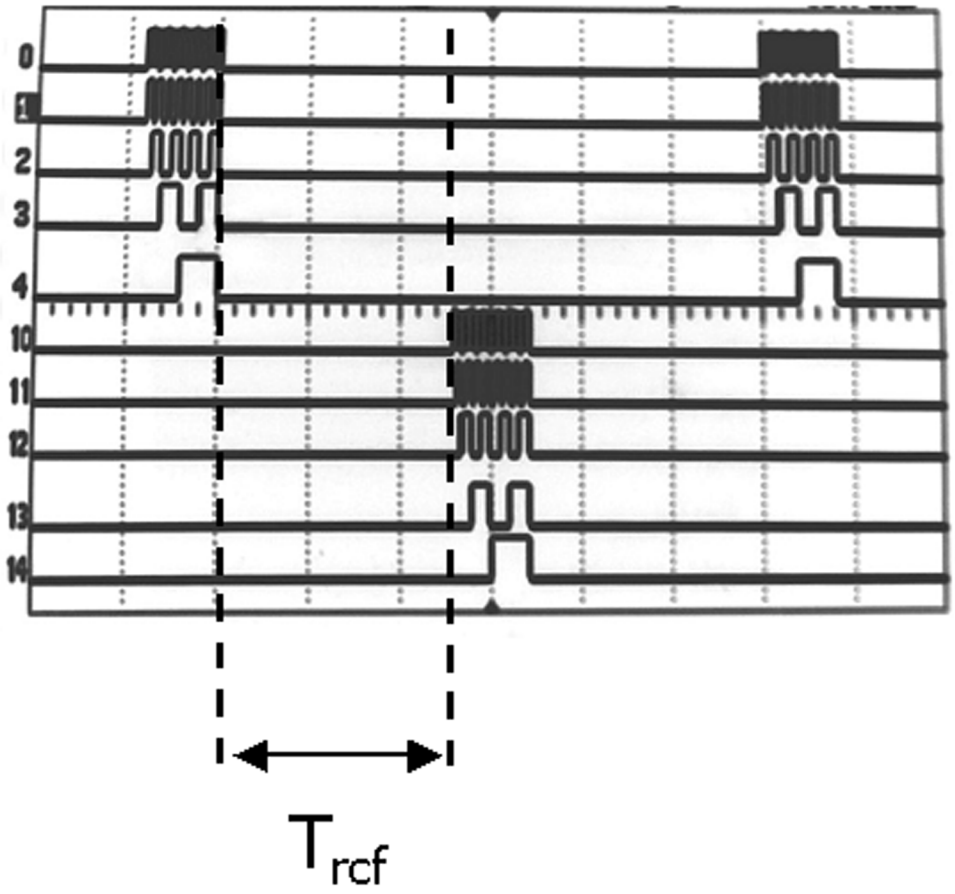
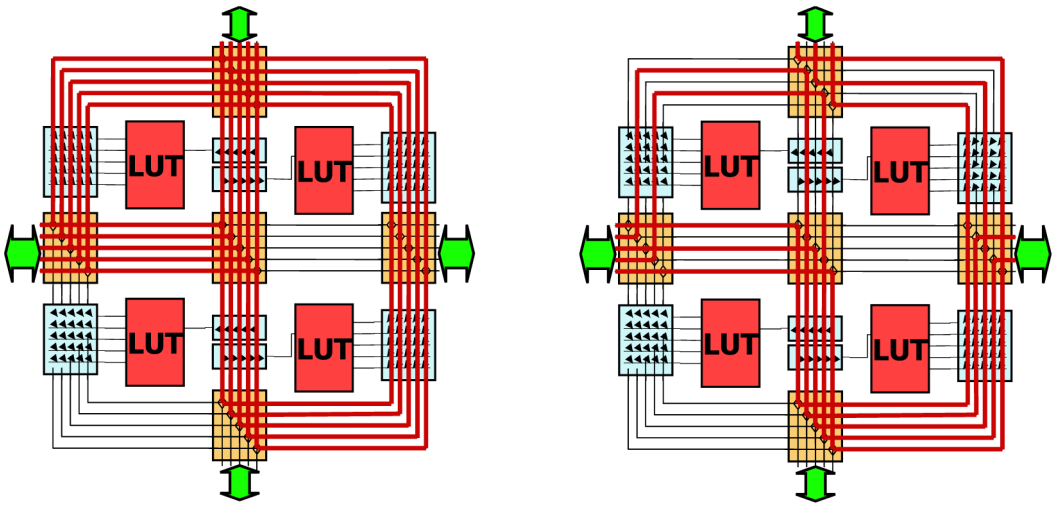



**Figure 2-17.** Output of the OGPA chip undergoing reconfiguration. The time $T_{rcf}$ required to switch between the two output sets an upper bound for the reconfiguration time we can achieve with the device under test, and is measured to be 127 μs.

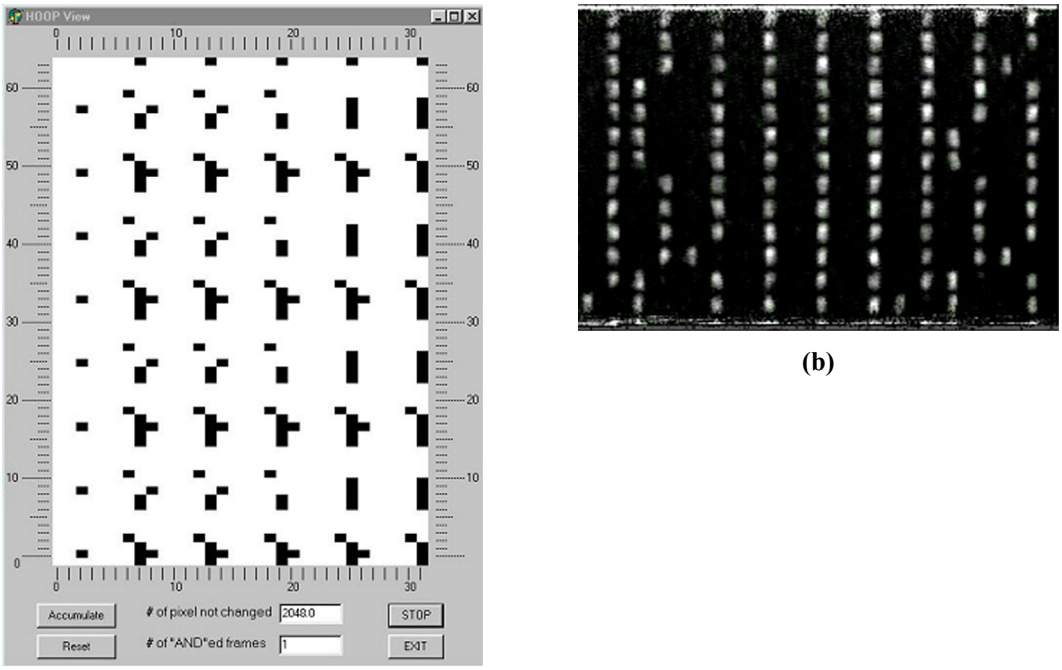
is the photodetector integration, and not the transfer of the reconfiguration data to the logic, the time would be about the same for a full reconfiguration.

The functionality of the two reconfigurations contained in the second data page is illustrated in Figure 2-18. The first one separates the lines of the bus into two groups, odd (1, 3) and even (0, 2, 4). The odd lines are routed to the upper I/O port, while the even lines are routed to the right I/O port. The second reconfiguration of the second data page broadcasts all five bus lines to both I/O ports. By rotating the holographic material, we can reconstruct that second data page. The corresponding APS readout and a detail of the data page hologram are shown in Figure 2-19, while the oscilloscope output is shown in Figure 2-20, confirming the successful configuration of the device.




**(a)** **(b)**

**Figure 2-18.** Functionality of the reconfiguration templates contained in the second data page. Thick lines denote the connections implemented in this reconfiguration.




**(a)** **(b)**

**Figure 2-19.** (a) APS readout of the second data page. (b) Detail of the hologram containing the second data page.

In conclusion, we demonstrated the successful reconfiguration of the OPGA chip using holographic data, contained in a peristrophically multiplexed database. The

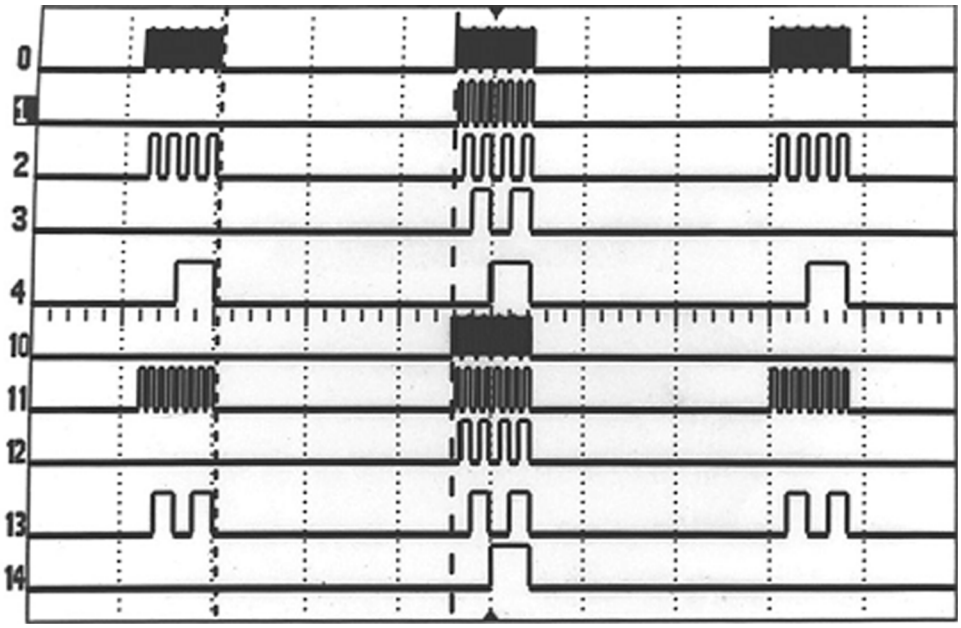



**Figure 2-20.** Output of the OGPA chip undergoing reconfiguration.

extension to larger APS arrays, more reconfigurable logic blocks, and holographic reconfiguration databases with more data pages should be straightforward.

## 2.5 Discussion

In this chapter we discussed the integration of a holographic memory and a silicon circuit combining photodetectors and reconfigurable logic elements, to build the OGPA device. We presented briefly the elements of the device, and focused on the silicon circuit. We characterized the APS array and reconfigurable logic of a demo unit that was fabricated. Finally, we built a setup where we interfaced a holographic memory to the circuit, and demonstrated successful reconfiguration, using optically reconstructed data.

From the characterization of the APS array we conclude that our design can achieve error-free reconstruction for a range of incident intensities. This error-free intensity range is limited, among other things, by the output latches, a point that should be addressed in future implementations. The differential encoding scheme has a good performance, but its complexity exceeds the benefits, therefore it should be dropped. This is especially true if we use MEMS as the addressing

device, since in that case we have enough power to produce reconstruction of the holographic data with high SNR.

From the OPGA demonstration setup we obtained a reconfiguration time $T_{rfc} = 127$ μs, limited by the integration time of the detectors. A clear direction for future research is to investigate how much further this time can be reduced.

Our design should be, in principle, scalable, both in terms of data pages, and in terms of pixels per data page. The implementation of larger holographic databases should be straightforward, therefore, future research should focus on the implementation of larger, distributed pixel sensor arrays.

## 2.6 References

[2-1] D. A. B. Miller, *Rationale and challenges for optical interconnects to electronic chips.* Proceedings of the IEEE, 2000. **88**(6): pp. 728-749.

[2-2] E. Towe, ed. *Heterogeneous optoelectronics integration*. 2000, SPIE: Bellingham.

[2-3] J. Mumbru, G. Zhou, S. Ay, X. An, G. Panotopoulos, F. Mok, and D. Psaltis. *Optically reconfigurable processors*. in *Euro-American Workshop on Optoelectronic Information Processing*. 1999: SPIE.

[2-4] J. Mumbru, G. Panotopoulos, D. Psaltis, X. An, F. Mok, S. Ay, S. Barna, and E. Fossum. *Optically programmable gate array.* in *Proc. SPIE of Optics in Computing*. 2000.

[2-5] D. Psaltis and F. Mok, *Holographic memories.* Scientific American, 1995. **273**(5): pp. 70-76.

[2-6] D. Psaltis and G. W. Burr, *Holographic data storage.* Computer, 1998. **31**(2): p. 52.

[2-7] S. D. Brown, R. J. Francis, J. Rose, and Z. G. Vranesic, *Field-programmable gate arrays*. 1992, Norwell: Kluwer Academic Publishers.

[2-8] P. Stogiannos, A. Dollas, and V. Digalakis, *A configurable logic based architecture for real-time continuous speech recognition using hidden Markov models.* Journal of VLSI Signal Processing Systems for Signal Image and Video Technology, 2000. **24**(2-3): pp. 223-240.

[2-9] A. Benedetti and P. Perona, *A system for real-time 2-D feature detection based on field programmable gate arrays.* Integrated Computer-Aided Engineering, 2000. **7**(3): pp. 181-191.

[2-10] T. H. Szymanski, M. Saint-Laurent, V. Tyan, A. Au, and B. Supmonchai, *Field-programmable logic devices with optical input-output.* Applied Optics, 2000. **39**(5): pp. 721-732.

[2-11] J. Van Campenhout, H. Van Marck, J. Depreitere, and J. Dambre, *Optoelectronic FPGA's.* IEEE Journal of Selected Topics in Quantum Electronics, 1999. **5**(2): pp. 306-315.

[2-12] K. Curtis and D. Psaltis, *Characterization of the Dupont photopolymer for 3-dimensional holographic storage.* Applied Optics, 1994. **33**(23): pp. 5396-5399.

[2-13] R. M. Shelby, D. A. Waldman, and R. T. Ingwall, *Distortions in pixel-matched holographic data storage due to lateral dimensional change of photopolymer storage media.* Optics Letters, 2000. **25**(10): pp. 713-715.

[2-14] K. Curtis, C. Gu, and D. Psaltis, *Crosstalk in wavelength-multiplexed holographic memories.* Optics Letters, 1993. **18**(12): pp. 1001-1003.

[2-15] G. W. Burr, F. H. Mok, and D. Psaltis, *Angle and space multiplexed holographic storage using the 90-degrees geometry.* Optics Communications, 1995. **117**(1-2): pp. 49-55.

[2-16] G. Barbastathis, M. Levene, and D. Psaltis, *Shift multiplexing with spherical reference waves.* Applied Optics, 1996. **35**(14): pp. 2403-2417.

[2-17] K. Curtis, A. Pu, and D. Psaltis, *Method for holographic storage using peristrophic multiplexing.* Optics Letters, 1994. **19**(13): pp. 993-994.

[2-18] F. H. Mok, G. W. Burr, and D. Psaltis, *System metric for holographic memory systems.* Optics Letters, 1996. **21**(12): pp. 896-898.

[2-19] A. V. Krishnamoorthy, K. W. Goossen, L. M. F. Chirovsky, R. G. Rozier, P. Chandramani, W. S. Hobson, S. P. Hui, J. Lopata, J. A. Walker, and L. A. D'Asaro, *16 x 16 VCSEL array flip-chip bonded to CMOS VLSI circuit.* IEEE Photonics Technology Letters, 2000. **12**(8): pp. 1073-1075.

[2-20] E. R. Fossum, *CMOS image sensors: Electronic camera-on-a-chip.* IEEE Transactions on Electron Devices, 1997. **44**(10): pp. 1689-1698.

[2-21] R. H. Nixon, S.E. Kemeny, B. Pain, C.O. Staller, and E.R. Fossum, *256 × 256 CMOS active pixel sensor camera-on-a-chip.* IEEE Journal of Solid-State Circuits, 1996. **31**(12): pp. 2046-2050.

# 3 Applications for the OPGA

## 3.1 Introduction

The OPGA device that we presented in Chapter 2 is a reconfigurable optoelectronic processor, characterized by its very low reconfiguration time. The ability to change the functionality of the processor in a time scale similar to that of the actual computation, gives rise to a new computational paradigm, where the reconfiguration becomes part of the computation.

The most straightforward way to take advantage of the fast reconfiguration time of the OPGA device is to substitute several FPGA devices with a single OPGA device, which will circle through the functionality of FPGA's in real time. This can be though of as "multiplexing" several FPGA's in time, and an application along

those lines will be presented in Section 3.2, where we will show how the OPGA can be used for real-time image processing.

Taking the “multiplexing in time” approach a step further, we can think of applications where the next configuration is decided based on the outcome of the current computation, as opposed to going through a predefined sequence of configurations. Such an application is the classification of cursive digits, discussed in Section 3.3. The hardware presented in the previous chapter uses a read-only holographic memory. In a device with a read-write memory, we could even think of applications, where configurations are developed based on the outcome of computations, thus we have a circuit that can learn and adapt to an evolving environment.

Though the OPGA chip was designed as part of a compact reconfigurable processor, we can use it as a component of a number of different systems. One such system is proposed in Section 3.4, where we use the OPGA chip as a smart head to read out and perform queries on large holographic databases. In Section 3.5 we discuss briefly a number of other interesting applications that are not fully developed in this work.

## 3.2 Real-time video processing

The rapid reconfigurability of the OGPA makes it a good match for time-constrained applications, such as real-time video processing [3-1]. The use of OPGA’s for such an application is illustrated in Figure 3-1. The processing that we want to perform on the video stream is the application of 100 different convolution kernels, one of the most common image processing operations. Assume that we have a camera collecting frames at a rate of 30 frames per second (fps). Each frame is 512 × 512 pixels, monochrome with an 8-bit resolution. Once collected by the camera, each frame is buffered for 33 ms, allowing for a 330 μs processing time per convolution kernel. Assuming that the OGPA has a 64-bit bus and runs at 120 MHz, both rather conservative assumptions, the processing time for a 8 × 8 pixel kernel should be about 280 μs. A reconfiguration time of 50 μs, which is well within

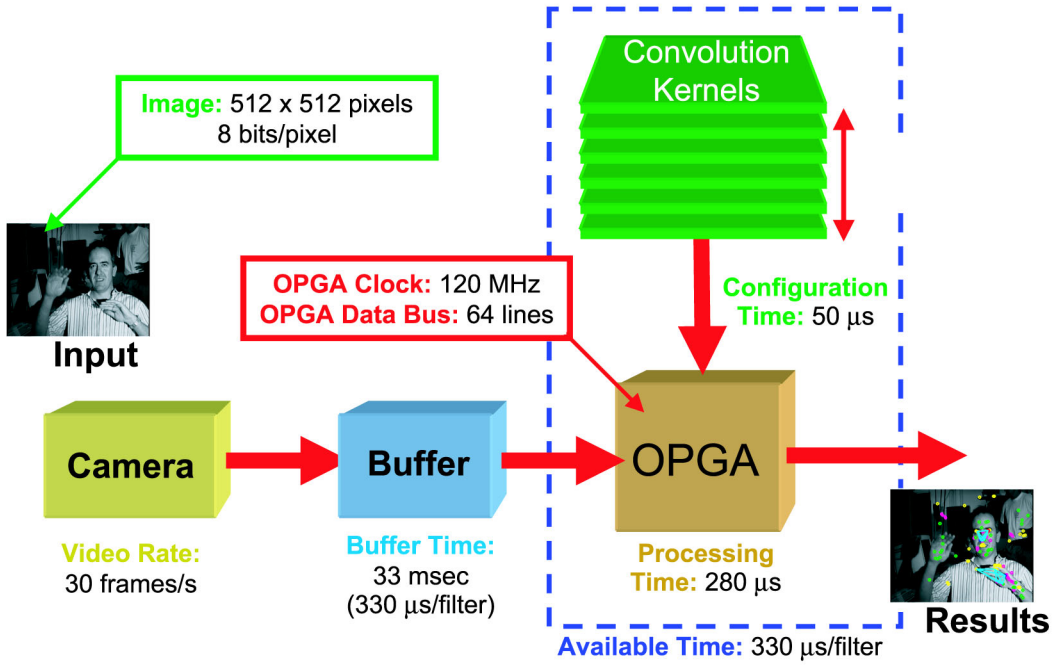



**Figure 3-1.** Schematic diagram of Real-Time Video Processing application using OGPA's.

the capabilities of the OPGA, brings the total processing time to 330 µs per frame, allowing the application of 100 convolution kernels per frame in real time.

By way of contrast, the reconfiguration time of an FPGA is in the order of ms, making it impossible for this device to apply more than one convolution per frame. Therefore a system consisting of FPGA's with the same processing power would have 100 FPGA's, with the associated costs. This comparison suggest an intuitive way to think about the operation of the OGPA in this application: It can be thought of as a device that multiplexes several FPGA's in time. The convolutions that we want to carry out could be performed in parallel by a hundred different processors. Instead, we use a single processor, which successively implements each of those hundred required processors. At this point we should stress the fact that, in this particular application, each reconfiguration is independent of the results of the computation performed by the previous configuration. In the next section we will discuss an application where the opposite is true.

# 3.3 Cursive digit classification

## 3.3.1 Introduction

In this section we discuss an algorithm for cursive digit classification that requires rapid reconfiguration, similar to that offered by the OGPA device. Any pattern classification problem is solved, in its most abstract form, by defining subspaces in a multidimensional space [3-2]. To make this statement more concrete, imagine that we have 8×8 pixel images of handwritten digits. These images lie in a 64-dimensional space. Classifying the digits as 0, 1, 2, etc., is equivalent to defining 10 hypersurfaces, each surrounding the occurrences of the corresponding class. The complexity of the classification lies in the requirement to specify these hypersurfaces.

This complexity becomes evident if we try to perform the classification using limited computational resources, such as the Neural Network (NN) shown in Figure 3-2,

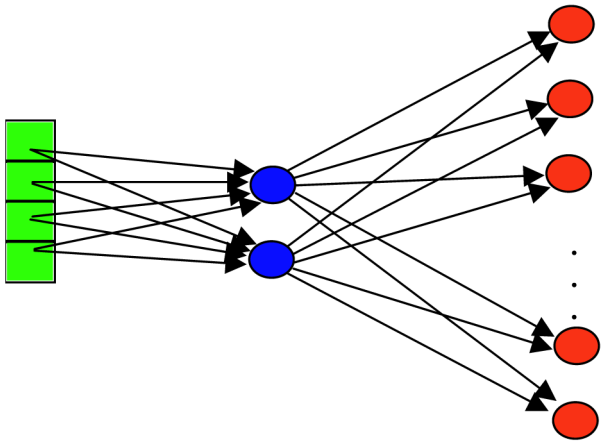


**Figure 3-2.** Simple neural network consisting of one hidden layer with two nodes.

which has one hidden layer with two nodes. We train the network using the back propagation with momentum algorithm [3-3, 3-4]. Our database consists of training set of 3823 digits and a test set of 1797 digits. In Figure 3-3 we plot the correct classification performance of the NN vs. the number of classes it attempts to classify in. Note that when there is one output class (essentially we ask the question: “Is this digit X?”), the probability of obtaining the correct answer is almost 100%. This

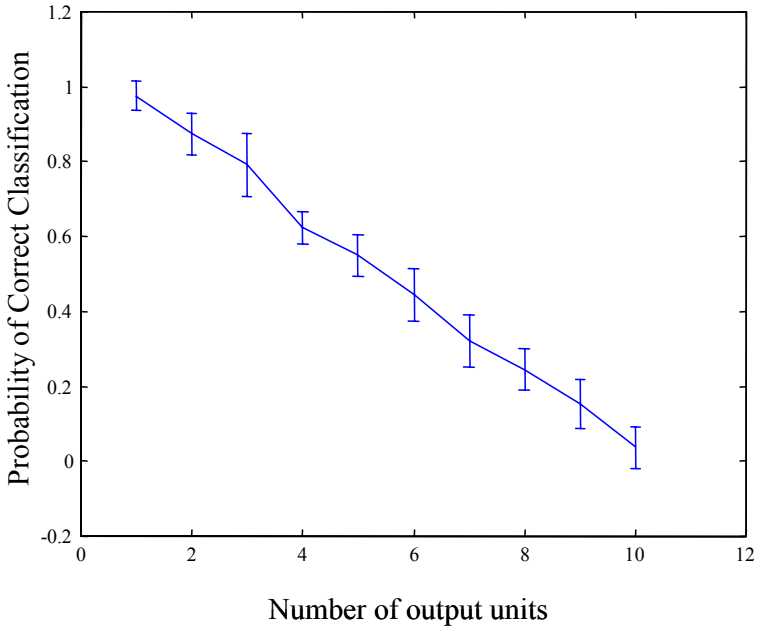



**Figure 3-3.** Probability of correct classification vs. the number of output classes for the Neural Network of Figure 3-2.

probability drops fast as we increase the number of output classes, and when we reach 10 classes, it's almost 10%, meaning we are equally well off assigning each input sample to a random class. The algorithm we present in this section provides a method to tackle the full classification problem, using such limited computational resources.

### 3.3.2 Divide and conquer classification algorithm

A way to simplify the full classification task is to break it down to simpler tasks, so that the final result is reached by navigating through a decision tree. This principle is illustrated in Figure 3-4, where, going back to our example, instead of directly classifying each digit to one of ten classes, we can initially classify it to one of the following three sets: {1,2,3,8}, {0,5,9}, and {4,6,7}[1]. Since we only have three output classes, the probability of correct classification is substantial. Once we have decided in which class the digit belongs, we can proceed to clarify which digit it is. Again, this is a fairly simple question, and we can expect to receive the correct

1. These sets were selected as to maximize the correlation of the digits within each set.

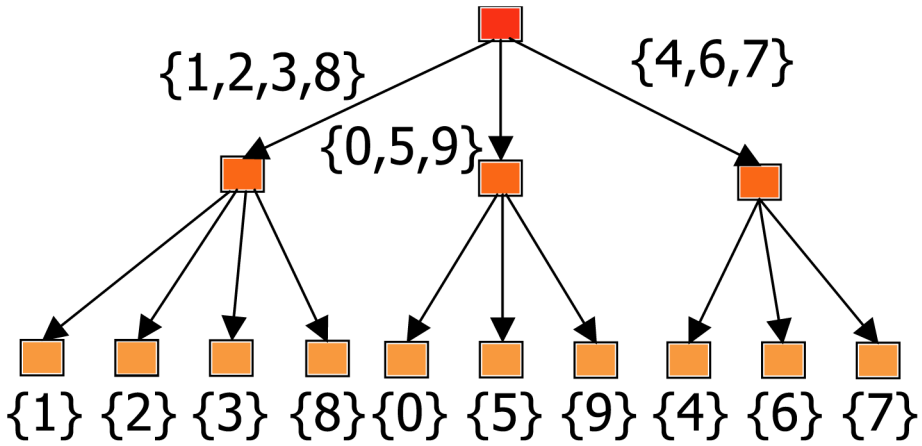



**Figure 3-4.** Decision tree for the divide and conquer classification algorithm.

answer with high probability. Finally, once we have reached the conclusion that this is most likely digit X, we can use a NN with a single output to verify this hypothesis. The graph in Figure 3-3 indicates that this is a virtually error-free decision, resulting in high accuracy for the complete algorithm.

In order to better illustrate the algorithm and understand how it maps to the OGPA device, let's walk through a specific example (see Figure 3-5): We initialize the OPGA device to implement the NN that distinguishes between the three sets {1,2,3,8}, {0,5,9}, and {4,6,7}. Assume that the digit 2 is presented to input, and that the NN correctly classifies it into the first set. Based on this outcome, the next reconfiguration will program the OPGA to implement the NN that distinguishes between digits 1,2,3, and 8. The same input is presented to this network, and let us assume, for the sake of argument, that in this case the NN erroneously concludes that this is the digit 3. This outcome will dictate the next reconfiguration, namely a single output NN that classifies its input as either the digit 3 on not. Since such a NN has very low probability of error, more likely than not it will detect the mistake of the previous layer. The OPGA device is reconfigured to its previous functionality, and the second most likely outcome is selected, say, that the input is the digit 2. We will have one last reconfiguration to the NN that classifies the input as the digit 2 or not, and again, most likely, this will confirm that the input is indeed the digit 2 and the final outcome will be presented at the output.

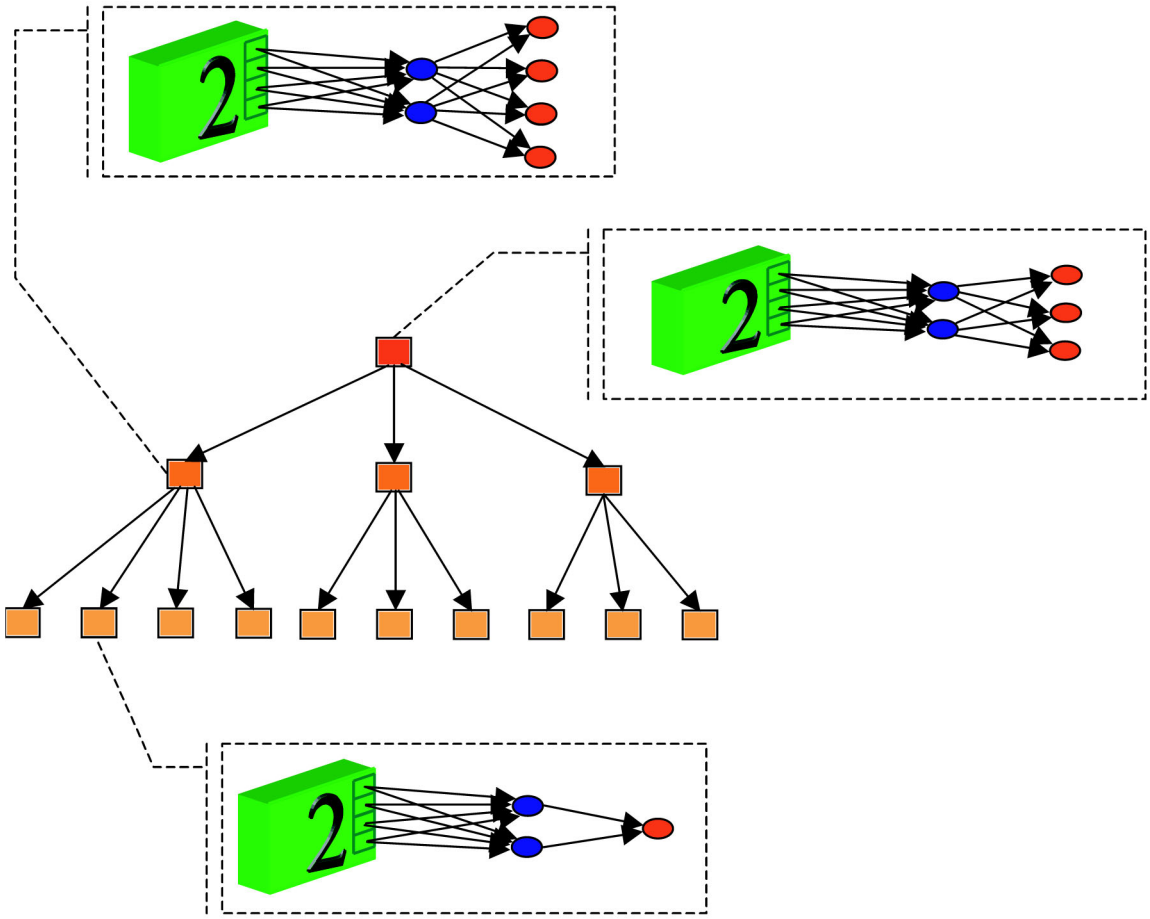



**Figure 3-5.** Mapping of the divide and conquer classification algorithm on the OPGA device.

### 3.3.3 Algorithm performance

To determine the performance of the algorithm we define three metrics:

- The probability of n reconfigurations $P(\mathrm{rec{=}n})$, defining the probability that the algorithm will produce an output after n reconfigurations, regardless of the correctness of that output.
- The probability of correct classification $P_{corr}$, defining the probability that the algorithm will classify the input correctly.
- The probability that the algorithm will end without classifying the input $P_{noClass}$.

The first metric, $P(\mathrm{rec{=}n})$, relates to the speed of the algorithm. Ideally we would like the algorithm to produce an output after as few reconfigurations as possible. The second metric, $P_{corr}$, defines the success rate of the algorithm, and the third

metric, which can be incorporated to the probability of error, quantifies the worse case scenario, where we have to navigate the whole decision tree without reaching any conclusions.

In order to determine the merits of our algorithm, we compared it with three other algorithms, which are outlined here:

- Fully Parallel Classification (FPC): This algorithm would be implemented by having ten FPGA's running in parallel, each implementing a NN with a single output. The input is be presented to all NN's in parallel and the highest output determines the class.
- Sequential Classification (SC): To implement this algorithm we would use an OGPA to successively implement the single output NN's corresponding to digits 0, 1, 2, etc. The output of each successive implementation is compared to a threshold. Once the threshold is exceeded we classify the digit in the corresponding class and abort.
- Exhaustive Sequential Classification (ESC): This algorithm is a sequential version of the Fully Parallel one. An OPGA implements successively the single output NN's. Their outputs are buffered and once we go through all ten of them they are compared, and the highest output determines the class.

Table 3-1 summarizes the results from the comparison of these three algorithms to

| | **Algorithm** | | | |
|---|---|---|---|---|
| **Parameter** | **Fully Parallel (FPC)** | **Divide and Conquer (DAC)** | **Sequential (SC)** | **Exhaustive Sequential (ESC)** |
| Correct classification probability for the training set | 0.98378 | 0.97253 | 0.97332 | 0.98378 |
| Correct classification probability for the test set | 0.93656 | 0.91096 | 0.89705 | 0.93656 |
| Average number of reconfigurations for the training set | 1 | 3.5124 | 5.3421 | 10 |
| Average number of reconfigurations for the test set | 1 | 3.9104 | 5.0534 | 10 |
| Required memory | 10 | 14 | 10 | 10 |
| Required computational resources | 10 | 1 | 1 | 1 |

**Table 3-1.** Performance comparison of the four algorithms used to classify cursive digits.

the divide and conquer (DAC) algorithm. Regarding the probability of correct clas-

sification for the test set, we notice that the FPC and ESC share the best performance, as expected, followed by the DAC and the SC algorithms. Regarding the average number of reconfigurations for the test set, the FPC has the least, the DAC comes second, and the SC and ESC algorithms follow. The FPC has the best performance, both in terms of speed and accuracy. Nevertheless this performance comes at a steep price, since it requires ten times the computational resources required by any other algorithm. The ESC has the same high accuracy performance, but in this case in comes to the expense of execution speed, since it requires the most number of reconfigurations. The DAC algorithm has better performance than the SC, both in terms of speed and accuracy. This performance comes with an overhead in memory requirement, but since we are implementing the memory holographically it is an abundant resource. Overall, the DAC algorithm that we introduced in the previous section offers a nice compromise between speed, accuracy, and computational resources, leveraging on the high capacity of holographic memories.

### 3.3.4 Modeling of algorithm performance

In order to gain a better understanding of the parameters that affect the performance of this algorithm, we formulated a theoretical model and compared it to simulation results. We first assume a 3-layer, balanced, binary decision tree, like the one shown in Figure 3-6. We assume that all the NN's at level $i$ have a specific

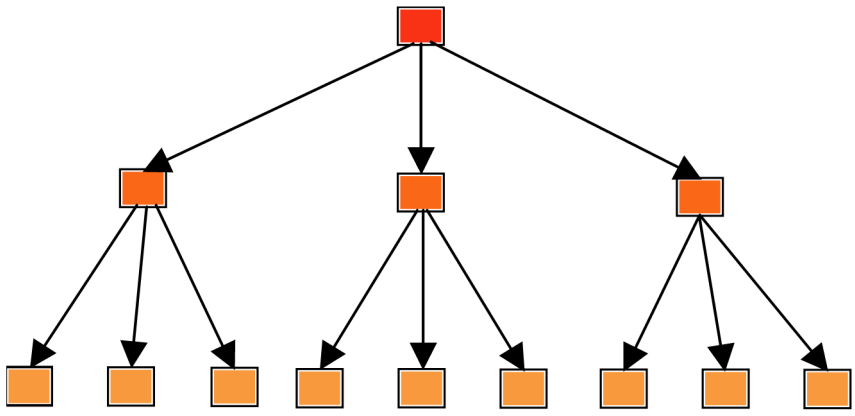


**Figure 3-6.** Three-layer, balanced, binary decision tree.

probability of providing the correct answer, $P_i$, with $i = 1$ being the root. The probabilities defined in Section 3.3.3 on page 37, $P(\text{rec=n})$, $P_{corr}$, and $P_{noClass}$, can be expressed in terms of the probabilities $P_i$. As an example we can calculate $P(\text{rec=3})$. We will have three reconfigurations in the following cases:

- A correct classification is made at all levels $P_1P_2P_3$.
- A correct classification is made at the root, the second layer misclassifies and the third does not catch the mistake $P_1(1-P_2)(1-P_3)$.
- The root makes an erroneous classification, and the third layer does not catch it $(1-P_1)(1-P_3)$.

Therefore $P(\text{rec=3}) = P_1P_2P_3 + P_1(1-P_2)(1-P_3) + (1-P_1)(1-P_3)$. From these terms only $P_1P_2P_3$ will contribute to $P_{corr}$, and none to $P_{noClass}$. All other results can be obtained using similar reasoning.

In order to test our analytical calculations we implemented a simulation algorithm, which implements the tree and goes through several thousand classification trajectories, dictated by the probabilities $P_i$. The theory and simulation are in excellent agreement, as shown in Figure 3-7 and Table 3-2. Note that we cannot have clas-

| Parameter | Analytical value | Simulated value | Normalized difference |
|---|---|---|---|
| $P_{corr}$ | 0.80627 | 0.80619 | $9.3 \times 10^{-5}$ |
| $P_{noClass}$ | 0.0729 | 0.07242 | $3.3 \times 10^{-4}$ |

**Table 3-2.** Comparison of analytical and simulated values for performance measures of the binary, three-layer decision tree.

sification after 5 reconfigurations, since at that point the OPGA would be implementing a node in the second layer of the tree.

We can use the analytical formulas for $P(\text{rec=n})$, $P_{corr}$, and $P_{noClass}$ to investigate the effect of the probabilities $P_i$ on correct classification and reconfiguration time. This information is crucial in determining how to best allocate our computational

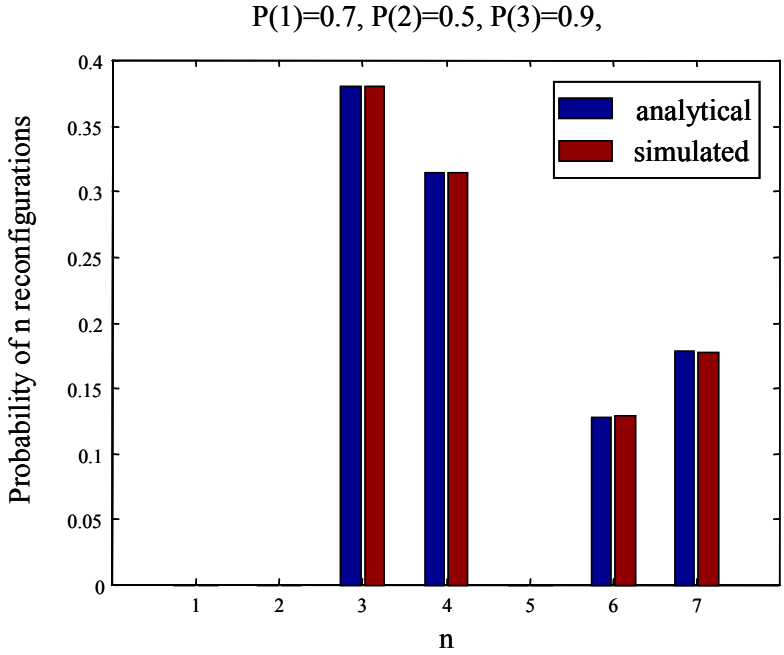



**Figure 3-7.** Probability of n reconfigurations for a binary, three-layer decision tree.

resources in order to optimize performance. In Figure 3-8 we plot the probability of correct classification $P_{corr}$ and the average number of reconfigurations vs. $P_2$ and $P_3$ for specific values of $P_3$. These and similar plots allow us to reach the following conclusions:

- The most important parameter in determining the overall probability of correct classification $P_{corr}$ is the probability of correct classification at the leaf level, $P_3$.
- The probability of correct classification at higher layers determines the average number of reconfigurations, assuming reasonably good performance of the leaf layer.

The first conclusion is intuitively appealing, since if the last layer is always correct, it can compensate for mistakes of higher layer, which only result in a penalty in terms of number of reconfigurations. The second conclusion doesn't come as a surprise either, given that, when a mistake is made at a node, the algorithm must visit the whole subtree before realizing the error. The higher the node, the bigger the subtree. From these two conclusions it follows that in order to optimize the overall performance of the algorithm, we should allocate more computational

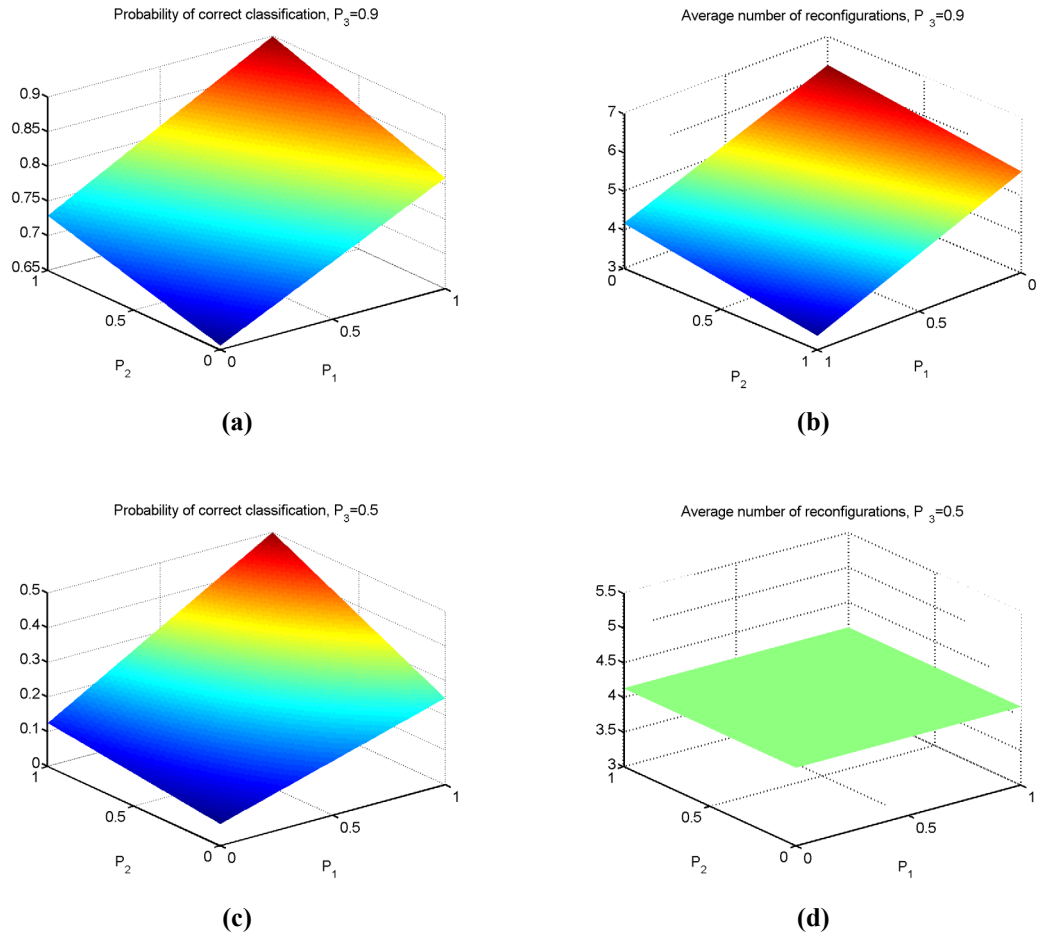



**Figure 3-8.** Performance of the Divide and Conquer classification algorithm as a function of the performance of the individual NN's. (a) Probability of correct classification vs. $P_1$ and $P_2$, given $P_3 = 0.9$. (b) Average number of reconfigurations vs. $P_1$ and $P_2$, given $P_3 = 0.9$. (c) Probability of correct classification vs. $P_1$ and $P_2$, given $P_3 = 0.5$. (d) Average number of reconfigurations vs. $P_1$ and $P_2$, given $P_3 = 0.5$.

resources at higher levels, and make sure that the NN's at the leaf level are as efficient as possible.

The theory was extended to account for 3-layer, balanced, ternary classification trees. In this case, for each level $i$ we define the probability that the right output will appear at order $j$ as $P_{ij}$, so that the probability that the nodes at the leaf level produce the right answer at the first try is $P_{31}$. In this case the theoretical calculations become intractable, therefore we developed an algorithm that extracts all possible trajectories through the search tree and produces the analytical formulas corre-

sponding to $P(\text{rec=n})$, $P_{corr}$, and $P_{noClass}$. A different algorithm, similar to the one developed for the binary case, is used to validate these analytical results. The comparison of the results, shown in Figure 3-9 and Table 3-3, shows excellent agree-

| Parameter | Analytical value | Simulated value | Normalized difference |
|---|---|---|---|
| $P_{corr}$ | 0.75406 | 0.75403 | $3.7 \times 10^{-5}$ |
| $P_{noClass}$ | 0.04305 | 0.04304 | $1.6 \times 10^{-4}$ |

**Table 3-3.** Comparison of analytical and simulated values for performance measures of the ternary, three-layer decision tree.

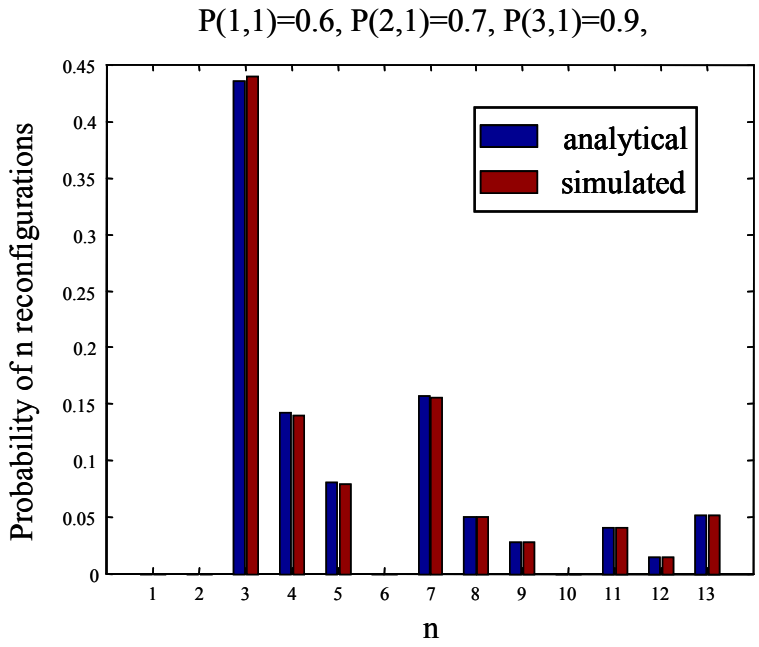



**Figure 3-9.** Probability of n reconfigurations for a ternary, three-layer decision tree.

ment between the two.

The considerations for overall performance optimization discussed in the case of the binary tree hold here as well. Still, there is one additional aspect we need take into account: In the case of the ternary tree, even when a NN does not produce the right answer, it makes a difference whether the correct answer is classified as second or third. This is an aspect that is not addressed by our training algorithm, that only takes into account whether the answer of the NN is correct or not. Incor-

porating this aspect of evaluation into the training algorithm would improve the overall performance of the Divide and Conquer classification algorithm.

### 3.3.5 Discussion

In this section we introduced a new algorithm that performs complex classification tasks by dividing them into simpler ones, which it resolves sequentially. This algorithm takes advantage of the unique properties of the OPGA device. It employes not only its ability of fast reconfiguration, but also the fact that the each reconfiguration can be dictated by the current outcome.

In order to demonstrate the performance of the algorithm, we used it to classify cursive digits, a problem that has been often addressed in literature [3-5]. Our algorithm is a nice compromise between required computational resources and execution time. It minimizes the computational resources requirement with a penalty in time, which is not substantial given the rapid reconfigurability of the OPGA device, and an overhead in memory, which is a resource abundantly available, since it is implemented holographically. Finally, we developed and tested an analytical model of the algorithm, that allows us to understand which are the parameters that affect its performance.

## 3.4 Fast queries in large-scale non-indexed holographic databases

### 3.4.1 Introduction

In Section 3.2 and Section 3.3, we discussed applications that can be mapped directly to the compact OPGA device. In this section we present an application that uses the OPGA chip as part of a larger system. Before going into the details of that system, it is useful to approach the OGPA on a more abstract level. The OGPA chip differs from a common FPGA chip in the fact that it processes an array of photodetectors. In that sense, we can think of it as an integrated circuit that, in addition to the commonly available electronic bus, has an optical input bus. This optical bus allows the input of data at a much higher rate than the electronic bus. The applica-

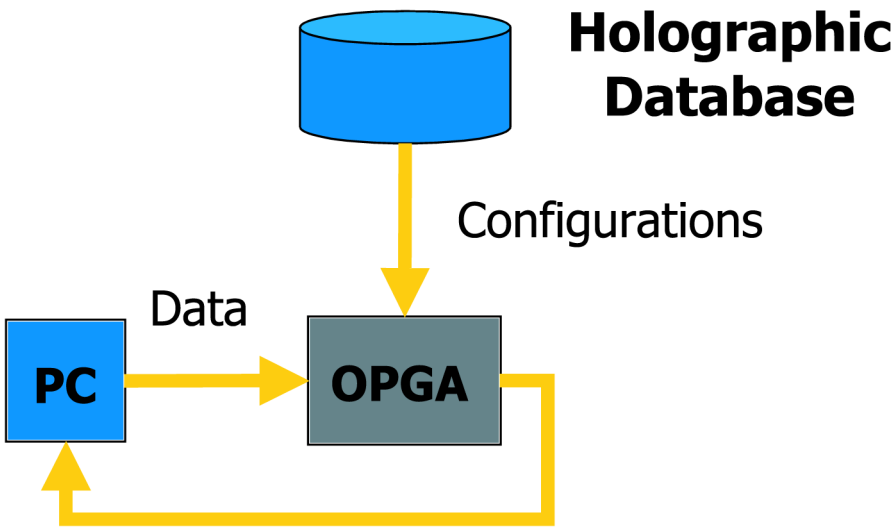
tions discussed previously required fast reconfiguration, therefore, the fast optical bus was used to transfer reconfiguration data. In this section we discuss an application where the operation performed by the OGPA chip is applied to large amounts of data, thus, the optical bus is used to input the data to the device and the electronic bus is used to reconfigure the device and output the outcome of the computation (see Figure 3-10.)

In Section 3.4.2 we introduce the Constellation Model algorithm, an image classification algorithm that relies on local photometry and global geometry. In Section 3.4.3 we discuss a system that performs queries in non-indexed, large, holographic databases, based on the Constellation Model algorithm.

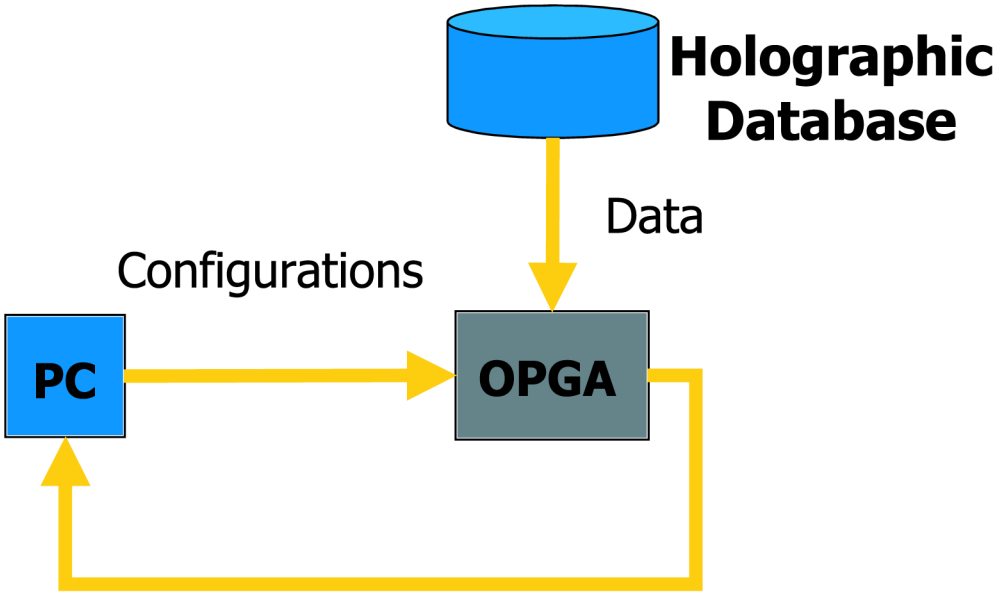
### 3.4.2 The constellation model

The detection of image classes, though performed effortlessly by humans, is a particularly hard problem if we try to approach it using digital computers. The main reason for that is that objects within a category can vary to a large extent, and the extraction of the common properties of the members of the class is far from being straightforward. As an example we can mention trees. Even a child has no problem classifying as trees specimens ranging from bonsai to sequoias, but it is very hard to extract the essence of “treeness,” in a way that would allow a computer to make the distinction between a tree and a nontree.

The Constellation Model algorithm developed by Burl et al. [3-6] proposes the use of local photometry and global geometry to implement recognition of object classes. The main idea behind the algorithm is that objects that belong in a class have common features that have small variations from instance to instance. As an example we can think of human faces. Though it would be particularly hard to find a template that would correlate well with a large number of faces, the opposite is true if we are targeting more restricted features, such as eyes, nostrils, or mouth corners. The problem with the detection of simple features is that usually the algorithms used to implement it produce false positives. The answer to that problem is obtained through the use of global photometry. Going back to the example of human faces, in Figure 3-11 it is easy to distinguish between the assortment of fea-




(a)




(b)

**Figure 3-10.** Different uses of the optical and electronic buses of the OPGA. (a) For applications that require fast reconfiguration the optical bus is used to input the reconfiguration data. (b) When the same reconfiguration is used to process large amounts of data it makes sense to use the optical bus to input that data. In both cases the outcome is output through the electronic bus.

tures that correspond to a human face and those that do not. The Constellation Model algorithm works as follows: First it uses several detectors to detect a number

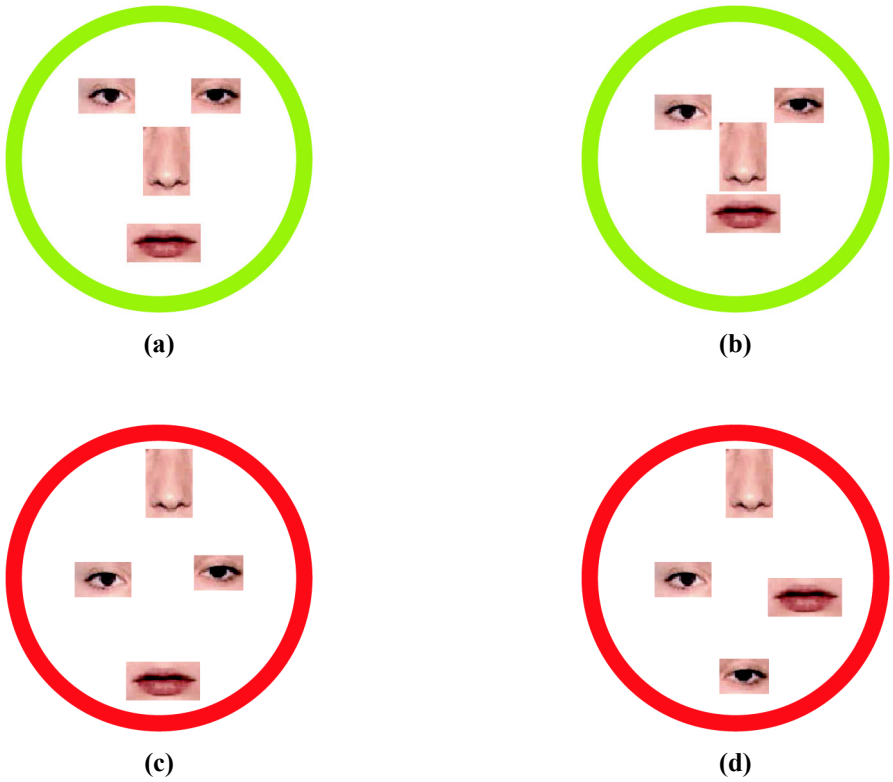



**Figure 3-11.** Constellations composed of face features. The human observer can immediately assess that cases (a) and (b) correspond to human faces, while cases (c) and (d) do not.

of different features, keeping track of which feature was detected at which location. Next it generates combinations of the detected features to form a constellation[1]. For each generated constellation, it computes the probability that it corresponds to the object class in question. Finally, this probabilistic score is compared to a threshold, and thus we reach the conclusion as to whether an object of the class appears in the image.

Further work has been done on the automatic extraction of the Constellation Model by Weber et al. [3-7]. In this context an unsupervised algorithm processes a number of images, and uses a clustering technique to identify features. It then proceeds to the extraction of Constellation Models for the object classes in the training images.

1. The model takes occlusion into account, as a result some of the generated constellations do not have all features.

### 3.4.3 The application

In this section we discuss how the Constellation Model algorithm can be used to perform queries into large, non-indexed, holographic databases. Assume that we have large number of images stored in a Holographic Disk [3-8, 3-9], and we want to identify those stored images that contain objects of a specific class. We first feed several instances of the class to the algorithm that extracts the Constellation Model for that class (see Figure 3-12). Once the features of the model have been

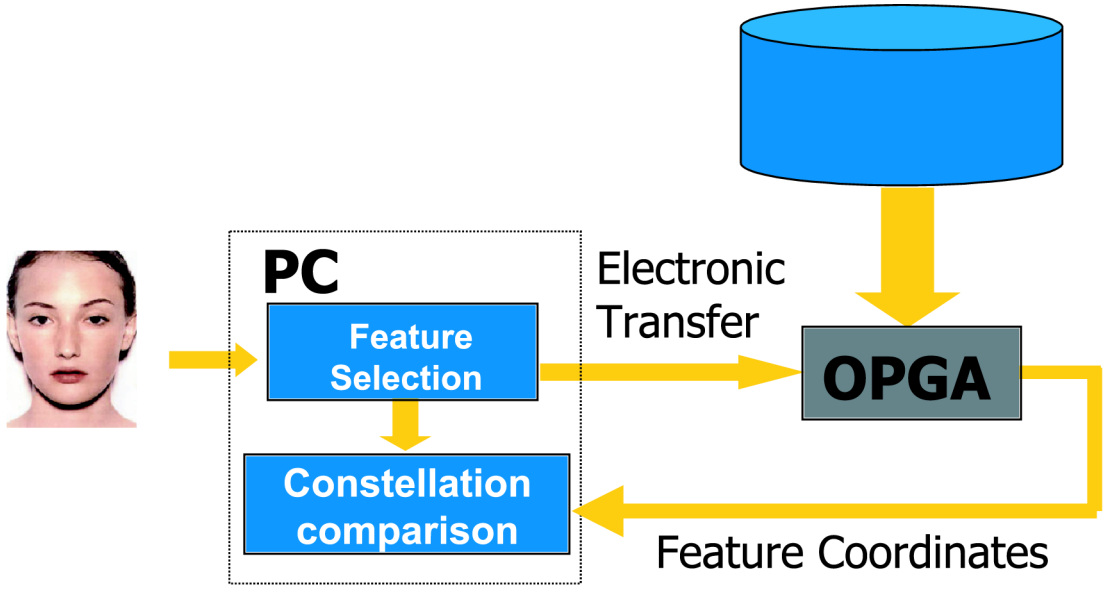



**Figure 3-12.** Schematic diagram of the system that performs queries in large, non-indexed holographic databases using the Constellation Model algorithm. In this case OPGA's are used as smart heads that detect features in images stored in the database.

selected, we program a number of OGPA's to detect those features. In this context OPGA chips act as smart readout heads for the holographic disk (see Figure 3-13). As the disk rotates images are transferred to the OPGA photodetector array and convolutions are run to locate features. The output of each OPGA chip for each image is the coordinates of the feature it has been programmed to detect on that image. The outputs of all OPGA heads are collected and combined to form models, which are then assigned probabilistic scores. These scores allow the selection of those images that contain objects of the target class. Note that the OPGA chip is used in this case as an extremely efficient interface between the high volume holo-

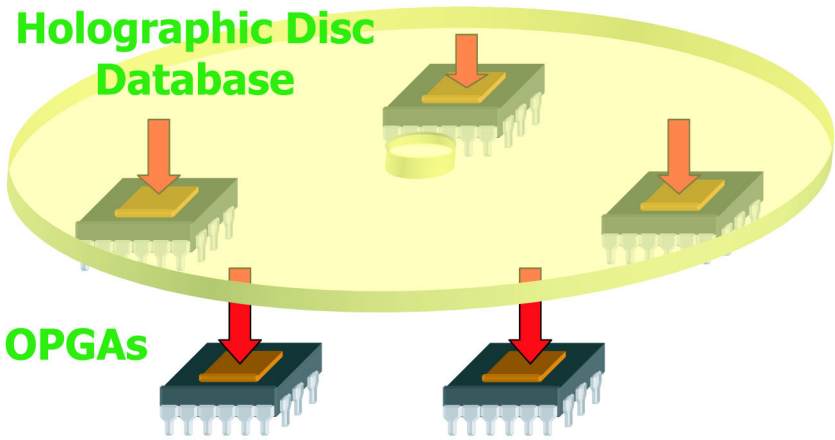



**Figure 3-13.** OPGA's act as smart heads that read out the data from the holographic disk database.

graphic database (images) and the low dimensional output (coordinates of constellation features).

## 3.5 Other applications

A number of other applications have been developed for the OGPA that are not covered in this work. One of the most interesting ones is the use of OGPA's for neural prosthetics, developed by Shenoy et al. [3-10, 3-11]. In that project neural signals produced by the motion planning centers of the brain are used to control a robotic arm. The compactness and processing power of the OPGA could contribute significantly to the creation of implantable devices.

Another application is the use of the OPGA chip for 4-D imaging (3 spatial dimensions and frequency) [3-12]. In this case the optical input to the OPGA chip is provided by a natural scene that is diffracted from a holographic element in a way that

maps the 4-D information of the scene to the 2-D surface of the chip, where it can be processed using the available logic (see Figure 3-14.)

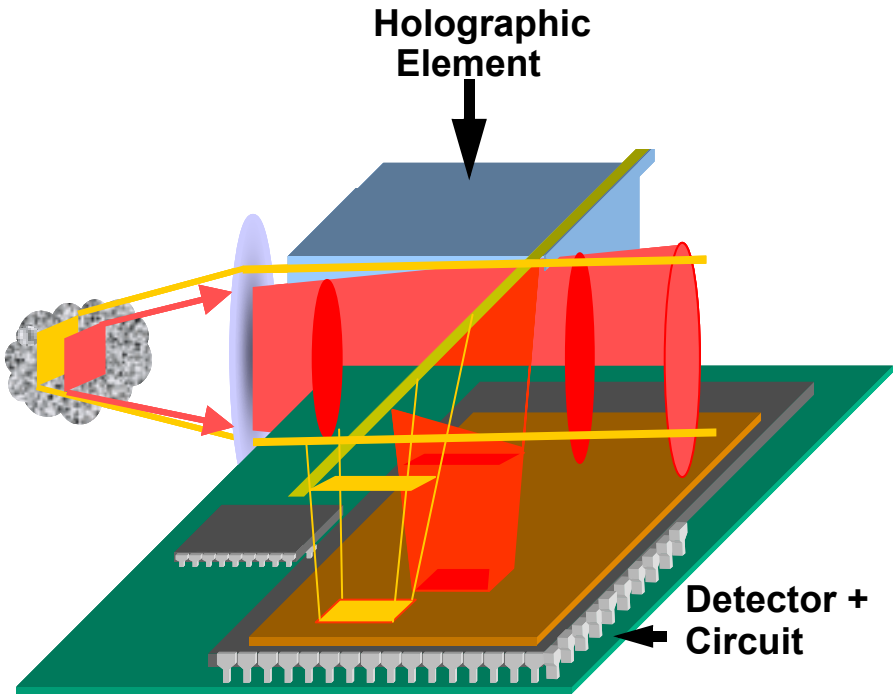



**Figure 3-14.** Extension of the OGPA chip used as a detector for a 4-D holographic imaging system.

## 3.6 Discussion

In this chapter we presented a number of applications that utilize the unique properties of the OPGA. The capability of the device to reconfigure itself in a time scale comparable to that associated with computation, gives rise to a new computational paradigm, where the reconfiguration becomes part of the computation. We presented several applications that illustrate the point. In Section 3.2 we showed how a single OPGA can substitute a large number of FPGA's in a real-time video processing task. In Section 3.3 we mapped a divide and conquer classification algorithm to the OPGA, taking advantage of the ability of the device to reconfigure itself selectively, based on the outcome of the previous computation.

We also investigated how the OPGA chip can be incorporated in various information-processing systems, such as large holographic databases, as we discussed in

Section 3.4. Finally, we mentioned some even more far-reaching extensions of the OPGA architecture in Section 3.5.

From the wealth of applications presented in this chapter, it becomes evident that much is to be gained by the combination of photodetectors and processing logic in a single circuit. The large data-transfer rates, achieved through the optical interface, allow us to take advantage of the high processing bandwidth that is available on-chip, and which cannot be put in full use otherwise, due to the restrictive, low-bandwidth, electronic interface.

## 3.7 References

[3-1] A. Benedetti and P. Perona, *A system for real-time 2-D feature detection based on field programmable gate arrays.* Integrated Computer-Aided Engineering, 2000. **7**(3): pp. 181-191.

[3-2] J. Daugman, *Face and gesture recognition: Overview.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997. **19**(7): pp. 675-676.

[3-3] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning representations by back-propagating errors.* Nature, 1986. **323**(6088): pp. 533-536.

[3-4] D. Plaut, S. Nawlan, and G. Hinton, *Experiments of learning by back propagation. Technical Report CMU-CS-86-126*. 1986, Carnegie Mellon University: Pittsburgh.

[3-5] J. L. Blue, G. T. Candela, P. J. Grother, R. Chellappa, and C. L. Wilson, *Evaluation of pattern classifiers for fingerprint and OCR applications.* Pattern Recognition, 1994. **27**(4): pp. 485-501.

[3-6] M. C. Burl, L. Asker, P. Smyth, U. Fayyad, P. Perona, L. Crumpler, and J. Aubele, *Learning to recognize volcanoes on Venus.* Machine Learning, 1998. **30**(2-3): pp. 165-194.

[3-7] M. Weber, M. Welling, and P. Perona, *Unsupervised learning of models for recognition.* Computer Vision - ECCV 2000, Pt I, Proceedings, 2000. **1842**: pp. 18-32.

[3-8] D. Psaltis and A. Pu, *Holographic 3-D disks.* Optoelectronics-Devices and Technologies, 1995. **10**(3): pp. 333-342.

[3-9] A. Pu and D. Psaltis, *High-density recording in photopolymer-based holographic three-dimensional disks.* Applied Optics, 1996. **35**(14): pp. 2389-2398.

[3-10] K. V. Shenoy, S. A. Kureshi, D. Meeker, B. L. Gillikin, D. J. Dubowitz, A. P. Batista, C. A. Buneo, S. Cao, J. W. Burdick, and R. A. Andersen. *Toward prosthetic systems controlled by parietal cortex*. Society for Neuroscience, 1999(**25**).

[3-11] J. Mumbru, *Optoelectornic circuits using holographic elements*, Ph.D. Thesis, California Institute of Technology, 2002.

[3-12] W. H. Liu, D. Psaltis, and G. Barbastathis, *Real-time spectral imaging in three spatial dimensions.* Optics Letters, 2002. **27**(10): pp. 854-856.

# 4 Strong volume gratings in the reflection geometry

## 4.1 Introduction

In this chapter we investigate the role of strong volume gratings as a means to control information-bearing light beams. The original motivation stems from the use of volume gratings to implement interconnects for a chip. This approach is closely related to the Optical Programmable Gate Array device, presented in Chapter 2. Indeed, if we replace the holographic memory of the device with a holographic element where we superimpose strong gratings, we can use those to allow different parts of the chip to communicate with each other, using Vertical Cavity Surface Emitting Lasers (VCSEL) to transmit information and photodetectors to receive it, as shown in Figure 4-1.

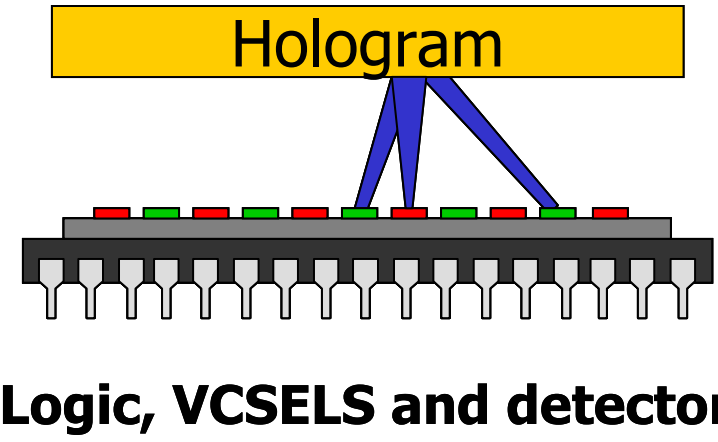



**Figure 4-1.** Holographic element with strong reflection gratings, used to implement intra-chip interconnects. VCSEL's are denoted by red and photodetectors by green.

The same gratings can be used in the field of Wavelength Division Multiplexing (WDM) optical fiber communications, both as filters [4-1, 4-2] and as de-multiplexers [4-3]. Using holography we can record periodic refractive index structures with tens of thousands of periods, in devices of length to the order of cm. Other technologies, such as thin-film filters, offer much stronger refractive index modulation, but a substantially smaller number of periods. The high number of periods results in filters with very fine Bragg selectivity, while the lack of modulation depth can be easily complemented by the use of longer devices. In addition, the free-space nature of volume gratings gives rise to degrees of freedom not present in more established technologies, such as fiber Bragg gratings.

In the remainder of this chapter we will develop the theoretical background necessary for the treatment of diffraction from strong volume gratings. We will start with a general treatment of waves in two-dimensional periodic media in Section 4.2. Then we will focus on the reflection geometry and investigate some of the finer aspects of the coupled mode theory, with applications to polarization dependence, angular, wavelength, and thermal detuning, in Section 4.3. At the end of the section we will present experimental results confirming our theories. Finally, in Section 4.4, we will discuss the implications of our findings to the design of WDM filters.

A portion of the theoretical material in this chapter is well documented in literature [4-4]. It is presented here with minor extensions, for the sake of completeness and notation uniformity, but mainly to give us the opportunity to accentuate what we consider to be points of particular interest, and to compare and contrast different approaches to the problem of strong volume gratings in reflection geometry.

# 4.2 Waves in two-dimensional periodic media

## 4.2.1 Introduction

In this section we use the Bloch formalism to draw conclusions on the properties of mode-coupling in 2-D periodic media. Note that, in this analysis, we assume that both the medium and the periodicity are two-dimensional. Some of the properties will be used to justify, otherwise intuitive, arguments that will be made in Section 5.2.2, where we will derive the coupled-mode equation for the $90^{o}$ geometry.

## 4.2.2 Modes in periodic media

We know that in isotropic media eigenmodes are plain waves. What happens if we impose a periodic modulation to the dielectric coefficient of the medium $\varepsilon$? Assume a medium with $\varepsilon(\grave{r}) = \varepsilon(\grave{r} + \grave{a})$ and magnetic susceptibility $\mu$ constant. The question that we are going to answer here is, which solutions to the wave equation can exist independently in this medium.

Since the dielectric coefficient is periodic, we can express it as a Fourier Series:

$$\varepsilon(\grave{r}) = \sum_{l} \varepsilon_l e^{-jl\grave{g}\grave{r}}, \qquad \textbf{(EQ 4-1)}$$

where $l$ is an integer, $\grave{a}$ is the lattice vector, and $\grave{g} = \frac{2\pi}{|a|^2}\grave{a}$ is the reciprocal-lattice vector.

There is some electric field $\grave{E}$ in the medium, for which at this point we know very little. Still this unknown field has a Fourier Transform:

$$\vec{E} = \int \vec{A}(\vec{k}) e^{-j\vec{k}\vec{r}} d\vec{k}. \quad \textbf{(EQ 4-2)}$$

The other thing we know about $\vec{E}$ is that it will satisfy Maxwell's equations:

$$\nabla\times(\nabla\times\vec{E}) - \omega^2\mu\varepsilon\vec{E} = 0 \quad \textbf{(EQ 4-3)}$$

Plugging Equation 4-1 and Equation 4-2 into Equation 4-3 we get[1]:

$$\nabla\times(\nabla\times\int \vec{A}(\vec{k}) e^{-j\vec{k}\vec{r}} d\vec{k}) - \omega^2\mu\varepsilon\int \vec{A}(\vec{k}) e^{-j\vec{k}\vec{r}} d\vec{k} = 0 \Rightarrow$$

$$\int \vec{k}\times(\vec{k}\times\vec{A}(\vec{k})) e^{-j\vec{k}\vec{r}} d\vec{k} + \omega^2\mu\sum_l \int \varepsilon_l \vec{A}(\vec{k}) e^{-j(l\vec{g}+\vec{k})\vec{r}} d\vec{k} = 0 \Rightarrow$$

$$\int \vec{k}\times(\vec{k}\times\vec{A}(\vec{k})) e^{-j\vec{k}\vec{r}} d\vec{k} + \omega^2\mu\sum_l \int \varepsilon_l \vec{A}(\vec{k}-l\vec{g}) e^{-j\vec{k}\vec{r}} d\vec{k} = 0.$$

So we have:

$$\int \vec{k}\times(\vec{k}\times\vec{A}(\vec{k})) e^{-j\vec{k}\vec{r}} d\vec{k} + \omega^2\mu\sum_l \int \varepsilon_l \vec{A}(\vec{k}-l\vec{g}) e^{-j\vec{k}\vec{r}} d\vec{k} = 0. \quad \textbf{(EQ 4-4)}$$

For Equation 4-4 to be true, all the coefficients of $e^{-j\vec{k}\vec{r}}$ must be equal to $0$, so we have:

$$\vec{k}\times(\vec{k}\times\vec{A}(\vec{k})) + \omega^2\mu\sum_l \varepsilon_l \vec{A}(\vec{k}-l\vec{g}) = 0. \quad \textbf{(EQ 4-5)}$$

Now comes a subtle argument: From Equation 4-5 it should be obvious that out of all the Fourier components of a possible solution $\vec{E}$, only those separated by $\vec{g}$ are coupled. So, if a field component of specific spatial frequency $\vec{K}$ is present in the medium, the complete solution should be composed of it and other frequency components separated by integer multiples of $\vec{g}$. Another way of stating this is that the periodicity of the medium imposes a coupling between harmonic components of the electric field. This last statement will form the basis of Coupled-Mode Analysis, which will be presented in Section 4.3.

1. Taking the FT of $\vec{E}$ is a nice way to be able to handle the double curl in a tractable fashion.

Let us assume a solution corresponding to a specific spatial frequency $\vec{k}$. We have $\vec{E} = \sum_l \vec{A}(\vec{k} - l\vec{g}) e^{-j(\vec{k} - l\vec{g})\cdot\vec{r}} = e^{-j\vec{k}\cdot\vec{r}} \sum_l \vec{A}(\vec{k} - l\vec{g}) e^{-jl\vec{g}\cdot\vec{r}} = e^{-j\vec{k}\cdot\vec{r}} \vec{E}_K(\vec{r})$. So the solution is in the form:

$$\vec{E}(\vec{r}) = e^{-j\vec{k}\cdot\vec{r}} \vec{E}_K(\vec{r}), \qquad \textbf{(EQ 4-6)}$$

with

$$\vec{E}_K(\vec{r}) = \sum_l \vec{A}(\vec{k} - l\vec{g}) e^{-jl\vec{g}\cdot\vec{r}}, \qquad \textbf{(EQ 4-7)}$$

for which obviously

$$\vec{E}_K(\vec{r} + \vec{a}) = \vec{E}_K(\vec{r}). \qquad \textbf{(EQ 4-8)}$$

### 4.2.3 Modes in media with two-dimensional periodicity

We assume that we have a periodic modulation of the dielectric constant in the $xy$-plane. Then Equation 4-1 becomes

$$\varepsilon(x, y) = \sum_m \sum_n \varepsilon_{mn} e^{-j\left(\frac{2\pi}{\Lambda_x} mx + \frac{2\pi}{\Lambda_y} ny\right)}. \qquad \textbf{(EQ 4-9)}$$

Accordingly Equation 4-5 will become

$$\vec{k} \times [\vec{k} \times \vec{A}(\vec{k})] + \omega^2 \mu \sum_m \sum_n \varepsilon_{mn} \vec{A}(\vec{k} - m\vec{g}_x - n\vec{g}_y) = 0, \qquad \textbf{(EQ 4-10)}$$

where $\vec{g}_x = \frac{2\pi}{\Lambda_x}\hat{x}$ and $\vec{g}_y = \frac{2\pi}{\Lambda_y}\hat{y}$, with $\hat{x}$ and $\hat{y}$ the unit vectors along the $x$ and $y$ axes respectively. Again using the same arguments we have:

$$\vec{E}_K(x, y) = \sum_m \sum_n \vec{A}(\vec{k} - m\vec{g}_x - n\vec{g}_y) e^{-j\left(\frac{2\pi}{\Lambda_x} mx + \frac{2\pi}{\Lambda_y} ny\right)}, \qquad \textbf{(EQ 4-11)}$$

and

$$\vec{E}(x, y) = \vec{E}_K(x, y) e^{-j(k_x x + k_y y)}. \qquad \textbf{(EQ 4-12)}$$

We have assumed $\vec{E}(x,y) = E(x,y)\hat{z}$, therefore $\vec{A}(\vec{k}) = A(\vec{k})\hat{z}$. We have also assumed $\vec{k} = k_x\hat{x} + k_y\hat{y}$, so we have $\vec{k}\times\vec{A} = -k_xA\hat{y} + k_yA\hat{x}$ and

$\vec{k}\times(\vec{k}\times\vec{A}) = (k_x\hat{x} + k_y\hat{y})\times(-k_xA\hat{y} + k_yA\hat{x}) = -\hat{z}(k_x^2 + k_y^2)A$. So Equation 4-10 will become

$$k^2A(\vec{k}) - \omega^2\mu\sum_m\sum_n\varepsilon_{mn}\vec{A}(\vec{k} - m\vec{g}_x - n\vec{g}_y) = 0, \qquad \textbf{(EQ 4-13)}$$

where we adopt the notation $k = |\vec{k}|$.

As mentioned before, and as we can see from Equation 4-13, only modes separated by the Fourier components of the dielectric constant will be coupled with each other. Still which modes $A(\vec{k})$ are strongly coupled? In answering this question it would be helpful to write a few terms of Equation 4-13:

$$\begin{aligned} &k^2A(\vec{k}) - \omega^2\mu\varepsilon_{00}A(\vec{k}) - \omega^2\mu\varepsilon_{10}A(\vec{k} - \vec{g}_x) - \omega^2\mu\varepsilon_{01}A(\vec{k} - \vec{g}_y) \\ &- \omega^2\mu\varepsilon_{11}A(\vec{k} - \vec{g}_x - \vec{g}_y) - \ldots = 0. \end{aligned} \qquad \textbf{(EQ 4-14)}$$

Solving for $A(\vec{k})$ we get

$$A(\vec{k}) = \frac{1}{k^2 - \omega^2\mu\varepsilon_{00}}(\omega^2\mu\varepsilon_{10}A(\vec{k} - \vec{g}_x) + \omega^2\mu\varepsilon_{01}A(\vec{k} - \vec{g}_y) + \ldots). \qquad \textbf{(EQ 4-15)}$$

By setting $\vec{k} = \vec{k} - \vec{g}_x$ in Equation 4-13, Equation 4-14 becomes

$$\begin{aligned} &|\vec{k} - \vec{g}_x|^2A(\vec{k} - \vec{g}_x) - \omega^2\mu\varepsilon_{00}A(\vec{k} - \vec{g}_x) - \omega^2\mu\varepsilon_{10}A(\vec{k} - 2\vec{g}_x) \\ &- \omega^2\mu\varepsilon_{01}A(\vec{k} - \vec{g}_x - \vec{g}_y) - \omega^2\mu\varepsilon_{11}A(\vec{k} - 2\vec{g}_x - \vec{g}_y) - \omega^2\mu\varepsilon_{-10}A(\vec{k}) - \ldots = 0. \end{aligned} \qquad \textbf{(EQ 4-16)}$$

Solving for $A(\vec{k} - \vec{g}_x)$ we get

$$A(\vec{k} - \vec{g}_x) = \frac{1}{|\vec{k} - \vec{g}_x|^2 - \omega^2\mu\varepsilon_{00}}(\omega^2\mu\varepsilon_{-10}A(\vec{k}) + \omega^2\mu\varepsilon_{10}A(\vec{k} - 2\vec{g}_x) + \ldots). \qquad \textbf{(EQ 4-17)}$$

In the most general case we have

$$\begin{aligned} A(\vec{k} - m\vec{g}_x - n\vec{g}_y) &= \frac{1}{|\vec{k} - m\vec{g}_x - n\vec{g}_y|^2 - \omega^2\mu\varepsilon_{00}}(\omega^2\mu\varepsilon_{-m-n}A(\vec{k}) \\ &+ \omega^2\mu\varepsilon_{10}A(\vec{k} - (m+1)\vec{g}_x - n\vec{g}_y) + \omega^2\mu\varepsilon_{01}A(\vec{k} - m\vec{g}_x - (n+1)\vec{g}_y) + \ldots). \end{aligned} \qquad \textbf{(EQ 4-18)}$$

Assume now that we have a mode that meets the propagation condition:

$$k^2 \cong \omega^2 \mu \varepsilon_c \qquad \textbf{(EQ 4-19)}$$

and there exists $\grave{g}_0 = m_0 \grave{g}_x + n_0 \grave{g}_y$, such that:

$$\left| \grave{k} - \grave{g}_0 \right|^2 \cong \omega^2 \mu \varepsilon_c , \qquad \textbf{(EQ 4-20)}$$

where $\varepsilon_c = \varepsilon_{00}$. It should be obvious from Equation 4-15 to Equation 4-18 that only the modes $A(\grave{k})$ and $A(\grave{k} - \grave{g}_0)$ will be important, and all the rest can be neglected. If we neglect everything but these two modes these equations boil down to:

$$(k^2 - \omega^2 \mu \varepsilon_0) A(\grave{k}) - \omega^2 \mu \varepsilon_{m_0 n_0} A(\grave{k} - \grave{g}_0) = 0 , \qquad \textbf{(EQ 4-21)}$$

$$-\omega^2 \mu \varepsilon_{m_0 n_0}{}^* A(\grave{k}) + (\left| \grave{k} - \grave{g}_0 \right|^2 - \omega^2 \mu \varepsilon_0) A(\grave{k} - \grave{g}_0) = 0 , \qquad \textbf{(EQ 4-22)}$$

where $\varepsilon_{-m_0 - n_0} = \varepsilon_{m_0 n_0}{}^*$, since we assume that $\varepsilon(x, y)$ is real.

Let us now step back from the algebraic details, and give a more intuitive overview: We started with the assumption that we have a periodic modulation of the dielectric constant of the medium. Using that fact we proved that, from all possible plane waves that can exist in the medium, only those separated by the grating vectors (equivalently Fourier components of $\varepsilon$) can exist independently. Then, we went one step further, and showed that, if a Bragg-like condition is met for two of these modes, then only those two modes should be considered.

Now assume that we excite one of the propagating modes of such a pair in the medium. According to the theory we just went through, it suffices to just keep track of that mode and of the one coupled to it, in order to have a complete picture of the field in the medium. Notice that we did not make any assumptions about the relative magnitudes of the Fourier components of the dielectric coefficient. Our only assumption is that this coefficient is periodic in the $xy$-plane. All this gives us a very strong support to make the assumptions we will make while deriving the coupled mode equations for the $90^o$ geometry, in Section 5.2.2.

### 4.2.4 Photonic band gap

Given a frequency $\omega$, the assumption that $k_z = 0$, and that the electric field is polarized along the $z$ direction, we can use the results of the previous section to calculate $|\vec{k}|$ as a function of $\omega$, also known as the dispersion relation.

Equation 4-21 and Equation 4-22 form a homogeneous system. In order for that system to have nontrivial solutions, its determinant needs to be zero. So, by adopting the notation $\varepsilon_{00} = \varepsilon_c$ and $\varepsilon_{m_0 n_0} = \Delta\varepsilon$, we get

$$\begin{vmatrix} k^2 - \omega^2\mu\varepsilon_c & -\omega^2\mu\Delta\varepsilon \\ -\omega^2\mu(\Delta\varepsilon)^* & |\vec{k} - \vec{g}|^2 - \omega^2\mu\varepsilon_c \end{vmatrix} = 0, \qquad \textbf{(EQ 4-23)}$$

or

$$(k^2 - \omega^2\mu\varepsilon_c)(|\vec{k} - \vec{g}|^2 - \omega^2\mu\varepsilon_c) - (\omega^2\mu|\Delta\varepsilon|)^2 = 0. \qquad \textbf{(EQ 4-24)}$$

This last equation is an implicit dispersion relation. The simplest way to proceed would be to solve for $\vec{k}$ as a function of $\omega$ and identify potential stop-bands, where $\vec{k}$ becomes complex, also known as band gaps [4-5]. Unfortunately, solving Equation 4-24 for $\vec{k}$ is not straightforward. So in an attempt to get something out of it let us rewrite it as:

$$\omega^4\mu^2(\varepsilon_c^2 - |\Delta\varepsilon|^2) - \omega^2\mu\varepsilon_c(k^2 + |\vec{k} - \vec{g}|^2) + k^2|\vec{k} - \vec{g}|^2 = 0. \qquad \textbf{(EQ 4-25)}$$

The corresponding discriminant is

$$\Delta = \mu^2\varepsilon_c^2[k^2 + |\vec{k} - \vec{g}|^2]^2 - 4\mu^2(\varepsilon_c^2 - |\Delta\varepsilon|^2)k^2|\vec{k} - \vec{g}|^2, \qquad \textbf{(EQ 4-26)}$$

or, after some trivial simplification:

$$\Delta = \mu^2\varepsilon_c^2[k^2 - |\vec{k} - \vec{g}|^2]^2 + 4\mu^2|\Delta\varepsilon|^2k^2|\vec{k} - \vec{g}|^2. \qquad \textbf{(EQ 4-27)}$$

It should be obvious that $\Delta \geq 0$. So for every $\vec{k}$ we have two solutions for $\omega$, given by:

$$\omega_{\pm}^2 = \frac{\mu\varepsilon_c[k^2 + |\vec{k} - \vec{g}|^2] \pm \sqrt{\mu^2\varepsilon_c^2[k^2 - |\vec{k} - \vec{g}|^2]^2 + 4\mu^2|\Delta\varepsilon|^2k^2|\vec{k} - \vec{g}|^2}}{2\mu^2(\varepsilon_c^2 - |\Delta\varepsilon|^2)}. \qquad \textbf{(EQ 4-28)}$$

By inspection of Equation 4-28 we can see that $\omega_{\pm} \geq 0$. We can easily calculate the difference between the two frequencies for any given $\hat{k}$. Again, from Equation 4-28 we get

$$\omega_{+}^2 - \omega_{-}^2 = \frac{\sqrt{\varepsilon_c^2[k^2 - |\hat{k} - \hat{g}|^2]^2 + 4|\Delta\varepsilon|^2 k^2 |\hat{k} - \hat{g}|^2}}{\mu^2(\varepsilon_c^2 - |\Delta\varepsilon|^2)}. \quad \textbf{(EQ 4-29)}$$

When is the difference between the two frequencies minimized? To figure that out we need to take the derivative of $\varepsilon_c^2[k^2 - |\hat{k} - \hat{g}|^2]^2 + 4|\Delta\varepsilon|^2 k^2 |\hat{k} - \hat{g}|^2$ with respect to $\hat{k}$ and equate it to $0$. Calculating this derivative requires some pretty straightforward algebra and the final result is

$$2\varepsilon_c^2 g^2(2\hat{k} - \hat{g}) + 4|\Delta\varepsilon|^2 \hat{k}(\hat{k} - \hat{g})(2\hat{k} - \hat{g}) = 0 \quad \textbf{(EQ 4-30)}$$

The solution to the last equation is $\hat{k} = \frac{1}{2}\hat{g}$, the Bragg condition in reflection geometry. What are the frequencies that correspond to this $\hat{k}$? The answer can be obtained by plugging the last equality in Equation 4-28, but it's simpler to go back to Equation 4-24, plug in and solve. In any case the final solution is

$$\omega_{\pm} = \frac{g}{2\sqrt{\mu(\varepsilon_c \mp |\Delta\varepsilon|)}}. \quad \textbf{(EQ 4-31)}$$

These frequencies have the smallest distance between them, and correspond to the Bragg condition. What happens when we are at a frequency between $\omega_{+}$ and $\omega_{-}$? To answer this question let's see what happens close to the Bragg match condition, i.e.,

$$\hat{k} = \frac{1}{2}\hat{g} + \hat{x}, \quad \textbf{(EQ 4-32)}$$

where $|\hat{x}| \ll |\hat{g}|$. In that case Equation 4-24 becomes

$$\left(\left|\frac{1}{2}\hat{g}+\hat{x}\right|^2-\omega^2\mu\varepsilon_c\right)\left(\left|\frac{1}{2}\hat{g}+\hat{x}-\hat{g}\right|^2-\omega^2\mu\varepsilon_c\right)-(\omega^2\mu|\Delta\varepsilon|)^2 = 0 \Rightarrow$$

$$\left(\frac{g^2}{4}+x^2+gx\cos\theta-\omega^2\mu\varepsilon_c\right)\left(\frac{g^2}{4}+x^2-gx\cos\theta-\omega^2\mu\varepsilon_c\right)-(\omega^2\mu|\Delta\varepsilon|)^2 = 0 \Rightarrow$$

$$\left(\frac{g^2}{4}+x^2-\omega^2\mu\varepsilon_c\right)^2-(gx\cos\theta)^2-(\omega^2\mu|\Delta\varepsilon|)^2 = 0 \Rightarrow$$

$$\frac{g^4}{16}+x^4+(\omega^2\mu\varepsilon_c)^2+\frac{g^2x^2}{2}-\frac{g^2\omega^2\mu\varepsilon_c}{2}-2x^2\omega^2\mu\varepsilon_c-(gx\cos\theta)^2-(\omega^2\mu|\Delta\varepsilon|)^2 = 0,$$

where $\theta$ is the angle between $\hat{g}$ and $\hat{x}$. The terms $x^4$, $\frac{g^2x^2}{2}$, and $(gx\cos\theta)^2$ vanish compared to $g^4$, so solving for $x^2$ we get

$$x^2 = \frac{\left(\frac{g^2}{4}-\omega^2\mu\varepsilon_c\right)^2-(\omega^2\mu|\Delta\varepsilon|)^2}{2\omega^2\mu\varepsilon_c}. \qquad \textbf{(EQ 4-33)}$$

So now the question is, when is $\hat{k}$ real? For $\hat{k}$ to be real we must have $x^2>0$ which we can easily see corresponds to

$$\omega^2 < \frac{g^2}{4\mu(\varepsilon_c+|\Delta\varepsilon|)}, \qquad \textbf{(EQ 4-34)}$$

or

$$\omega^2 > \frac{g^2}{4\mu(\varepsilon_c-|\Delta\varepsilon|)}, \qquad \textbf{(EQ 4-35)}$$

which are the same results obtained in Equation 4-31. So these two frequencies define the limits of the stop band, and any field with frequency $\omega$ such that $\omega_- < \omega < \omega_+$ will decay exponentially.

### 4.2.5 Band gap width

From Equation 4-31 we can calculate the single-side bandwidth of the photonic band gap, defined as $\Delta\omega = \omega_+ - \omega_0$, where $\omega_0 = \frac{g}{2\sqrt{\mu\varepsilon_c}}$. We have

$$\Delta\omega = \omega_+ - \omega_0 =$$

$$\frac{2\pi}{\lambda\sqrt{\mu(\varepsilon_c - \Delta\varepsilon)}} - \frac{2\pi}{\lambda\sqrt{\mu\varepsilon_c}} =$$

$$\frac{2\pi}{\lambda\sqrt{\mu\varepsilon_c}}\left(\frac{1}{\sqrt{1-\frac{\Delta\varepsilon}{\varepsilon_c}}} - 1\right) \cong$$

$$\frac{2\pi}{\lambda\sqrt{\mu\varepsilon_c}}\left(1 + \frac{\Delta\varepsilon}{2\varepsilon_c} - 1\right) \Rightarrow$$

$$\Delta\omega = \omega_0 \frac{\Delta\varepsilon}{2\varepsilon_c}. \qquad \textbf{(EQ 4-36)}$$

We see that the width of the band gap is proportional to the modulation depth $\frac{\Delta\varepsilon}{\varepsilon_c}$.

# 4.3 Reflection geometry using coupled-mode analysis

## 4.3.1 Introduction

In this section we start by presenting the Coupled-Mode formalism, first introduced by Kogelnick [4-6]. We focus on mode coupling in free-space, as opposed to waveguides, and derive the Coupled-Mode equations. Then we discuss in detail how the coupling constant is affected by the anisotropy that characterizes the most commonly used photorefractive crystals, such as $LiNbO_3$. We proceed to an analytical solution of the Coupled-Mode equations in reflection geometry. At this point, we start viewing the strong reflection gratings as filters, and investigate how each parameter of the grating affects the characteristics of the resulting filter. We then compare the results with those derived using the Bloch formalism in Section 4.2.5. Finally, we investigate the effects of angular and thermal detuning on filters imple-

mented using strong volume gratings in reflection geometry, and propose a method for athermal design.

### 4.3.2 Coupled mode equations in free space

Assume that we have a medium, whose dielectric coefficient has a sinusoidal modulation:

$$\varepsilon(\vec{r}) = \varepsilon_c + \Delta\varepsilon(e^{-j\vec{K}\vec{r}} + e^{+j\vec{K}\vec{r}}), \quad \textbf{(EQ 4-37)}$$

where $\varepsilon_c$ is the dielectric constant of the crystal and $\Delta\varepsilon$ is the amplitude of the dielectric coefficient modulation[1]. Assume also that the medium is a photorefractive crystal, therefore both $\varepsilon_c$ and $\Delta\varepsilon$ are tensors. An additional assumption is that dielectric constant modulation is weak, i.e., $\frac{\Delta\varepsilon}{\varepsilon_c} \ll 1$.

Since the perturbation of the dielectric coefficient is small, the crystal modes remain quasi-plain waves:

$$\vec{E}(\vec{k}, \vec{r}) = \hat{e}e^{-j\vec{k}\vec{r}}, \quad \textbf{(EQ 4-38)}$$

where $\hat{e}$ is a unit vector along the polarization direction of $\vec{E}(\vec{r})$. Based on Equation 4-38, we have the orthogonality condition:

$$\int \vec{E}(\vec{k}_1, \vec{r})\vec{E}^*(\vec{k}_2, \vec{r})d\vec{r} = \delta(\vec{k}_1 - \vec{k}_2) \quad . \quad \textbf{(EQ 4-39)}$$

The total field in the crystal can be described as a superposition of different modes:

$$\vec{E}(\vec{r}) = \int A(\vec{k})\hat{e}e^{-j\vec{k}\vec{r}}d\vec{k}. \quad \textbf{(EQ 4-40)}$$

Starting from the wave equation, we can write in the perturbed medium:

$$\nabla \times (\nabla \times \vec{E}) - \omega^2\mu\varepsilon\vec{E} = 0 \Rightarrow$$
$$\nabla(\nabla \bullet \vec{E}) - \nabla^2\vec{E} - \omega^2\mu\varepsilon\vec{E} = 0 \Rightarrow$$
$$\nabla^2\vec{E} + \omega^2\mu\varepsilon\vec{E} = 0,$$

where we assumed that $\nabla \bullet \vec{E} = 0$. From $\nabla \bullet \vec{D} = 0$, we get

1. This notation is strictly equivalent to the one introduced in Section 4.2.4 on page 60.

$$\nabla \bullet \vec{E} = -\frac{\vec{E} \bullet \nabla \varepsilon}{\varepsilon}. \qquad \textbf{(EQ 4-41)}$$

The right-hand side of Equation 4-41 becomes zero when $\vec{E}$ is perpendicular to $\vec{K}$, the grating vector. This is exactly true when the wave vector $\vec{k}$ of the field is parallel to $\vec{K}$, and approximately true when the two vectors form a small angle. Therefore, in the case of the reflection geometry, $\nabla \bullet \vec{E} = 0$ is a valid approximation.

In the unperturbed crystal the wave equation becomes

$$\nabla^2 \vec{E} + \omega^2 \mu \varepsilon_c \vec{E} = 0. \qquad \textbf{(EQ 4-42)}$$

Plugging Equation 4-40 in the wave equation we get

$$\nabla^2 \int A(\vec{k}) \hat{e} e^{-j\vec{k}\vec{r}} d\vec{k} + \omega^2 \mu \varepsilon_c \int A(\vec{k}) \hat{e} e^{-j\vec{k}\vec{r}} d\vec{k} = 0 \Rightarrow$$

$$\int A(\vec{k})(\omega^2 \mu \varepsilon_c - |\vec{k}|^2) \hat{e} e^{-j\vec{k}\vec{r}} d\vec{k} = 0. \qquad \textbf{(EQ 4-43)}$$

The last equation is the FT (Fourier transform) of the quantity $A(\vec{k})\hat{e}(\omega^2 \mu \varepsilon_c - |\vec{k}|^2)$, and since it is equal to zero the quantity itself must be zero:

$$A(\vec{k})(\omega^2 \mu \varepsilon_c - |\vec{k}|^2)\hat{e} = 0. \qquad \textbf{(EQ 4-44)}$$

We assume the boundaries of the medium to be perpendicular to the $z$ axis, as a result the quasi plane waves will get coupled and exchange energy along that axis. Therefore $A = A(\vec{k}, z)$, and Equation 4-40 in the perturbed medium will become

$$\vec{E}(\vec{r}) = \int A(\vec{k}, z) \hat{e} e^{-j\vec{k}\vec{r}} d\vec{k}. \qquad \textbf{(EQ 4-45)}$$

Plugging Equation 4-45 in Equation 4-42 we get

$$\nabla^2 \int A(\vec{k}, z) \hat{e} e^{-j\vec{k}\vec{r}} d\vec{k} + \omega^2 \mu \varepsilon \int A(\vec{k}, z) \hat{e} e^{-j\vec{k}\vec{r}} d\vec{k} = 0 \Rightarrow$$

$$\int [\hat{e} \nabla^2 (A(\vec{k}, z) e^{-j\vec{k}\vec{r}}) + \omega^2 \mu \varepsilon \hat{e} A(\vec{k}, z) e^{-j\vec{k}\vec{r}}] d\vec{k} = 0. \qquad \textbf{(EQ 4-46)}$$

We have $\nabla^2 (A(\vec{k}, z) e^{-j\vec{k}\vec{r}}) = \left( \frac{\partial^2}{\partial z^2} A(\vec{k}, z) - 2j\vec{k}\hat{z} \frac{\partial}{\partial z} A(\vec{k}, z) + A(\vec{k}, z) |\vec{k}|^2 \right) e^{-j\vec{k}\vec{r}}$.

Assuming that the energy exchange between modes is much slower than the oscillation of the field[1] we can use the approximation:

$$\nabla^2(A(\vec{k},z)e^{-j\vec{k}\vec{r}}) \cong \left(-2jk_z\frac{\partial}{\partial z}A(\vec{k},z)+A(\vec{k},z)k^2\right)e^{-j\vec{k}\vec{r}}, \qquad \textbf{(EQ 4-47)}$$

where we have defined $k_z = \vec{k}\hat{z}$ and $k = |\vec{k}|$. Plugging Equation 4-37 and Equation 4-47 into Equation 4-46, we get

$$\int\left\{\left(-2jk_z\frac{\partial}{\partial z}A(\vec{k},z)+A(\vec{k},z)k^2\right)\hat{e}e^{-j\vec{k}\vec{r}}+ \right. \qquad \textbf{(EQ 4-48)}$$

$$\left.\omega^2\mu[\varepsilon_c+\Delta\varepsilon(e^{-j\vec{K}\vec{r}}+e^{+j\vec{K}\vec{r}})]A(\vec{k},z)\hat{e}e^{-j\vec{k}\vec{r}}\right\}d\vec{k} = 0.$$

Since the perturbation $\Delta\varepsilon$ is small, $A(\vec{k},z)$ will satisfy Equation 4-44, and therefore Equation 4-48 is simplified to:

$$\int\left[-2jk_z\frac{\partial}{\partial z}A(\vec{k},z)\hat{e}e^{-j\vec{k}\vec{r}}+\omega^2\mu\Delta\varepsilon\hat{e}(e^{-j\vec{K}\vec{r}}+e^{+j\vec{K}\vec{r}})A(\vec{k},z)e^{-j\vec{k}\vec{r}}\right]d\vec{k} = 0. \qquad \textbf{(EQ 4-49)}$$

We multiply this last equation with $e^{jk_{y,1}y}$ and integrate over $y$[2]. By changing the order of integration, taking into account the orthogonality relation described in Equation 4-39, and the assumption that the medium is infinite along the $y$ direction we get $\int\left[-2jk_z\frac{\partial}{\partial z}A(k_{y,1},k_z,z)\hat{e}e^{-jk_zz}+\omega^2\mu\Delta\varepsilon\hat{e}e^{-jK_zz}A(k_{y,1}-K_y,k_z,z)e^{-jk_zz}+\right.$

$$\left.\omega^2\mu\Delta\varepsilon\hat{e}e^{+jK_zz}A(k_{y,1}+K_y,k_z,z)e^{-jk_zz}\right]dk_z = 0.$$

This last equation allows us to deduce that the amplitude change of any quasi plane wave is a linear combination of the product of the perturbation with the amplitudes of the modes in the crystal:

1. This approximation is also known as the parabolic approximation.
2. By this integration we fix the value of $k_y$ to $k_{y,1}$.

$$2jk_{z,1}\frac{\partial}{\partial z}A(k_{y,1},k_{z,1},z)\hat{e}_1 e^{-jk_{z,1}z} =$$
$$\int \omega^2\mu\Delta\varepsilon\hat{e}[e^{-jK_z z}A(k_{y,1}-K_y,k_z,z)+e^{+jK_z z}A(k_{y,1}+K_y,k_z,z)]e^{-jk_z z}dk_z.$$

This can be rewritten as:

$$\frac{\partial}{\partial z}A(k_1,z) = -j\frac{\omega^2\mu\hat{e}_1{}^{\dagger}\Delta\varepsilon}{2k_{z,1}}(\int A(k_{y,1}-K_y,k_z,z)\hat{e}e^{-j(k_z-k_{z,1}+K_z)z}dk_z +$$
$$\int A(k_{y,1}+K_y,k_z,z)e^{-j(k_z-k_{z,1}-K_z)z}dk_z).$$

where $\hat{e}_1{}^{\dagger}$ is the complex conjugate of $\hat{e}_1$.

Solving for $A(k_1,z)$ we get

$$A(k_1,z) = -j\frac{\omega^2\mu\hat{e}_1{}^{\dagger}\Delta\varepsilon}{2k_{z,1}}[\int(\int A(k_{y,1}-K_y,k_z,z)\hat{e}e^{-j(k_z-k_{z,1}+K_z)z}dz)dk_z +$$
$$\int(\int A(k_{y,1}+K_y,k_z,z)e^{-j(k_z-k_{z,1}-K_z)z}dz)dk_z].$$

The integrals with respect to $z$ will average to $0$ unless $k_z-k_{z,1}+K_z\cong 0$, or $k_z-k_{z,1}-K_z\cong 0$. For the sake of argument let us assume $K_z\cong k_{z,1}-k_z$. Then we have

$$\frac{\partial}{\partial z}A(k_1,z) = -j\frac{\omega^2\mu\hat{e}_1{}^{\dagger}\Delta\varepsilon}{2k_{z,1}}\int A(k_{y,1}-K_y,k_z,z)\hat{e}e^{-j(k_z-k_{z,1}+K_z)z}dk_z. \qquad \textbf{(EQ 4-50)}$$

Going back to the physics of the problem we notice that since we have defined the $y$ component of $A(k_{y,1}-K_y,k_z,z)$ and $k_z\cong k_{z,1}-K_z$, the propagation condition imposed by the wave equation $\omega^2\mu\varepsilon_c-k^2=0$ allows us to define unambiguously $k_z=k_{z,2}$. By defining $k_{y,2}=k_{y,1}-K_y$ we finally have

$$\frac{\partial}{\partial z}A(k_1,z) = -j\frac{\omega^2\mu\hat{e}_1{}^{\dagger}\Delta\varepsilon\hat{e}_2}{2k_{z,1}}A(k_2,z)e^{-j(k_{z,2}-k_{z,1}+K_z)z}. \qquad \textbf{(EQ 4-51)}$$

We can also write symmetrically:

$$\frac{\partial}{\partial z}A(k_2, z) = -j\frac{\omega^2 \mu \hat{e}_2{}^{\dagger} \Delta\varepsilon \hat{e}_1}{2k_{z,2}} A(k_1, z) e^{+j(k_{z,2} - k_{z,1} + K_z)z}. \quad \textbf{(EQ 4-52)}$$

Using the definitions $A_i(z) = A(k_i, z)$,

$$\kappa_{12} = \frac{\omega^2 \mu \hat{e}_1{}^{\dagger} \Delta\varepsilon \hat{e}_2}{2k_{z,1}}, \quad \textbf{(EQ 4-53)}$$

$$\kappa_{21} = \frac{\omega^2 \mu \hat{e}_2{}^{\dagger} \Delta\varepsilon \hat{e}_1}{2k_{z,2}}, \quad \textbf{(EQ 4-54)}$$

$$\Delta k_z = k_{z,2} - k_{z,1} + K_z \quad \textbf{(EQ 4-55)}$$

we can finally write Equation 4-51 and Equation 4-52 as:

$$\begin{aligned} \frac{\mathrm{d}A_1}{\mathrm{d}z} &= -j\kappa_{12} A_2 e^{-j\Delta k_z z} \\ \frac{\mathrm{d}A_2}{\mathrm{d}z} &= -j\kappa_{21} A_1 e^{+j\Delta k_z z} \end{aligned}, \quad \textbf{(EQ 4-56)}$$

which are the Coupled-Mode equations.

## 4.3.3 Mode coupling in anisotropic media

### 4.3.3.1 Introduction

In this section we discuss the effects of anisotropy on mode coupling. We focus on lithium niobate ($LiNbO_3$), which is the most commonly used volume holographic material. Anisotropy is manifested by the fact that the refractive index $n$ perceived by an incident field depends on its polarization, or, equivalently, that the electro-optic coefficients are given by a tensor [4-4, 4-7]:

$$\Delta\left(\frac{1}{n^2}\right) = \begin{bmatrix} r_{11} & 0 & r_{13} \\ -r_{11} & 0 & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{42} & 0 \\ r_{42} & 0 & 0 \\ 0 & -r_{11} & 0 \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}. \quad \textbf{(EQ 4-57)}$$

Since $n = \sqrt{\frac{\varepsilon}{\varepsilon_0}}$, it is straightforward to prove that, for small perturbations, $\Delta\varepsilon = -\frac{\varepsilon^2}{\varepsilon_0}\Delta\left(\frac{1}{n^2}\right)$. We define $q_{ij} = -\frac{\varepsilon_{ij}^2}{\varepsilon_0} r_{ij}$ and we have

$$\Delta\varepsilon = \begin{bmatrix} q_{11} & 0 & q_{13} \\ -q_{11} & 0 & q_{13} \\ 0 & 0 & q_{33} \\ 0 & q_{42} & 0 \\ q_{42} & 0 & 0 \\ 0 & -q_{11} & 0 \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}. \qquad \textbf{(EQ 4-58)}$$

At this point we need to define two reference frames, the incidence reference frame $xyz$ and the crystal reference frame $x'y'z'$. The direction and polarization of the incident and diffracted beams and the grating direction are defined in the incidence frame, while the crystal frame is defined by the crystallographic axes. In the most general case the two reference frames differ by some rotation.

#### 4.3.3.2 Grating parallel to the c-axis

Assume that the grating lies along the $z$-axis of the incidence frame. Also assume that the incidence and crystal frames are the same, therefore the direction of the recorded grating is parallel to the c-axis, and that the incident and diffracted beams propagate at an angle $\theta$ (see Figure 4-2.)

In that case the space charge field is

$$\vec{E}_{sc} = E_{sc}\hat{z}. \qquad \textbf{(EQ 4-59)}$$

Therefore, according to Equation 4-58, we have

$$\Delta\varepsilon = \begin{bmatrix} q_{13} & 0 & 0 \\ 0 & q_{13} & 0 \\ 0 & 0 & q_{33} \end{bmatrix} E_{sc} = \begin{bmatrix} \Delta\varepsilon_{or} & 0 & 0 \\ 0 & \Delta\varepsilon_{or} & 0 \\ 0 & 0 & \Delta\varepsilon_{eo} \end{bmatrix}, \qquad \textbf{(EQ 4-60)}$$

where the subscripts denote ordinary and extraordinary polarization.

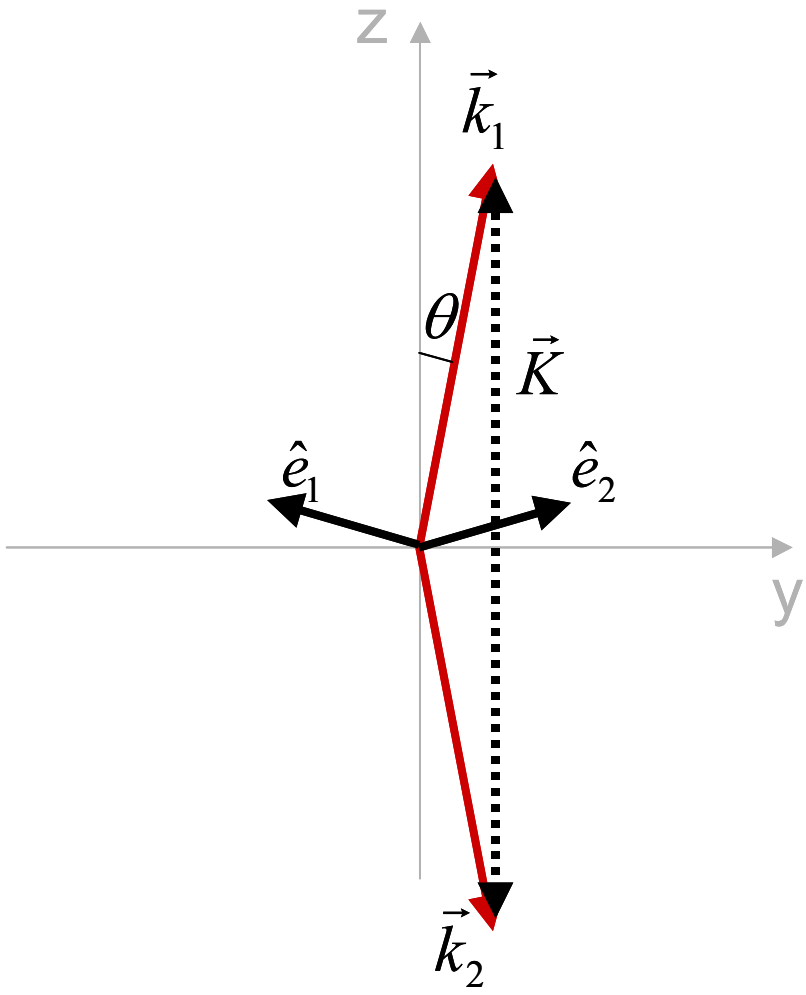



**Figure 4-2.** Simple mode-coupling geometry. The grating vector $\vec{K}$ is parallel to the $z$-axis. Both the incident and diffracted beam wavevectors form an angle $\theta$ with the grating vector. The polarization vectors shown correspond to in-plane, or extraordinary polarization.

Since $\Delta\varepsilon$ is diagonal from Equation 4-53 and Equation 4-54, we deduce that there will be no polarization coupling.

**Ordinary polarization.** In this case we have $\hat{e}_1 = \hat{e}_2 = \hat{x}$, so combining Equation 4-60, Equation 4-53, and Equation 4-54 we get

$$\kappa_{12,o} = \frac{\omega^2 \mu \Delta\varepsilon_{or}}{2k\cos\theta}, \qquad \textbf{(EQ 4-61)}$$

$$\kappa_{21,\,o} = -\frac{\omega^2 \mu \Delta\varepsilon_{or}}{2k\cos\theta}, \qquad \textbf{(EQ 4-62)}$$

The expressions for $\kappa_{12}$ and $\kappa_{21}$ can be simplified further. We have

$$\frac{\omega^2 \mu \Delta\varepsilon_{or}}{2k\cos\theta} = \frac{\omega^2 \mu \varepsilon_0 \dfrac{\Delta\varepsilon_{or}}{\varepsilon_0}}{2k_0 n\cos\theta} = \frac{k_0^2 \dfrac{\Delta\varepsilon_{or}}{\varepsilon_0}}{2k_0 n\cos\theta} = \frac{k_0 \dfrac{\Delta\varepsilon}{2\varepsilon_0 n}}{\cos\theta}.$$

We can also write

$$n = \sqrt{\frac{\varepsilon}{\varepsilon_0}} \Rightarrow$$

$$\Delta n = \frac{1}{2}\frac{1}{\sqrt{\varepsilon\varepsilon_0}}\Delta\varepsilon \Rightarrow$$

$$\frac{\Delta\varepsilon}{2\varepsilon_0 n} = \Delta n.$$

Therefore,

$$\kappa_{12,\,o} = \frac{k_0 \Delta n}{\cos\theta}, \qquad \textbf{(EQ 4-63)}$$

$$\kappa_{21,\,o} = -\frac{k_0 \Delta n}{\cos\theta}. \qquad \textbf{(EQ 4-64)}$$

These last expressions for the coupling constants are the ones most often encountered in literature. Notice the factor $\cos\theta$ in the denominator. This factor is due to the fact we are only interested in the $\hat{k}$ component along the axis of power exchange, which in this case is $z$. Note also that sometimes there is factor of 2 in the denominator. The presence of this factor depends on how the perturbation is defined. In our case it is defined as $\varepsilon(\hat{r}) = \varepsilon_c + \Delta\varepsilon(e^{-jKz} + e^{+jKz})$, so there is no such factor. If we were to define $\varepsilon(\hat{r}) = \varepsilon_c + \Delta\varepsilon\cos Kz$ then the factor of 2 would appear in the denominator.

**Extraordinary polarization.** Again, referring to Figure 4-2, we have $\hat{e}_1 = \begin{bmatrix} 0 \\ -\cos\theta \\ \sin\theta \end{bmatrix}$ and $\hat{e}_2 = \begin{bmatrix} 0 \\ \cos\theta \\ \sin\theta \end{bmatrix}$.

From Equation 4-60 we get $\hat{e}_1^{\dagger}\Delta\varepsilon\hat{e}_2 = \hat{e}_2^{\dagger}\Delta\varepsilon\hat{e}_1 = \Delta\varepsilon_{eo}\sin^2\theta - \Delta\varepsilon_{or}\cos^2\theta$, so we finally have

$$\kappa_{12} = \frac{k_0(\Delta\varepsilon_{eo}\sin^2\theta - \Delta\varepsilon_{or}\cos^2\theta)}{2\varepsilon_0 n_{eo}(\theta)\cos\theta}, \qquad \textbf{(EQ 4-65)}$$

$$\kappa_{21} = -\frac{k_0(\Delta\varepsilon_{eo}\sin^2\theta - \Delta\varepsilon_{or}\cos^2\theta)}{2\varepsilon_0 n_{eo}(\theta)\cos\theta}. \qquad \textbf{(EQ 4-66)}$$

Notice that in this case there can be no further simplification, since

$$n_{eo}(\theta) = \frac{n_{or}n_{eo}}{\sqrt{(n_{eo}\cos\theta)^2 + (n_{or}\sin\theta)^2}}, \qquad \textbf{(EQ 4-67)}$$

still $\kappa_{12} = -\kappa_{21}$.

#### 4.3.3.3 Grating off the c-axis

In this case we assume that there is an angle between the grating and the c-axis, still the c-axis lies within the plane defined by the grating and the propagation directions of the two beams, i.e., the c-axis lies is the $yz$-plane of the incidence frame. Under these constraints we can move from one reference system to the other using two rotations, one around the $x$-axis by $\phi_1$, and the other around the $z$-axis by $\phi_3$. These rotations can be expressed as a matrix:

$$R = \begin{bmatrix} \cos\phi_3 & -\cos\phi_1\sin\phi_3 & \sin\phi_1\sin\phi_3 \\ \sin\phi_3 & \cos\phi_1\cos\phi_3 & -\sin\phi_1\cos\phi_3 \\ 0 & \sin\phi_1 & \cos\phi_1 \end{bmatrix}. \qquad \textbf{(EQ 4-68)}$$

The matrix $R$ takes a vector from the incidence frame $xyz$ to the crystal frame $x'y'z'$, and its transpose $R^T$ does the opposite. It is easy to check that the $\hat{z}'$ vector

in the crystal frame (c-axis) gets mapped to $\begin{bmatrix} 0 & \sin\phi_1 & \cos\phi_1 \end{bmatrix}^T$, thus lies in the $yz$ plane, as required. The electro-optic coefficients in the crystal frame are given by Equation 4-58. The space-charge field lies along the $z$-axis in the incidence frame, therefore the electro-optic coefficients are given by

$$\Delta\varepsilon = \begin{bmatrix} q_{11} & 0 & q_{13} \\ -q_{11} & 0 & q_{13} \\ 0 & 0 & q_{33} \\ 0 & q_{42} & 0 \\ q_{42} & 0 & 0 \\ 0 & -q_{11} & 0 \end{bmatrix} \begin{bmatrix} \cos\phi_3 & -\cos\phi_1\sin\phi_3 & \sin\phi_1\sin\phi_3 \\ \sin\phi_3 & \cos\phi_1\cos\phi_3 & -\sin\phi_1\cos\phi_3 \\ 0 & \sin\phi_1 & \cos\phi_1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ E_{sc} \end{bmatrix},$$

or, in matrix form:

$$\Delta\varepsilon = E_{sc} \begin{bmatrix} q_{13}\cos\phi_1 + q_{11}\sin\phi_1\sin\phi_3 & q_{11}\sin\phi_1\cos\phi_3 & q_{42}\sin\phi_1\sin\phi_3 \\ q_{11}\sin\phi_1\cos\phi_3 & q_{13}\cos\phi_1 - q_{11}\sin\phi_1\sin\phi_3 & -q_{42}\sin\phi_1\cos\phi_3 \\ q_{42}\sin\phi_1\sin\phi_3 & -q_{42}\sin\phi_1\cos\phi_3 & q_{33}\cos\phi_1 \end{bmatrix}.$$

**Ordinary polarization.** In this case we have $\hat{e}_1 = \hat{e}_2 = \hat{x}$, so by Equation 4-53 on page 68 we get

$\kappa_{12} = \frac{\omega^2\mu}{2k_{z,1}} \left( R \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right)^\dagger \Delta\varepsilon R \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, which, using some straightforward algebra, boils down to:

$$\kappa_{12} = \frac{\omega^2\mu}{2k_{z,1}} E_{sc}[q_{13}\cos\phi_1 + q_{11}(1 + 2\cos(2\phi_3))\sin\phi_1\sin\phi_3]. \qquad \textbf{(EQ 4-69)}$$

Note that for $\phi_3 = \frac{\pi}{2}$ and $\phi_1 = \operatorname{atan}\frac{q_{13}}{q_{11}}$ we have $\kappa_{12} = 0$.

Also note that, since both $k_{z,1}$ and $k_{z,2}$ see the same refractive index, we will have $\kappa_{12} = -\kappa_{21}$.

**Extraordinary polarization.** In this case we have $\kappa_{12} = \frac{\omega^2 \mu}{2k_{z,1}} \left( R \begin{bmatrix} 0 \\ -\cos\theta \\ \sin\theta \end{bmatrix} \right)^{\dagger} \Delta\varepsilon R \begin{bmatrix} 0 \\ \cos\theta \\ \sin\theta \end{bmatrix}$. Again, cranking through the algebra we get

$$\kappa_{12} = \frac{\omega^2 \mu}{2k_{z,1}} \frac{E_{sc}}{8} (q_{11}[\cos(2\theta + \phi_1 - 3\phi_3) + \cos(2\theta - \phi_1 + 3\phi_3) + \qquad \textbf{(EQ 4-70)}$$

$$\cos(\phi_1 + 3\phi_3) - \cos(\phi_1 - 3\phi_3) - \cos(2\theta - \phi_1 - 3\phi_3)$$

$$- \cos(3(\phi_1 - \phi_3)) + \cos(3(\phi_1 - \phi_3)) - \cos(3(\phi_1 + \phi_3))$$

$$- \cos(2\theta + \phi_1 + 3\phi_3)] - 2(q_{13} + q_{33})\cos(2\theta - \phi_1)$$

$$- 2(q_{13} - q_{33} - q_{42})\cos\phi_1 - 2q_{13}\cos 3\phi_1 + 2q_{33}\cos 3\phi_1 - 4q_{42}\cos 3\phi_1$$

$$- 2q_{13}\cos(2\theta + \phi_1)).$$

In this case, since $k_{z,1}$ and $k_{z,2}$ do not see the same refractive index, we will have $|\kappa_{12}| \neq |\kappa_{21}|$.

**Cross-polarization coupling.** Since the matrix $\Delta\varepsilon$ is not diagonal, we expect to have coupling from ordinary polarization to extraordinary, and vice versa. Indeed, we have $\left( R \begin{bmatrix} 0 \\ -\cos\theta \\ \sin\theta \end{bmatrix} \right)^{\dagger} \Delta\varepsilon R \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = E_{sc} q_{11} \cos(\theta - \phi_1)\cos\phi_3(1 - 2\cos 2\phi_3)\sin\phi_1$, so

$$\kappa_{1e2o} = \frac{\omega^2 \mu}{2k_{z,1e}} E_{sc} q_{11} \cos(\theta - \phi_1)\cos\phi_3(1 - 2\cos 2\phi_3)\sin\phi_1. \qquad \textbf{(EQ 4-71)}$$

Similarly we get $\left(R\begin{bmatrix}0\\ \cos\theta\\ \sin\theta\end{bmatrix}\right)^{\dagger}\Delta\varepsilon R\begin{bmatrix}1\\0\\0\end{bmatrix} = E_{sc}q_{11}\cos(\theta+\phi_1)\cos\phi_3(1-2\cos 2\phi_3)\sin\phi_1$,

so

$$\kappa_{2e1o} = \frac{\omega^2\mu}{2k_{z,2e}}E_{sc}q_{11}\cos(\theta+\phi_1)\cos\phi_3(1-2\cos 2\phi_3)\sin\phi_1. \qquad \textbf{(EQ 4-72)}$$

Note that for $\phi_1 = 0$ cross-polarization coupling drops to zero as expected, since $\vec{K}$ lies across the c-axis. Also note that for $\phi_3 = \pm\frac{\pi}{6}$ there is no cross-polarization coupling, no matter what the value of $\phi_1$ .This is just an algebraic effect. We have the term $R^{\dagger}\Delta\varepsilon R$ in the coupling coefficient, so there exists some $R$ that will diagonalize $\Delta\varepsilon$, thus reducing cross-polarization coupling to zero.

### 4.3.4 Solution of coupled-mode equations

Assume that the grating is along the c-axis, as discussed in Section 4.3.3.2 on page 69. In that case we have $\kappa_{12} = -\kappa_{21}$, so defining $\kappa = \kappa_{12}$, Equation 4-56 on page 68 becomes

$$\begin{aligned}\frac{dA_1}{dz} &= -j\kappa A_2 e^{-j\Delta k_z z} \\ \frac{dA_2}{dz} &= j\kappa A_1 e^{+j\Delta k_z z}.\end{aligned} \qquad \textbf{(EQ 4-73)}$$

The general solution of this system of differential equations, with boundary conditions $A_1(0)$ and $A_2(L)$, is [4-4]

$$A_1(z) = e^{j\frac{\Delta k_z}{2}z}\left[\frac{s\cosh[s(L-z)]+j\frac{\Delta k_z}{2}\sinh[s(L-z)]}{s\cosh(sL)+j\frac{\Delta k_z}{2}\sinh(sL)}A_1(0)\right. \qquad \textbf{(EQ 4-74)}$$

$$\left.-\frac{j\kappa e^{j\frac{\Delta k_z}{2}L}\sinh(sz)}{s\cosh(sL)+j\frac{\Delta k_z}{2}\sinh(sL)}A_2(L)\right],$$

$$A_2(z) = e^{-j\frac{\Delta k_z}{2}z}\left[\frac{-j\kappa\sinh[s(L-z)]}{s\cosh(sL)+j\frac{\Delta k_z}{2}\sinh(sL)}A_1(0) + e^{j\frac{\Delta k_z}{2}L}\frac{s\cosh(sz)+j\frac{\Delta k_z}{2}\sinh(sz)}{s\cosh(sL)+j\frac{\Delta k_z}{2}\sinh(sL)}A_2(L)\right], \quad \textbf{(EQ 4-75)}$$

where

$$s^2 = \kappa^2 - \left(\frac{\Delta k_z}{2}\right)^2, \quad \textbf{(EQ 4-76)}$$

and we have assumed there is no absorption in the medium.

In the context of filters, the boundary conditions become $A_1(0) = 1$ and $A_2(L) = 0$. In that case we are interested in the Diffraction Efficiency (DE) of the grating, which is defined as:

$$DE = \left|\frac{A_2(0)}{A_1(0)}\right|^2. \quad \textbf{(EQ 4-77)}$$

Plugging the boundary conditions into Equation 4-74 and Equation 4-75, and substituting in Equation 4-77 we get

$$DE = \left|\frac{\kappa\sinh sL}{s\cosh sL + j\frac{\Delta k_z}{2}\sinh sL}\right|^2. \quad \textbf{(EQ 4-78)}$$

## 4.3.5 Bragg detuning of strong volume gratings in reflection geometry

### 4.3.5.1 Introduction

In this section we investigate the effect of Bragg detuning on the diffraction efficiency $DE$ of strong volume gratings in reflection geometry. After some general remarks regarding the dependence of $DE$ on Bragg detuning, we proceed to examine the spectral response of holographic filter, as well as the combined effects of angular and wavelength detuning.

Diffraction efficiency $DE$ is a function of $\Delta k_z$, the Bragg detuning, defined in Equation 4-55 on page 68. From Equation 4-78 and Equation 4-76 we conclude that when $\Delta k_z = 0$, we get $s = \kappa$, and $DE = \tanh^2 \kappa L$. Therefore, when the Bragg condition is met, given sufficient length, the $DE$ becomes equal to unity. If we increase $\Delta k_z$, we still have this exponential behavior of $DE$ vs. $L$, as long as $s$ remains real. When $\left|\frac{\Delta k_z}{2}\right|$ becomes larger than $\kappa$, $s$ becomes imaginary, and the behavior of $DE$ becomes oscillatory. In other words, when Bragg detuning is small:

$$\left|\frac{\Delta k_z}{2}\right| < \kappa, \qquad \textbf{(EQ 4-79)}$$

power is coupled unidirectionally from the incident beam to the diffracted beam, and, provided the grating is long enough, the incident beam will become extinct and all its power will be reflected through the diffracted beam. When the Bragg detuning exceeds the threshold specified in Equation 4-79, energy is exchanged back and forth between the two beams, and the total exchange quickly averages to zero, leaving the incident beam to go through the grating practically unperturbed.

A more intuitive approach to Bragg detuning is to think of the diffracted beam as the superposition of the parts of the incident beam that get reflected at different points of the grating. When the Bragg condition is met, the reflected beams combine constructively, producing a strong diffracted beam. Using energy conservation, we deduce that, in that case, the incident beam is depleted. When detuning is large, the superposition of the reflected beams averages to zero, therefore the incident beam is transmitted through the grating.

The length of the grating $L$ is not of crucial importance, as long as it exceeds a certain minimum, which will allow the energy to be coupled from the incident to the diffracted beam. This is due to the fact that the energy exchange is exponential with length in the stop-band. The effect of $L$ out of the stop-band is limited to controlling the period of the oscillations of the side bands.

#### 4.3.5.2 Wavelength detuning

In order to calculate the spectral response of a holographic filter, it suffices to express Bragg detuning $\Delta k_z$ in terms of wavelength detuning $\Delta\lambda$, and use Equation 4-78. From the k-sphere diagram shown in Figure 4-3, assuming

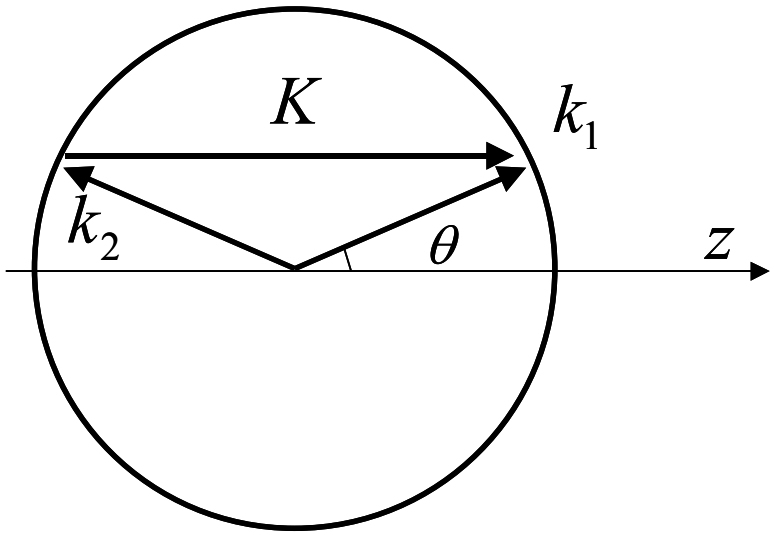



**Figure 4-3.** The k-sphere diagram for the reflection geometry.

$k = \frac{2\pi}{\lambda} n_c$, $k' = \frac{2\pi}{\lambda + \Delta\lambda} n_c$, and taking into account that $\frac{2\pi}{\Lambda} = 2\frac{2\pi}{\lambda} n_c \cos\theta$, we get:

$$\Delta k_z(\Delta\lambda) = -k'\cos\theta - k'\cos\theta + \frac{2\pi}{\Lambda} = -2\frac{2\pi}{\lambda + \Delta\lambda} n_c \cos\theta + 2\frac{2\pi}{\lambda} n_c \cos\theta \Rightarrow$$

$$\Delta k_z(\Delta\lambda) = 4\pi n_c \cos\theta \left( \frac{1}{\lambda} - \frac{1}{\lambda + \Delta\lambda} \right). \qquad \textbf{(EQ 4-80)}$$

Plugging Equation 4-80 into Equation 4-78 gives an expression of $DE$ vs. $\Delta\lambda$. We have plotted this function in Figure 4-4, assuming $\theta = 0$, $\lambda = 1550$ nm, $\Delta n = 5e-4$, and $L = 2$ cm. For $\Delta\lambda$ close to zero we see that the diffraction efficiency $DE$ is close to unity. The dashed vertical lines indicate the point where $\left| \frac{\Delta k_z}{2} \right| = \kappa$. Outside those lines the behavior of $DE$ becomes oscillatory, and gradually decreases to zero.

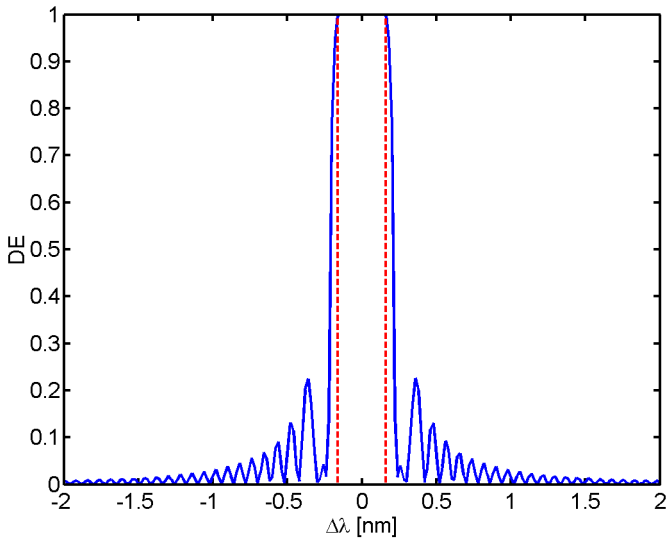



**Figure 4-4.** Diffraction efficiency $DE$ of a strong reflection volume grating vs. wavelength detuning $\Delta\lambda$. The parameters for this plot are $\theta = 0$, $\lambda = 1550$ nm, $\Delta n = 2.5\times10^{-4}$, and $L = 2$ cm. The dashed vertical lines denote the extent of the stop-band.

#### 4.3.5.3 Filter bandwidth vs. photonic band gap width

In this section we compare the bandwidth of the filter, assuming ordinary polarization incidence, with the Photonic Band Gap width derived in Section 4.2.5 on page 62. The edge of the bandwidth is given by the solution of the equation:

$$\left|\frac{\Delta k_z(\omega)}{2}\right| = \kappa. \qquad \textbf{(EQ 4-81)}$$

Taking into account that $k = \frac{\omega}{c}$, where $c$ the speed of light in the medium, and Equation 4-55 on page 68, we get

$$\Delta k_z = \frac{2\Delta\omega}{c}. \qquad \textbf{(EQ 4-82)}$$

From Equation 4-61 on page 70, for $\theta = 0$, we get

$$\kappa = \frac{\omega_0^2 \mu \Delta\varepsilon}{2k}. \qquad \textbf{(EQ 4-83)}$$

Plugging Equation 4-82 and Equation 4-83 into Equation 4-81, and using $c = \frac{1}{\sqrt{\mu\varepsilon_c}}$, we finally get for the single-side bandwidth:

$$\Delta\omega = \omega_0 \frac{\Delta\varepsilon}{2\varepsilon_c}, \qquad \textbf{(EQ 4-84)}$$

which is the same result obtained in Equation 4-36 on page 63.

This is not surprising, since both approaches describe the same effect. Namely, for some frequencies, the partial reflections on each period of the grating interfere constructively, generating a strong diffracted beam, and depleting the incident beam.

#### 4.3.5.4 Angular and wavelength detuning

In Equation 4-80 on page 78 we expressed Bragg detuning $\Delta k_z$ as a function of wavelength detuning $\Delta\lambda$. Assuming that we have both angular and wavelength detuning, we get

$$\Delta k_z(\Delta\theta, \Delta\lambda) = -k'\cos(\theta + \Delta\theta) - k'\cos(\theta + \Delta\theta) + 2\frac{2\pi}{\lambda} n_c \cos\theta =$$
$$-2\frac{2\pi}{\lambda + \Delta\lambda} n_c \cos(\theta + \Delta\theta) + 2\frac{2\pi}{\lambda} n_c \cos\theta \Rightarrow$$

$$\Delta k_z(\Delta\theta, \Delta\lambda) = \frac{4\pi}{\lambda} n_c \cos\theta \left( 1 - \frac{1}{1 + \frac{\Delta\lambda}{\lambda}} \frac{\cos(\theta + \Delta\theta)}{\cos\theta} \right). \qquad \textbf{(EQ 4-85)}$$

We can obviously choose $\Delta\lambda$ and $\Delta\theta$ such that $\Delta k_z(\Delta\theta, \Delta\lambda) = 0$, therefore angular and wavelength detuning can compensate each other.

Another way of approaching this result is by viewing an angle change as a modification of the Bragg condition. As a result of such a change a different wavelength is required to meet this new condition, essentially yielding a filter with a different center frequency.

**Constant grating.** Assume that we record a strong grating in a holographic material. Also assume that when the incidence angle[1] is $\theta_0$ the center wavelength of

the corresponding filter is $\lambda_0$, and that we use ordinary light polarization. Therefore, the spatial frequency of the grating $\Lambda$ will satisfy the equation:

$$2k_0 \cos\theta_0 = \frac{2\pi}{\Lambda}, \qquad \textbf{(EQ 4-86)}$$

where $k_0 = \frac{2\pi}{\lambda_0} n_c$. Assume now that the filter is rotated, so that the new incidence angle is $\theta$. In order to have Bragg match it must be

$$2\frac{2\pi}{\lambda_0} n_c(\lambda_0)\cos\theta_0 = 2\frac{2\pi}{\lambda} n_c(\lambda)\cos\theta. \qquad \textbf{(EQ 4-87)}$$

Since $n_c = n_c(\lambda)$ is known for $LiNbO_3$ [4-8], Equation 4-87 can be solved numerically. The results are plotted in Figure 4-5. It should be evident that by rotating the

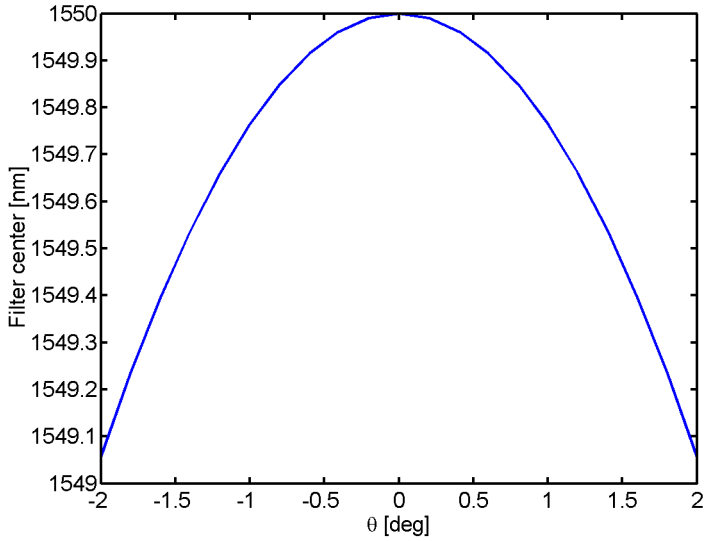



**Figure 4-5.** Dependence of the filter center of a fixed grating on the incidence angle $\theta$. The parameters for this plot are $\theta_0 = 0$ and $\lambda_0 = 1550$ nm.

filter, we are able to tune its center wavelength. Notice that the range of angles $\theta$ is fairly limited. This is to avoid Polarization Mode Dispersion (PMD), which will occur in larger angles, since, in principle, we do not have control of the input polar-

1. All incidence angles are internal to the medium, unless otherwise specified.

ization. Notice also that the angular tuning range is in the order of nanometers, therefore, this method cannot be used as the main tuning mechanism for, say, WDM tunable filters.

Since the angular tuning range is limited, we can assume that $n_c(\lambda_0) \cong n_c(\lambda)$, in which case Equation 4-87 can be simplified to:

$$\lambda = \lambda_0 \frac{\cos\theta}{\cos\theta_0}. \quad \textbf{(EQ 4-88)}$$

Equation 4-87 and Equation 4-88 yield practically the same result, confirming the validity of the approximation.

Next we consider the effect of angular tuning of the filter on the double-side bandwidth, defined as the absolute difference of the two wavelengths/frequencies that satisfy Equation 4-81 on page 79. Since

$$\kappa = \frac{\pi \Delta n}{\lambda \cos\theta}, \quad \textbf{(EQ 4-89)}$$

we can write

$$\left| \frac{4\pi}{\lambda_{edge}} n_c(\lambda_{edge}) \cos\theta - \frac{4\pi}{\lambda_0} n_c(\lambda_0) \cos\theta_0 \right| = 2 \frac{\pi \Delta n}{\lambda_{edge} \cos\theta}. \quad \textbf{(EQ 4-90)}$$

Assuming again that $n_c(\lambda_0) \cong n_c(\lambda)$ we get

$$BW = \frac{\lambda_0 \Delta n}{n_c \cos\theta \cos\theta_0}. \quad \textbf{(EQ 4-91)}$$

In Figure 4-6 we plot the numerical solution of Equation 4-90 and the approximate solution given by Equation 4-91. Obviously, in this case, the approximation is no longer valid. In Figure 4-7 we plot the numerical solution of Equation 4-90. We conclude that in the angular range in which we are confined the bandwidth is pretty much stable.

There is a simple explanation of why larger angles $\theta$ lead to larger bandwidths. As mentioned in Section 4.3.5.2 on page 78 the bandwidth depends on the coupling strength $\kappa$. From Equation 4-89 we see that as $\theta$ increases $\kappa$ becomes larger.

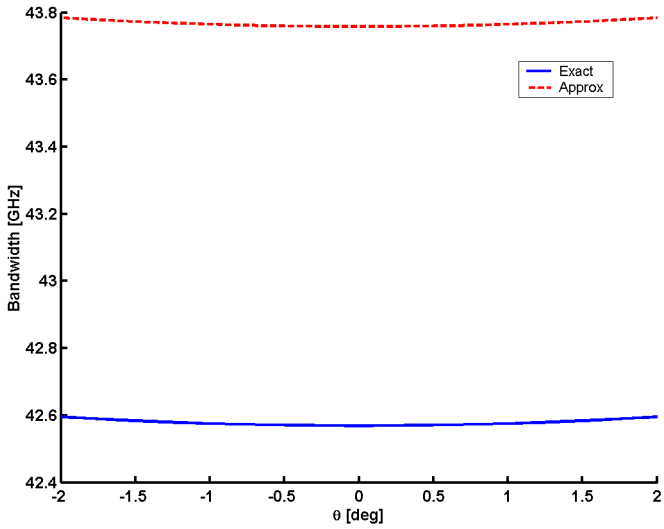



**Figure 4-6.** Dependence of the bandwidth of a fixed grating on the incidence angle $\theta$. The solid line corresponds to the exact solution of Equation 4-90, and the dashed line to the approximate solution, given by Equation 4-91. The results do not agree, therefore the approximate solution is not valid. The parameters for this plot are $\theta_0 = 0$, $\lambda_0 = 1550$ nm, and $\Delta n = 2.5\times10^{-4}$.

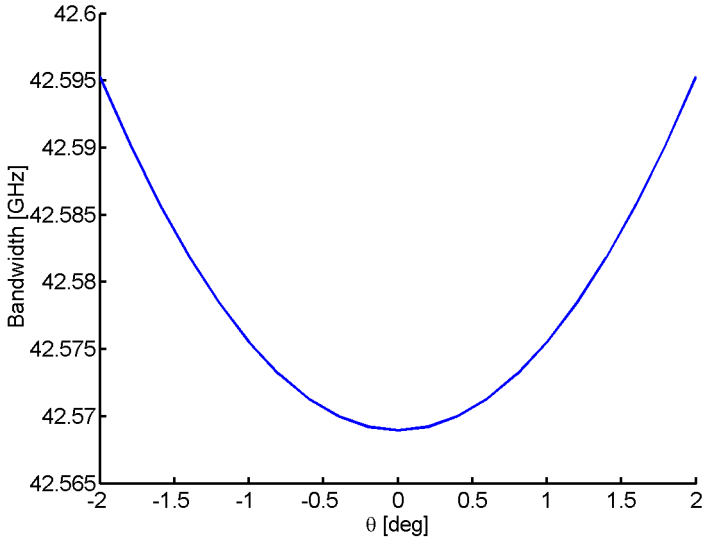



**Figure 4-7.** Dependence of the bandwidth of a fixed grating on the incidence angle $\theta$, based on the exact solution of Equation 4-90. The parameters for this plot are $\theta_0 = 0$, $\lambda_0 = 1550$ nm, and $\Delta n = 2.5\times10^{-4}$.

Again, this is easy to understand intuitively, since, when light is incident at increasing angles, the interaction length per grating period increases.

**Constant filter center.** Let us now start from a different assumption, namely, that we have a specific target center wavelength for our filters, and that for each filter we record a different grating, each of them corresponding to a different incidence angle. In this case we have

$$2k\cos\theta_i = \frac{2\pi}{\Lambda_i}, \qquad \textbf{(EQ 4-92)}$$

where $i$ indexes across different filters. Despite the fact that all filters have the same center they have different bandwidths. For the edge wavelength $\lambda_{edge}$ we have, again from Equation 4-81 on page 79:

$$\left|\frac{4\pi}{\lambda_{edge}} n_c(\lambda_{edge})\cos\theta_i - \frac{4\pi}{\lambda} n_c(\lambda)\cos\theta_i\right| = 2\frac{\pi\Delta n}{\lambda_{edge}\cos\theta_i}, \qquad \textbf{(EQ 4-93)}$$

which we can solve numerically, thus obtaining the bandwidth. In Figure 4-8 we plot the bandwidth $BW$ for both the fixed grating $\Lambda$ and the fixed center $\lambda$ case. Though in the fixed center case the bandwidth is higher, it remains practically constant .

**Consideration of external angle.** Previously we assumed that the internal incidence angle $\theta$ was constant vs. wavelength $\lambda$. In practice, we can only control the external incidence angle $\theta_{out}$. The internal incidence angle (from now on denoted as $\theta_{in}$), and the external incidence angle $\theta_{out}$ are related by Snell's law:

$$n_c(\lambda)\sin\theta_{in} = \sin\theta_{out}. \qquad \textbf{(EQ 4-94)}$$

Assume that we have a specific grating, described by it's spatial period $\Lambda$. We previously figured out the center wavelength $\lambda$ corresponding to every angle $\theta_{in}$. For these angles and wavelengths we can calculate the corresponding external angles $\theta_{out}$ using Equation 4-94.

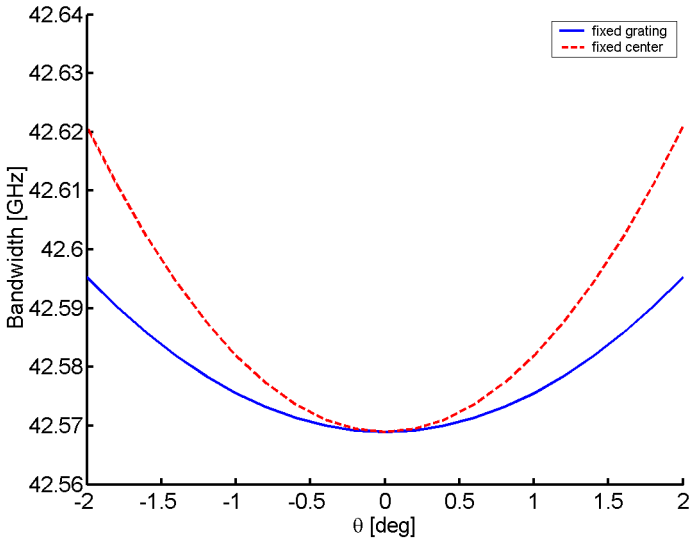



**Figure 4-8.** Dependence of the bandwidth on the incidence angle $\theta$. The solid line corresponds to a fixed grating, and the dashed line to different gratings, each having the same center $\lambda_0$ for the given incidence angle. The bandwidth is smaller for the fixed grating. The parameters for this plot are $\theta_0 = 0$, $\lambda_0 = 1550$ nm, and $\Delta n = 2.5\times10^{-4}$.

Assume now that we set the external angle to $\theta_{out}$. We know the corresponding filter center, and want to calculate the bandwidth. As we scan the wavelength the internal angle $\theta_{in}$ will change, so Equation 4-81 on page 79 becomes

$$\left|\frac{4\pi}{\lambda_{edge}} n_c(\lambda_{edge})\cos\left[\operatorname{asin}\left(\frac{\sin\theta_{out}}{n_c(\lambda_{edge})}\right)\right] - \frac{4\pi}{\lambda} n_c(\lambda)\cos\left[\operatorname{asin}\left(\frac{\sin\theta_{out}}{n_c(\lambda)}\right)\right]\right| = \quad \textbf{(EQ 4-95)}$$

$$\frac{2\pi\Delta n}{\lambda_{edge}\cos\left[\operatorname{asin}\left(\frac{\sin\theta_{out}}{n(\lambda_{edge})}\right)\right]}.$$

Again, we can solve Equation 4-95 numerically, thus obtaining the bandwidth, which we plot in Figure 4-9. It is clear from this plot that keeping $\theta_{out}$ constant results in smaller bandwidth than assuming that $\theta_{in}$ is constant. This is expected, since changing the wavelength will also change $\theta_{in}$ due to the wavelength dependence of the refractive index $n_c$, thus leading to faster Bragg detuning.

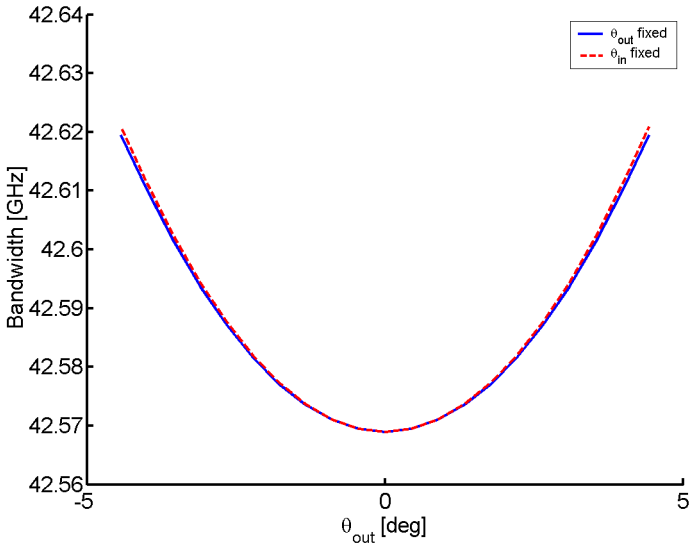



**Figure 4-9.** Dependence of the bandwidth on the incidence angle $\theta$, for different gratings, each having the same center $\lambda_0$ for the given incidence angle. The solid line corresponds to a fixed external angle $\theta_{out}$, and the dashed line to a fixed internal angle $\theta_{in}$. The results for the two cases are similar. The parameters for this plot are $\theta_0 = 0$, $\lambda_0 = 1550$ nm, and $\Delta n = 2.5\times10^{-4}$.

### 4.3.5.5 Experiments

**Experimental methods.** In order to provide experimental support to the theory presented in the previous sections, we recorded strong volume gratings in the reflection geometry, using an iron-doped lithium niobate crystal ($LiNbO_3$:Fe). The crystal has dimensions 3×10×20 $mm^3$, with the c-axis along the longest dimension. The doping level is $1.8\times10^{25}$ atoms/$m^3$ and the concentration ratio $c_{Fe^{2+}}/c_{Fe^{3+}}$ is 0.01.

Since we want to record filters with centers in the vicinity of 1550 nm, and $LiNbO_3$:Fe has very limited absorption in that range of wavelengths, we recorded using a Nd:YAG laser operating at 532 nm, by creating an interference pattern along the 10×20 $mm^2$ face of the crystal[1]. In Figure 4-10 we plot the k-sphere dia-

1. This technique is more commonly referred to as “recording from the top.”

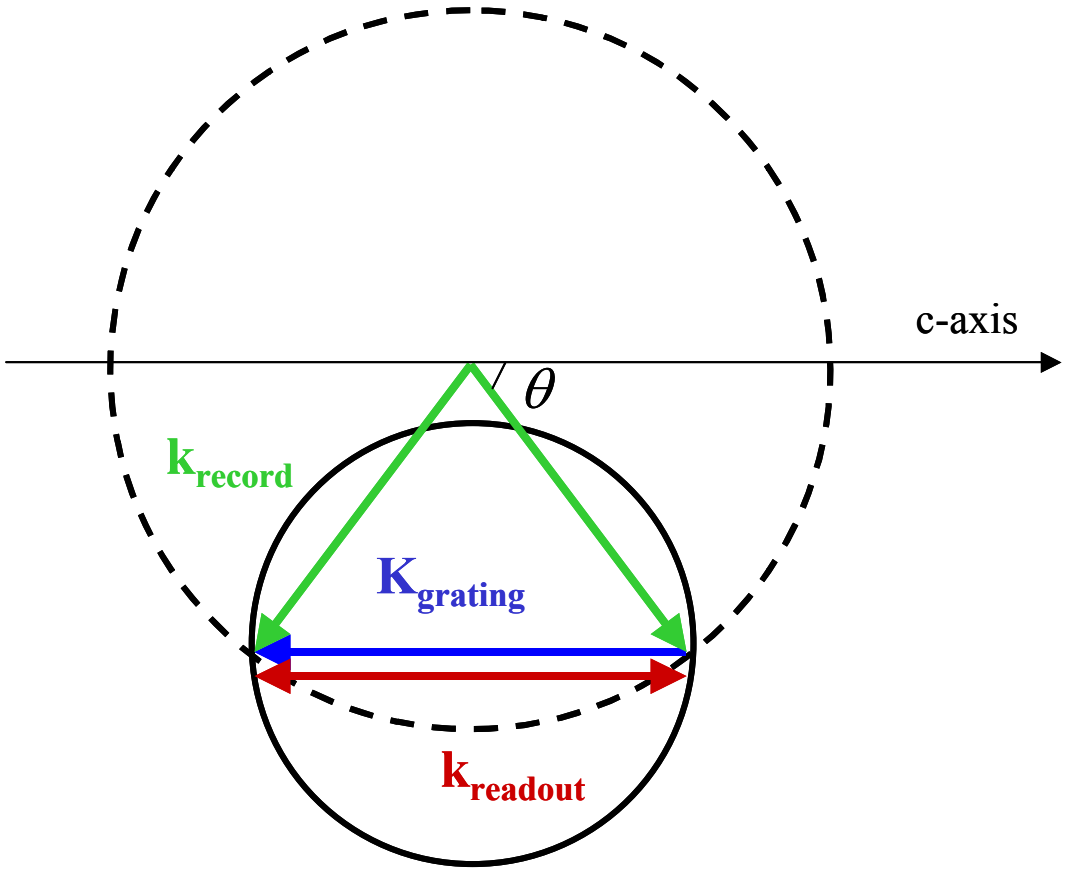



**Figure 4-10.** k-sphere diagram illustrating recording from the top. The dashed sphere corresponds to the shorter recording wavelength. The solid sphere corresponds to the longer readout wavelength.

grams for the recording and readout wavelengths, shown with dashed and solid lines, respectively. From these diagrams we get: $\frac{2\pi}{\lambda_{532}} n_{532} \cos\theta_{in} = \frac{2\pi}{\lambda_{1550}} n_{1550} \Rightarrow$

$$\cos\theta_{in} = \frac{\lambda_{532}}{\lambda_{1550}} \frac{n_{1550}}{n_{532}}. \qquad \textbf{(EQ 4-96)}$$

Using Snell's law[1]

$$n_{532} \cos\theta_{in} = \theta_{out}, \qquad \textbf{(EQ 4-97)}$$

we finally obtain $\theta_{out} = 40.6^{\circ}$, where $\theta_{out}$ is the angle between the face of the crystal and the recording beam.

1. This equation is different from Equation 4-94 on page 84 because of the different definition of angles.

In order to obtain filters with reasonable bandwidth, we need to attain modulation depths of $\Delta n$ in the order of $10^{-4}$. This requires recording times in the order of hours, which in turn necessitates the use of active stabilization. Indeed, for our experiments we used a stabilization scheme first introduced in [4-9], and described in more detail in Section 7.3.3 on page 148.

Once the grating is recorded, we use a different setup to measure the filter properties. The crystal is mounted on a rotation stage, and we monitor the transmitted light coming from a tunable laser.

**Experimental results.** We recorded for one hour using the parameters mentioned in the previous section. A typical recording curve ($DE$ at $\lambda = 532$ nm vs. time) is shown in Figure 4-11.

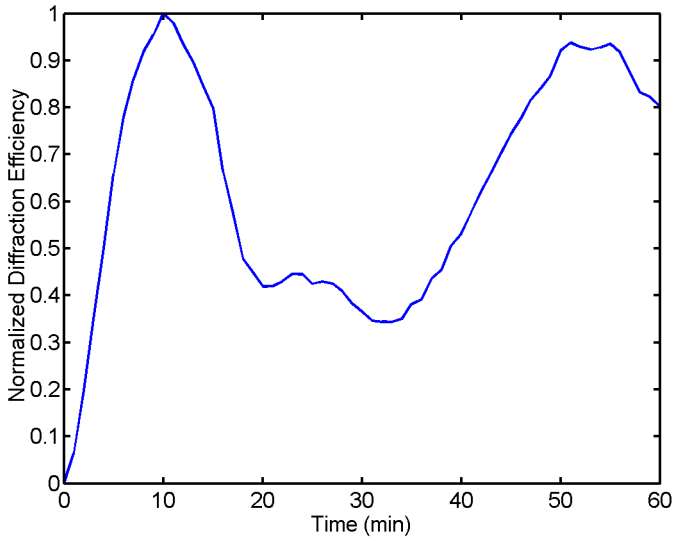



**Figure 4-11.** Recording curve of a strong grating, using an active stabilization system.

Once recording was complete, we measured the filter shapes corresponding to different angles. In Figure 4-12 we plot some of the resulting filters. We can clearly see that rotation results in filters of different centers. Deviation of the filter shape from the one predicted by theory (see Figure 4-4 on page 79) is due to the fact that

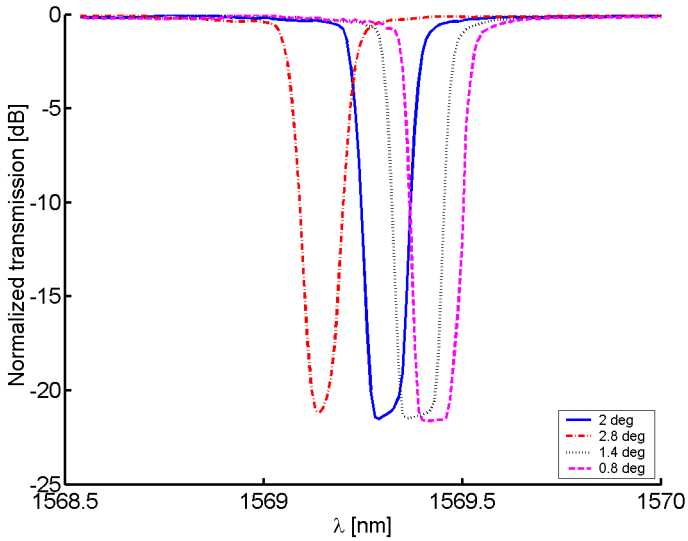



**Figure 4-12.** Filter shapes of strong volume reflection gratings, for different values of incidence angle. Increasing incidence angles result in shorter filter centers.

the recorded grating is not perfect, which in turn can be traced back to the deviation of the recording beams from perfect plane waves.

In Figure 4-13 we plot the filter center vs. incidence angle. We fit these data to:

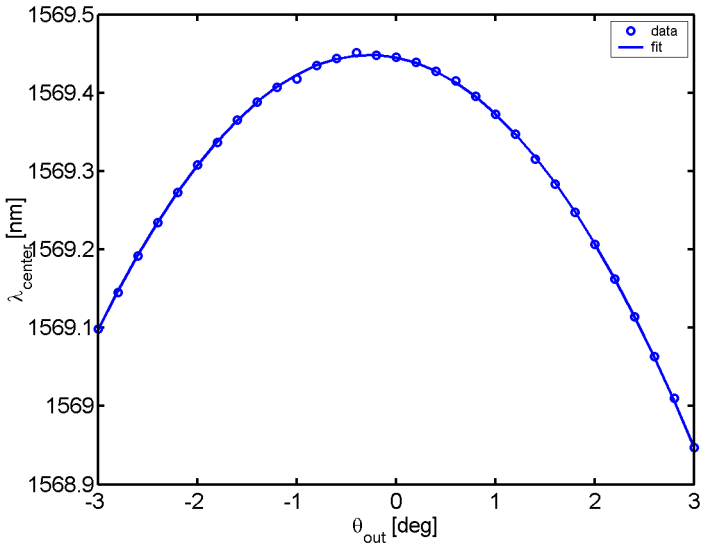



**Figure 4-13.** Filter center $\lambda_{center}$ vs. incidence angle. The dots are experimental results, and the solid line a fit using Equation 4-98.

$$\lambda_{center} = \lambda_{center,0}\cos\left[\frac{\mathrm{asin}(\theta_{out}-\theta_{out,off})}{n_c}\right], \quad \textbf{(EQ 4-98)}$$

which is Equation 4-88 on page 82, where we assume $\theta_0 = 0$, take into account the external angle, and correct for a possible offset of the rotation stage by $\theta_{out,off}$. The experimental data are in excellent agreement with theory, and the resulting parameters are shown in Table 4-1.

| Parameter | Value and 95% CI |
|---|---|
| $\lambda_{center,0}$ | 1569.4000 ± 9×10$^{-4}$ nm |
| $\theta_{out,off}$ | -0.00460 ± 7×10$^{-5}$ rad |
| $n_c$ | 2.260 ± 5×10$^{-3}$ |

**Table 4-1.** Parameter values and 95% Confidence Intervals (CI) resulting from the fitting of the angular filter tuning data to Equation 4-98.

## 4.3.6 Thermal detuning of strong volume gratings in reflection geometry

### 4.3.6.1 Introduction

In Section 4.3.5 we discussed how the characteristics of a strong volume grating filter are affected as we change the parameters of the incident beam, namely angle and wavelength. In this section we investigate the effects of temperature on these filters. We start be presenting the parameters of the filter that are affected by temperature changes, namely spatial period $\Lambda$ and average refractive index $n_c$, and we proceed with a discussion on how these effects can be combined with Bragg detuning to yield an athermal design for a holographic filter.

### 4.3.6.2 Effect of temperature on center wavelength

The Bragg condition for reflection gratings at temperature $T_0$ can be written as:

$$2k\cos\theta_{in,0} = \frac{2\pi}{\Lambda_0}, \quad \textbf{(EQ 4-99)}$$

where the subscript $0$ corresponds to the temperature $T_0$. Using Snell's law, we can rewrite Equation 4-99 in terms of $\theta_{out}$ for a temperature $T$:

$$\lambda_{Bragg} = 2n_c(\lambda_{Bragg}, T)\Lambda(T)\cos\left[\operatorname{asin}\left(\frac{\sin\theta_{out}}{n(\lambda_{Bragg}, T)}\right)\right]. \qquad \textbf{(EQ 4-100)}$$

Since both $\Lambda$ and $n_c$ are functions of temperature the center wavelength of a volume strong grating filter $\lambda_{Bragg}$ will also be a function of temperature.

For the case of $LiNbO_3$, the thermal expansion coefficient and the dependence of the refractive index on temperature are both documented in literature [4-10], therefore Equation 4-100 can be solved numerically. In Figure 4-14 we plot the center

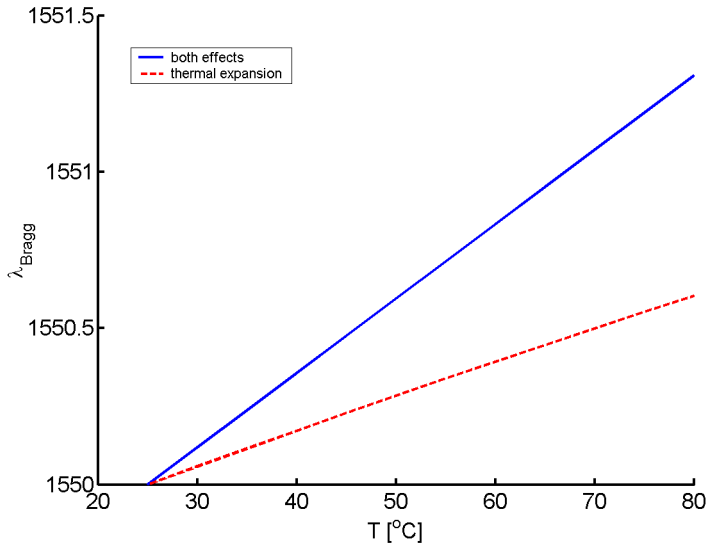



**Figure 4-14.** Effect of temperature on the filter center of strong volume reflection gratings. The solid curve takes into account both thermal dispersion and thermal expansion. The dashed curve only takes into account thermal expansion. We can conclude that the two effects act in the same direction, and that they are comparable.

wavelength vs. temperature taking into account thermal expansion only (dashed curve), or both effects (solid curve), assuming $T_0 = 25$ °C, $\lambda_{Bragg,\,0} = 1550$ nm, and $\theta_{out} = 0$. From this plot we conclude that thermal expansion and temperature dependence of the refractive index have a comparable effect, and act in the same

direction. Note that, since we assumed the grating to lie along the c-axis, there is no need to account for the anisotropic thermal expansion of the crystal.

#### 4.3.6.3 Angular compensation of thermal filter center drift

At a temperature $T = T_0 + \Delta T$, Equation 4-99 becomes

$$2k\cos\theta_{in} = \frac{2\pi}{\Lambda}. \qquad \textbf{(EQ 4-101)}$$

Since in both cases $k$ is the same, by dividing the two equations we get

$$\theta_{in} = \operatorname{acos}\left(\frac{n_{c,0}}{n_c}\frac{\Lambda_0}{\Lambda}\cos\theta_{in,0}\right), \qquad \textbf{(EQ 4-102)}$$

which means that we can compensate the effects of thermal filter center drift by changing the incidence angle. Using Snell's law we can rewrite Equation 4-102 as:

$$\theta_{out}(T) = \operatorname{asin}\left\{n_c \sin\left[\operatorname{acos}\left(\frac{n_{c,0}}{n_c}\frac{\Lambda_0}{\Lambda}\cos\theta_{in,0}\right)\right]\right\}. \qquad \textbf{(EQ 4-103)}$$

This last equation defines which external angle of incidence we should use to mitigate the effects of temperature change on the center of the filter.

#### 4.3.6.4 Experiments

**Experimental methods.** We recorded a strong reflection volume grating similarly to what is described in Section 4.3.5.5 on page 86. The readout setup was fitted with a thermo-electric coupler (TEC) to change the temperature of the crystal, and a thermistor to monitor the temperature. Initially we measured the filter properties at room temperature $T_0$ and $\theta_{out,0} = 0$, and determined the filter center $\lambda_{center,0}$. Then we used the TEC to adjust the temperature of the crystal, and for each temperature measured the angle $\theta_{out}$ that resulted in a filter center as close as possible to $\lambda_{center,0}$.

**Experimental results.** In Figure 4-15 we plot the experimentally obtained angles required for the compensation of thermal detuning vs. temperature. In Figure 4-16 we plot the filter center deviation from $f_{center,0}$ obtained by rotating the filter by the

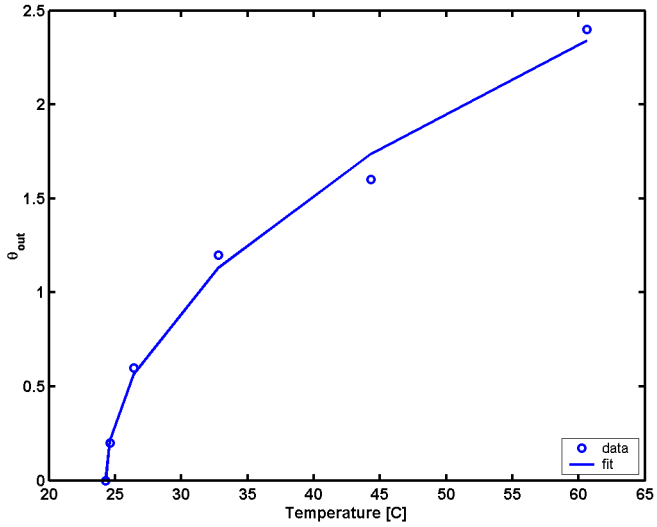



**Figure 4-15.** Angles required for the compensation of thermal detuning vs. temperature. Circles indicate experimental data points. The solid line is the fit curve, obtained using Equation 4-106.

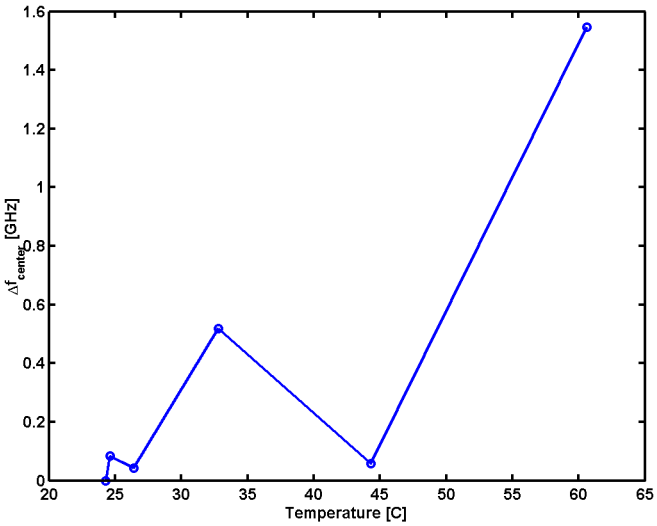



**Figure 4-16.** Deviation of the filter center obtained using angular compensation of thermal detuning, from the target filter center vs. temperature.

thermal compensation angles. The deviation can be minimized further by using a rotation stage with finer step.

In order to fit the thermal compensation data, and since the temperature range is fairly limited, we can assume that both the refractive index $n_c$ and the spatial period of the grating $\Lambda$ vary linearly with temperature:

$$n_c = n_{c,0}(1 + \alpha \Delta T), \quad \textbf{(EQ 4-104)}$$

$$\Lambda = \Lambda_0(1 + \beta \Delta T). \quad \textbf{(EQ 4-105)}$$

Plugging Equation 4-104 and Equation 4-105 into Equation 4-103 on page 92, and taking into account that $\theta_{in,0} = 0$, we get

$$\theta_{out}(T) = \operatorname{asin}\left\{ n_{c,0}(1 + \alpha \Delta T) \sin\left[ \operatorname{acos}\left( \frac{1}{(1 + \alpha \Delta T)(1 + \beta \Delta T)} \right) \right] \right\}. \quad \textbf{(EQ 4-106)}$$

This last equation was used to fit the data, yielding a satisfactory fitting curve, which is plotted in Figure 4-15.

Finally, we characterized the filter shapes of the angularly compensated filters. This was done by measuring the bandwidth at the -3, -5, -10, and -20 dB level, as shown in Figure 4-17 for a particular temperature. The results for all temperature

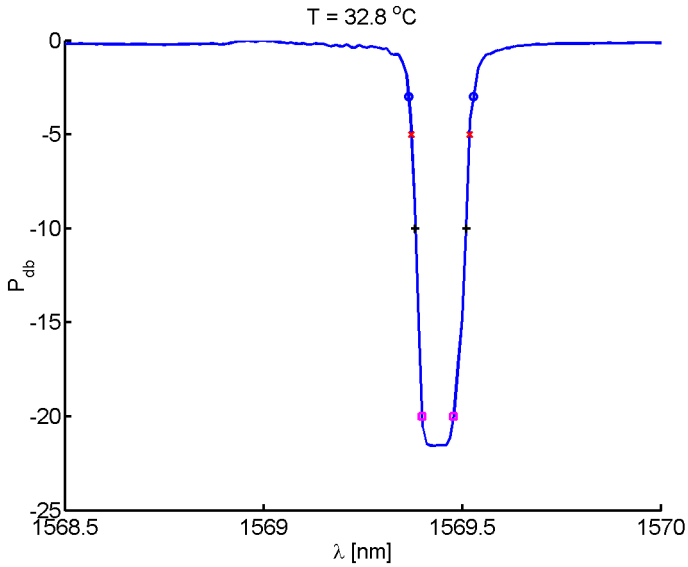



**Figure 4-17.** Filter shape corresponding to a grating at 32.8 °C, after angular compensation of thermal detuning. Pointers indicate the -3, -5, -10, and -20 dB levels.

are plotted in Figure 4-18. All filters have similar characteristics. The tendency of

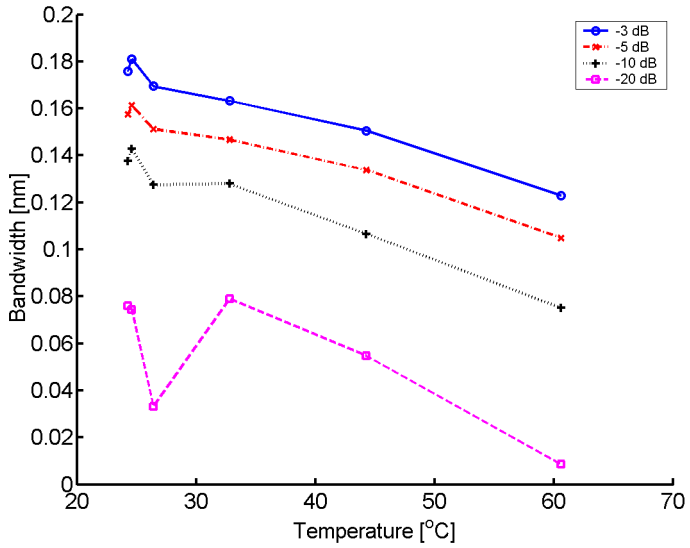



**Figure 4-18.** Experimentally measured bandwidths, corresponding to the -3, -5, -10, and -20 dB levels, for filters whose thermal detuning has been compensated using rotation, vs. temperature. The tendency of the bandwidth to decrease as the temperature increases is due, mainly, to the fact that the grating we used for our experiments was not fixed, and as a result it was getting erased as we increased the temperature.

the bandwidth to decrease as the temperature increases is due, mainly, to the fact that the grating we used for our experiments was not fixed, and as a result it was getting erased as we increased the temperature.

## 4.4 Discussion

We started this chapter with a discussion on the propagation of waves in 2-D media in Section 4.2. We proceeded, in Section 4.3, with the derivation of the coupled-mode equations in the reflection geometry, and demonstrated the equivalence of the two approaches. We also investigated in detail the effect of polarization on mode-coupling in $LiNbO_3$, and concluded that, as long as the grating vector is parallel to the $c$-axis of the crystal, there is no coupling across polarizations.

We then presented a solution of the coupled mode equations, and used it to characterize the performance of strong reflection gratings as filters. We concluded that the two most important parameters are the period of the grating $\Lambda$, which defines the center wavelength of the filter, and the modulation of the refractive index $\Delta n$,

which controls the bandwidth. The length of the grating, once over a certain threshold, only affects the period of the side bands.

The next step was to investigate Bragg detuning. In fiber Bragg gratings, there are two mechanisms of detuning, namely, wavelength and temperature. In the case of volume gratings, since we are in free space, we have the additional degree of freedom of angular detuning. We examined the effect of both thermal and angular detuning on the center wavelength and bandwidth of strong volume reflection gratings. We also obtained experimental results that confirm our theory.

We concluded that neither of the two tuning methods offers a wide enough range to make them suitable for the tuning of filters in the context of WDM fiber communications. Nevertheless, we can use angular tuning to compensate for the effects of thermal detuning, thus obtaining an athermal design. In addition, due to the small sensitivity of the center wavelength of the filter to either method of tuning, we can use each method independently to make fine adjustments, which can be used to mitigate various effects, such as laser drift.

## 4.5 References

[4-1] V. Leyva, G. A. Rakuljic, and B. O'Conner, *Narrow bandwidth volume holographic optical filter operating at the Kr transition at 1547.82 nm.* Applied Physics Letters, 1994. **65**(9): pp. 1079-1081.

[4-2] J. Hukriede, D. Kip, and E. Kratzig, *Thermal tuning of a fixed Bragg grating for IR light fabricated in a $LiNbO_3$: Ti channel waveguide.* Applied Physics B-Lasers and Optics, 2000. **70**(1): pp. 73-75.

[4-3] S. Breer and K. Buse, *Wavelength demultiplexing with volume phase holograms in photorefractive lithium niobate.* Applied Physics B-Lasers and Optics, 1998. **66**(3): pp. 339-345.

[4-4] A. Yariv and P. Yeh, *Optical waves in crystals*. 1984, New York: John Wiley & Sons.

[4-5] E. Yablonovitch, *Photonic band-gap structures.* Journal of the Optical Society of America B-Optical Physics, 1993. **10**(2): pp. 283-295.

[4-6] H. Kogelnik, *Coupled wave theory for thick hologram gratings.* Bell System Technical Journal, 1969. **48**(9): pp. 2909-2947.

[4-7] J. F. Nye, *Physical properties of crystals, their representation by tensors and matrices*. 1985, New York: Oxford University Press.

[4-8] U. Schlarb and K. Betzler, *Refractive indices of lithium niobate as a function of temperature, wavelength, and composition: A generalized fit.* Physical Review B, 1993. **48**(21): pp. 15613-15620.

[4-9] J. Frejlich, L. Cescato, and G. F. Mendes, *Analysis of an active stabilization system for a holographic setup.* Applied Optics, 1988. **27**(10): pp. 1967-1976.

[4-10] R. T. Smith and F. S. Welsh, *Temperature dependence of the elastic, piezoelectric, and dielectric constants of lithium tantalate and lithium niobate.* Journal of Applied Physics, 1971. **42**(6): p. 2219.

# 5 Strong volume gratings in the 90$^{o}$ geometry

## 5.1 Introduction

In this chapter, we extend our investigation of the use of strong volume gratings as a means to control information-bearing light beams, to include the 90$^{o}$ geometry. This geometry takes advantage of the fact that we are dealing with beams propagating in free space in a more direct way than the reflection geometry. In free space, unlike waveguides, we have infinite directions of propagation, and a continuum of modes for each direction, as opposed to waveguides, where the boundary conditions impose mode discretization, and limit the possible directions of propagation. The use of the 90$^{o}$ geometry allows us to implement systems with two axes of propagation, which are perpendicular to each other.

The diffraction of strong gratings in the 90$^o$ geometry can be treated using a number of different approaches. Among them are the scalar field diffraction theory [5-1, 5-2, 5-3], where the Born approximation should *not* be used. Formulating the problem as a system of differential equations with boundary conditions and using purely numerical methods, such as FDTD, to solve it, is also an option [5-4, 5-5]. A third approach is to formulate and solve the coupled-mode equations in the 90$^o$ geometry. The last approach is the one we are going to undertake in this chapter, mainly because of its computational efficiency, in some particular cases, which turn out to be the most interesting ones.

We will start the remainder of this chapter by deriving the coupled-mode equations for the 90$^o$ geometry in Section 5.2, and presenting a solution for the Bragg-matched case. In Section 5.3 we will discuss the diffraction efficiency of 90$^o$ geometry gratings. In Section 5.4 we will present a numerical solution of the coupled-mode equations for the non Bragg-matched case, and will use it to investigate wavelength detuning. We will proceed in Section 5.5 with a discussion of the impulse response. Since the use of the 90$^o$ geometry has been partially investigated in the case of imaging systems [5-5], in Section 5.6 we will focus on issues pertaining to its use as a filter for WDM optical telecommunications.

# 5.2 Coupled mode equations in the 90$^o$ geometry

## 5.2.1 Introduction

The coupled-mode equations, first introduced by H. Kogelnik [5-7], can account for any angle between the incident and diffracted beams. As a result, it would seem that they could be used to investigate the 90$^o$ geometry. Nevertheless, more careful consideration reveals that, inherent in that derivation, is the fact that the boundary conditions are two parallel planes, or, equivalently, that the exchange of energy between the incident and diffracted beams takes place along a single axis.

In the 90$^o$ geometry, the boundary conditions are specified on a rectangle, and the orientation of the grating results in power exchange along two axes, which are per-

pendicular to each other. In order to account for that, we need to assume field amplitudes that will be function of two coordinates, say $A_i(x, y)$, as opposed to the solutions in the reflection geometry, which only depend on one variable (see Section 4.3.2 on page 64).

The derivation of the coupled-mode equations for the 90° geometry has been attempted in literature [5-8, 5-9]. The derivation that follows is similar to the one in [5-9], though it avoids some of the unfounded generalizations presented in that paper, namely the use of nonperpendicular grating boundaries and nonplane waves.

## 5.2.2 Derivation of coupled mode equations

Assume a cubic medium, of finite dimensions along the $x$ and $y$ axes, and infinite along the $z$ axis (see Figure 5-1). Also assume that in this medium we record a strong grating. The grating is in the $xy$ plane, the light field is along z. The grating vector $\vec{G}$ is given by:

$$\vec{G} = k_0(-\hat{x} + \hat{y}), \qquad \textbf{(EQ 5-1)}$$

where $k_0 = \frac{2\pi}{\lambda_0} n_c$, and $n_c$ is the refractive index of the unperturbed medium. The resulting modulation of the electro-optic coefficient can be expressed as:

$$\varepsilon(x, y) = \varepsilon_c + \Delta\varepsilon(e^{-j\vec{G}\vec{r}} + e^{+j\vec{G}\vec{r}}), \qquad \textbf{(EQ 5-2)}$$

where we assume that $\Delta\varepsilon \ll \varepsilon_c$.

Assume now that we come with an incident field $\vec{E}_1 = A_1\hat{z}e^{-jky}$, where $k \cong k_0$. According to Section 4.2.3 on page 57, we expect a solution in the form:

$$E = A_1(x, y)e^{-jky} + A_2(x, y)e^{-jkx}, \qquad \textbf{(EQ 5-3)}$$

where all fields are polarized along $z$, and thus, we will, from now on, convert to scalar notation.

The solution must satisfy the wave equation:

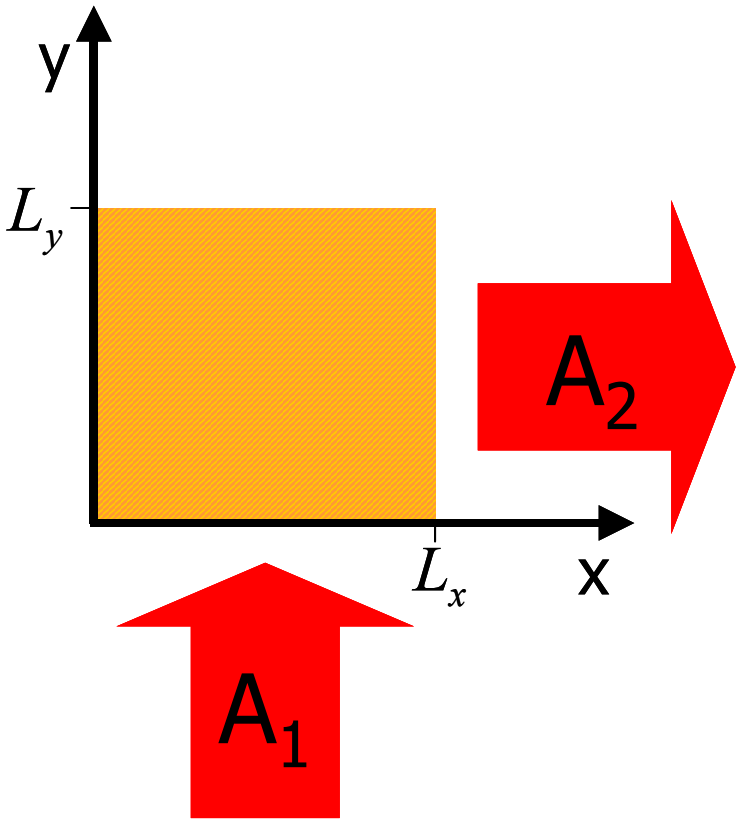



**Figure 5-1.** Grating, incident beam, and diffracted beam for the 90° geometry. $L_x$ and $L_y$ denote the dimensions of the crystal along each axis. $A_1$ and $A_2$ are the complex amplitudes of the incident and diffracted wave respectively.

$$\nabla^2 E + \omega^2 \mu \varepsilon E = 0 . \qquad \textbf{(EQ 5-4)}$$

We have

$$\nabla^2 [A_1(x,y)e^{-jky}] = \left( \frac{\partial^2}{\partial x^2} A_1 + \frac{\partial^2}{\partial y^2} A_1 - 2jk \frac{\partial A_1}{\partial y} - k^2 A_1 \right) e^{-jky} . \qquad \textbf{(EQ 5-5)}$$

Assuming that the energy exchange between modes is much slower than the oscillation of the field, we can use the approximation:

$$\nabla^2 [A_1(x,y)e^{-jky}] \cong \left( -2jk \frac{\partial A_1}{\partial y} - k^2 A_1 \right) e^{-jky} . \qquad \textbf{(EQ 5-6)}$$

Similarly, for $A_2(x,y)$, we have

$$\nabla^2[A_2(x,y)e^{-jkx}] \cong \left(-2jk\frac{\partial A_2}{\partial x} - k^2 A_2\right)e^{-jkx}. \quad \textbf{(EQ 5-7)}$$

Plugging Equation 5-2 and Equation 5-3 into Equation 5-4, and using Equation 5-6 and Equation 5-7, we get

$$-2jk\frac{\partial A_1}{\partial y}e^{-jky} - k^2 A_1 e^{-jky} - 2jk\frac{\partial A_2}{\partial x}e^{-jkx} - k^2 A_2 e^{-jkx} + \quad \textbf{(EQ 5-8)}$$

$$\omega^2\mu\varepsilon_c A_1 e^{-jky} + \omega^2\mu\varepsilon_c A_2 e^{-jkx} +$$

$$\omega^2\mu\Delta\varepsilon A_1(e^{-j\vec{G}\vec{r}} + e^{+j\vec{G}\vec{r}})e^{-jky} + \omega^2\mu\Delta\varepsilon A_2(e^{-j\vec{G}\vec{r}} + e^{+j\vec{G}\vec{r}})e^{-jkx} = 0.$$

Since we assumed that $\frac{\Delta\varepsilon}{\varepsilon_c} \ll 1$, we will have $\omega^2\mu\varepsilon_c A_1 - k^2 A_1 \cong 0$, and $\omega^2\mu\varepsilon_c A_2 - k^2 A_2 \cong 0$, therefore Equation 5-8 can be simplified to:

$$-2jk\frac{\partial A_1}{\partial y}e^{-jky} - 2jk\frac{\partial A_2}{\partial x}e^{-jkx} + \quad \textbf{(EQ 5-9)}$$

$$\omega^2\mu\Delta\varepsilon A_1(e^{-j\vec{G}\vec{r}} + e^{+j\vec{G}\vec{r}})e^{-jky} + \omega^2\mu\Delta\varepsilon A_2(e^{-j\vec{G}\vec{r}} + e^{+j\vec{G}\vec{r}})e^{-jkx} = 0.$$

Assuming that the spatial bandwidth of $\frac{\partial A_1}{\partial y}$ and $\frac{\partial A_2}{\partial x}$ are small compared to $\sqrt{2}k$, we can multiply Equation 5-9 with $e^{jky}$ and integrate over the $xy$ plane, thus obtaining:

$$-2jk\frac{\partial A_1}{\partial y} + \omega^2\mu\Delta\varepsilon A_2 e^{-j\vec{\Delta k}\vec{r}} = 0, \quad \textbf{(EQ 5-10)}$$

where

$$\vec{\Delta k} = \left(\frac{2\pi}{\lambda}n_c - k_0\right)\hat{x} - \left(\frac{2\pi}{\lambda}n_c - k_0\right)\hat{y}. \quad \textbf{(EQ 5-11)}$$

Similarly, multiplication with $e^{jkx}$ and integration over the $xy$ plane leads to

$$-2jk\frac{\partial A_2}{\partial x} + \omega^2\mu\Delta\varepsilon A_1 e^{+j\vec{\Delta k}\vec{r}} = 0. \quad \textbf{(EQ 5-12)}$$

Defining $\kappa = \frac{\omega^2\mu\Delta\varepsilon}{2k}$ we can finally write Equation 5-10 and Equation 5-12 as

$$\frac{\partial A_1}{\partial y} + j\kappa A_2 e^{-j\vec{\Delta k}\vec{r}} = 0, \quad \textbf{(EQ 5-13)}$$

$$\frac{\partial A_2}{\partial x} + j\kappa A_1 e^{+j\vec{\Delta k}\vec{r}} = 0, \quad \textbf{(EQ 5-14)}$$

which are the coupled mode equations in the 90° geometry.

### 5.2.3 Analytical solution of the coupled mode equations in the Bragg match case

In the Bragg match case, $|\vec{\Delta k}| = 0$, it is possible to solve the coupled mode equations analytically. The solution is presented in this section for completeness, and the reader may opt to skip to Equation 5-24 and Equation 5-25 at the end of the section, as the rest of the material in the chapter does not depend on the derivation.

For $\Delta k = 0$, Equation 5-13 and Equation 5-14 become

$$\frac{\partial}{\partial y} A_1(x, y) + j\kappa A_2(x, y) = 0, \quad \textbf{(EQ 5-15)}$$

$$\frac{\partial}{\partial x} A_2(x, y) + j\kappa A_1(x, y) = 0. \quad \textbf{(EQ 5-16)}$$

We take the Laplace Transform (LT) of Equation 5-15 with respect to $y$:

$$s\tilde{A}_1(x, s) - A_1(x, 0) + j\kappa\tilde{A}_2(x, s) = 0. \quad \textbf{(EQ 5-17)}$$

Assuming $A_1(x, 0)$ to be a constant $A_{10}$, we get

$$\tilde{A}_1(x, s) = \frac{A_{10} - j\kappa\tilde{A}_2(x, s)}{s}. \quad \textbf{(EQ 5-18)}$$

We now take the LT of Equation 5-16 with respect to $y$, obtaining $\frac{\partial}{\partial x}\tilde{A}_2(x, s) + j\kappa\tilde{A}_1(x, s) = 0$. Substituting $\tilde{A}_1(x, s)$ from Equation 5-18 we have

$$\frac{\partial}{\partial x}\tilde{A}_2(x, s) + \frac{\kappa^2}{s}\tilde{A}_2(x, s) = -j\frac{\kappa A_{10}}{s}. \quad \textbf{(EQ 5-19)}$$

The homogeneous solution corresponding to Equation 5-19 is $\frac{\partial}{\partial x}\tilde{A}_2(x, s) + \frac{\kappa^2}{s}\tilde{A}_2(x, s) = 0$, and its solution is

$$\tilde{A}_2 = ce^{-\frac{\kappa^2}{s}x}. \quad \textbf{(EQ 5-20)}$$

Assuming $c = c(x, s)$, and plugging Equation 5-20 in Equation 5-19 we get

$$c(x, s) = -j\frac{A_{10}}{\kappa}e^{\frac{\kappa^2}{s}x} + F(s), \quad \textbf{(EQ 5-21)}$$

and by substitution in Equation 5-20:

$$\tilde{A}_2(x, s) = -j\frac{A_{10}}{\kappa} + F(s)e^{-\frac{\kappa^2}{s}x}, \quad \textbf{(EQ 5-22)}$$

where $F(s)$ remains to be determined. In order to do that, we assume that $A_2(0, y) = A_{20}$, where $A_{20}$ is a constant, therefore: $\tilde{A}_2(0, s) = \frac{A_{20}}{s}$. Using Equation 5-22 we get $-j\frac{A_{10}}{\kappa} + F(s) = \frac{A_{20}}{s}$, so by substitution we finally get

$$\tilde{A}_2(x, s) = -j\frac{A_{10}}{\kappa}\left(1 - e^{-\frac{\kappa^2}{s}x}\right) + \frac{A_{20}}{s}e^{-\frac{\kappa^2}{s}x}. \quad \textbf{(EQ 5-23)}$$

Taking the Inverse Laplace Transform of this last equation we obtain

$$A_2(x, y) = A_{20}J_0(2\kappa\sqrt{xy}) - jA_{10}\sqrt{\frac{x}{y}}J_1(2\kappa\sqrt{xy}), \quad \textbf{(EQ 5-24)}$$

and, symmetrically:

$$A_1(x, y) = A_{10}J_0(2\kappa\sqrt{xy}) - jA_{20}\sqrt{\frac{y}{x}}J_1(2\kappa\sqrt{xy}). \quad \textbf{(EQ 5-25)}$$

# 5.3 Diffraction efficiency in the 90° geometry

## 5.3.1 Theory

In practice, we would like to use strong gratings in the 90° geometry as filters, pretty much in the same way we used strong reflection gratings in Chapter 4. In this case, we can assume that the incident field is $A_{10}e^{-jky}$, and that $A_{20} = 0$. The solutions to the coupled mode equations then become

$$A_1(x, y) = A_{10}J_0(2\kappa\sqrt{xy}), \quad \textbf{(EQ 5-26)}$$

$$A_2(x,y) = -jA_{10}\sqrt{\frac{x}{y}}J_1(2\kappa\sqrt{xy}). \qquad \textbf{(EQ 5-27)}$$

In Figure 5-2 we plot $A_1(x,y)$ and $A_2(x,y)$, using $\lambda = 800$ nm, $L_x = L_y = 1$ cm,

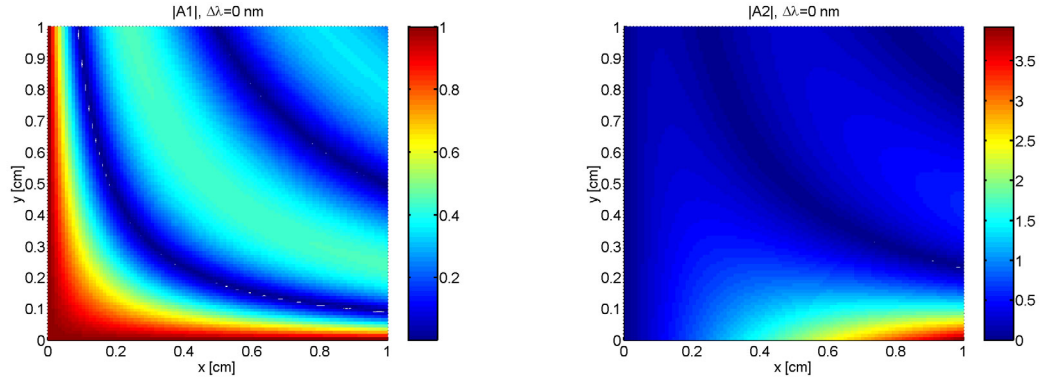



**Figure 5-2.** Color-coded plots of the field amplitude in the Bragg-matched case: (a) Incident beam. (b) Diffracted beam. Note that the color code is different for the two plots.

$n_c = 2.3$, and $\Delta n = 5\times10^{-5}$. Note that the field amplitudes $A_1$ and $A_2$ on the output boundaries ($y = L_y$ and $x = L_x$, respectively) are not constant, therefore the diffracted and transmitted beams are not plane waves. The intensity profile of the diffracted beam at the output boundary is a function of $y$ given by:

$$\left|A_2(L_x,y)\right|^2 = A_{10}^2\frac{L_x}{y}J_1^2(2\kappa\sqrt{L_x y}). \qquad \textbf{(EQ 5-28)}$$

In Figure 5-3 we plot the intensity profile vs. $y$ for various values of $\Delta n$. Note that as the modulation depth of the grating $\Delta n$ increases, more energy is diffracted from the area closer to the input boundary.

Since the diffracted beam is not a plane wave we can only define the Power Diffraction Efficiency $DE_P$, as opposed to the Intensity Diffraction Efficiency, commonly used with weak gratings.

The incident, or input, power will be proportional to $\int_0^{L_x}\left|A_1(x,0)\right|^2 dx = A_{10}^2 L_x$. The diffracted, or output, power will be proportional to:

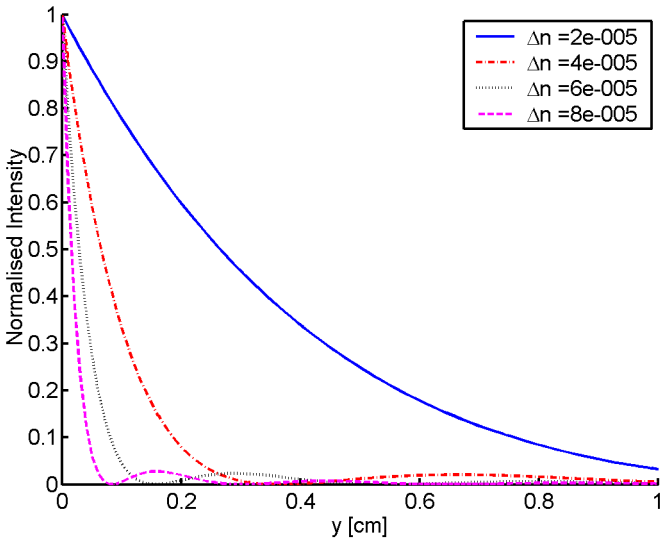



**Figure 5-3.** Intensity profile of the diffracted beam along the output boundary, for various values of refractive index modulation $\Delta n$. As $\Delta n$ increases, the grating becomes stronger, and more energy is diffracted from the area closer to the input boundary.

$$\int_0^{L_y} |A_2(L_x, y)|^2 dy = A_{10}^2 \int_0^{L_y} \frac{L_x}{y} J_1^2(2\kappa\sqrt{L_x y})dy . \qquad \textbf{(EQ 5-29)}$$

By using the change of variables $v = 2\kappa\sqrt{L_x y}$, and referring to integral tables we can calculate that last integral in closed form, obtaining:

$$\int_0^{L_y} |A_2(L_x, y)|^2 dy = A_{10}^2 L_x [1 - J_0^2(2\kappa\sqrt{L_x L_y}) - J_1^2(2\kappa\sqrt{L_x L_y})] . \qquad \textbf{(EQ 5-30)}$$

For $DE_P$ we have $DE_P = \dfrac{\int_0^{L_y} |A_2(L_x, y)|^2 dy}{\int_0^{L_x} |A_1(x, 0)|^2 dx} \Rightarrow$

$$DE_P = 1 - J_0^2(2\kappa\sqrt{L_x L_y}) - J_1^2(2\kappa\sqrt{L_x L_y}) . \qquad \textbf{(EQ 5-31)}$$

The $DE_P$ is an increasing function of $v = 2\kappa\sqrt{L_x L_y}$ (see Figure 5-4.) Indeed, increasing the coupling strength $\kappa$ should increase $DE_P$. Probably the best way to understand intuitively the effect of the term $\sqrt{L_x L_y}$, is to think of it as a measure of the surface over which the incident and diffracted waves interact. As long as we

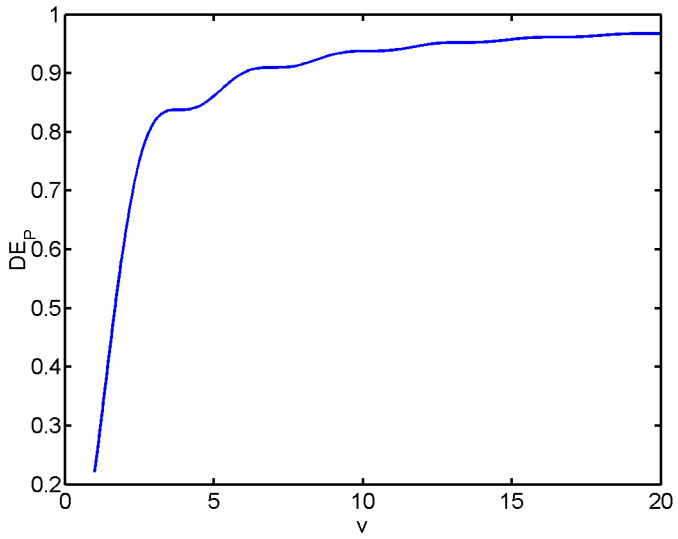



**Figure 5-4.** Diffraction efficiency $DE_P$ of a strong grating in 90$^o$ geometry vs. $v = 2\kappa\sqrt{L_x L_y}$.

are close to Bragg match, energy will get transferred from the incident to the diffracted beam, therefore an increase of the interaction surface should lead to an increase in $DE_P$.

In the previous derivation we defined $DE_P$ assuming that all the light on the output boundary can be used. In practice, assuming, say, a WDM filter that uses a strong grating in the 90$^o$ geometry, we would have to collect all the diffracted light, and couple it into a fiber. Given that the diffracted beam is *not* a plane wave, this is not straightforward, and, therefore, the $DE_P$ calculated here is an upper bound of what can be expected in practice.

### 5.3.2 Experiments

We used a 45$^o$-cut crystal to record strong gratings in the 90$^o$ geometry. The crystal is made of $LiNbO_3$, and is doped with iron and cerium, with $c_{Fe} = 0.05$ %wt and $c_{Ce} = 0.03$ %wt. The dimensions of the crystal are 10 × 10 × 10 mm$^3$. The source is a HeNe laser ($\lambda = 632$ nm.) The schematic of the recording setup is shown in Figure 5-5. The HeNe beam is filtered and collimated, then split in two beams of

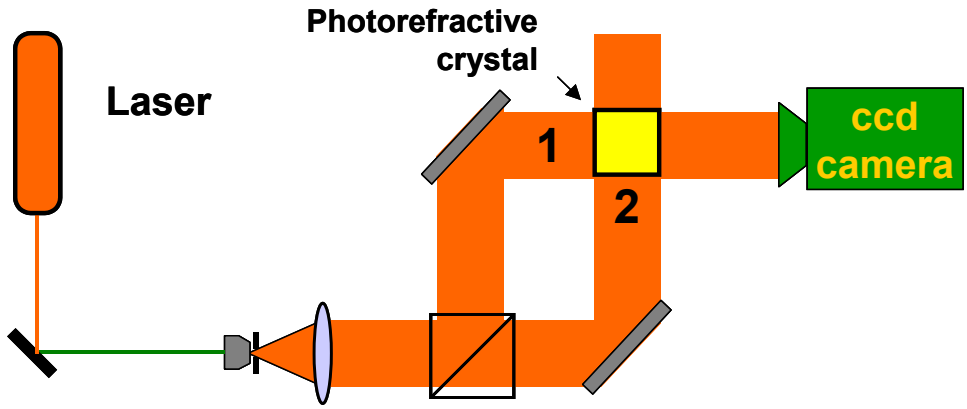



**Figure 5-5.** Schematic diagram of the experimental setup used to monitor the evolution of grating strength vs. recording time, for strong volume gratings in the 90° geometry.

equal intensity, which are interfered within the crystal volume to record the grating. The duration of the exposure is five hours, and every 15 minutes we block beam 1, and use the camera to record the pattern resulting from the diffraction of beam 2. In Figure 5-6 we show four of these frames. Notice that, as time goes by, the grating becomes stronger, and more energy is diffracted in the area closer to the input boundary, as predicted by Equation 5-28. In Figure 5-7 we plot several intensity profile curves corresponding to increasing recording times.

In order to test our theory we fit each intensity profile to the following equation:

$$I_i(y) = \frac{G_i}{y} J_1^2(c_i\sqrt{y - y_0}). \qquad \textbf{(EQ 5-32)}$$

The parameter $G_i$ is used to compensate for the auto-gain feature of the camera, which results in different gain for each frame. The parameter $y_0$ is used because we exclude the edges of the frame, in order to avoid scattering noise. Finally, the parameter $c_i$ will allow us to calculate the grating strength $\kappa_i$ corresponding to frame, according to:

$$\kappa_i = \frac{c_i}{2\sqrt{L_x}}. \qquad \textbf{(EQ 5-33)}$$

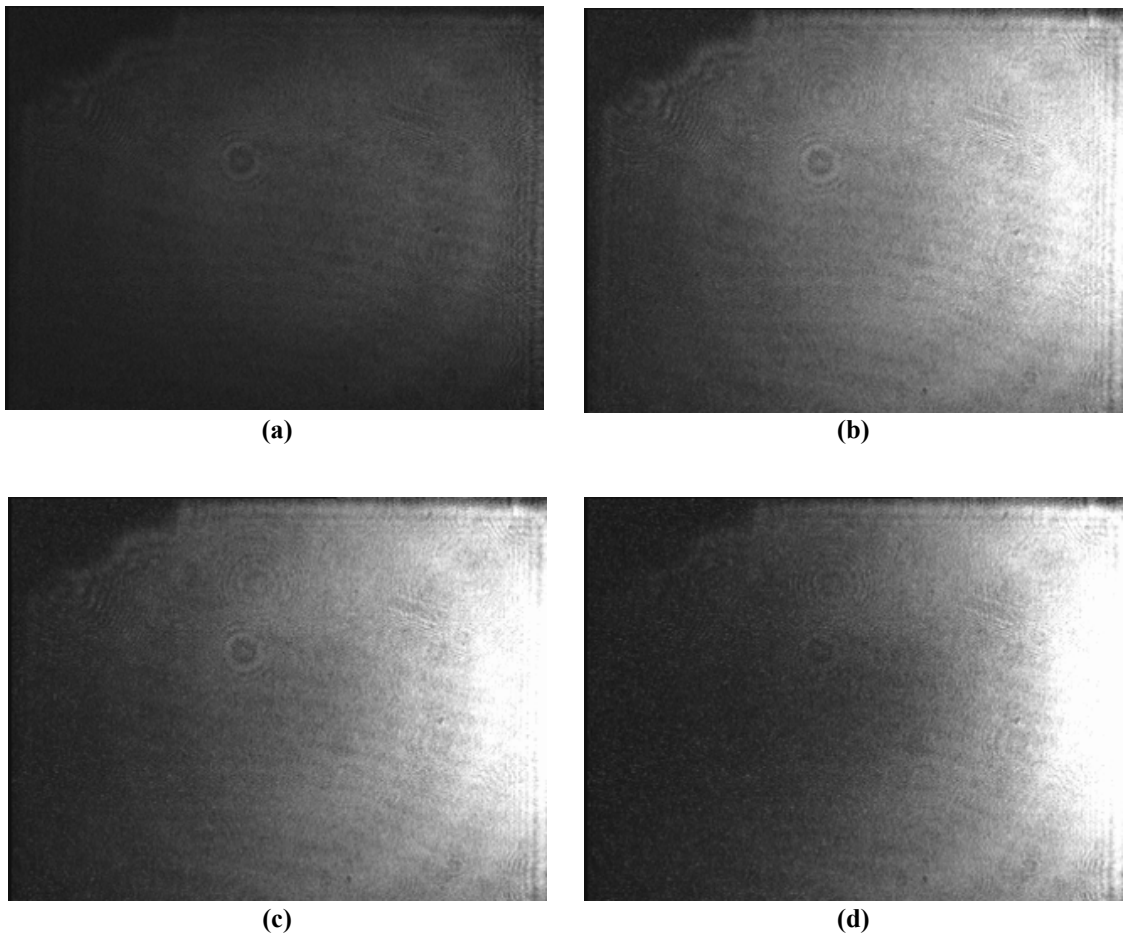



**Figure 5-6.** Diffracted beam profiles collected during the recording of a strong volume grating in the 90$^o$ geometry: (a) t=45 min, (b) t=120 min, (c) t=195 min, (d) t=285 min. As recording time increases the grating becomes stronger, and more energy is diffracted from the area closer to the input boundary.

With a derivation similar to the one in Section 4.3.3.2 on page 69 we can show that $\kappa = \frac{2\pi\Delta n}{\lambda}$, therefore we can calculate

$$\Delta n_i = \frac{\lambda c_i}{4\pi\sqrt{L_x}}. \qquad \textbf{(EQ 5-34)}$$

In Figure 5-8 we plot two data curves along with the corresponding fits, from which we can conclude that the fit is successful. In Figure 5-9 we plot the values of $\Delta n_i$ vs. recording time $t$. We see that the modulation depth of the grating increases with time and moves towards saturation. Due to the large confidence intervals resulting from the data fit (see Figure 5-10) it is not possible to conclude whether the growth

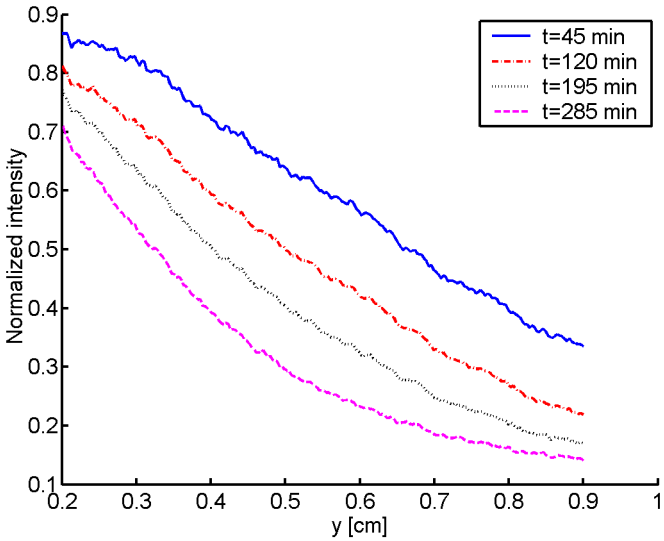



**Figure 5-7.** Intensity profile of the diffracted beam along the output boundary. Different curves correspond to different points in time during recording. As recording time increases the grating becomes stronger, and more energy is diffracted from the area closer to the input boundary. Note that we exclude 1 mm on each side of the crystal to avoid edge effects.

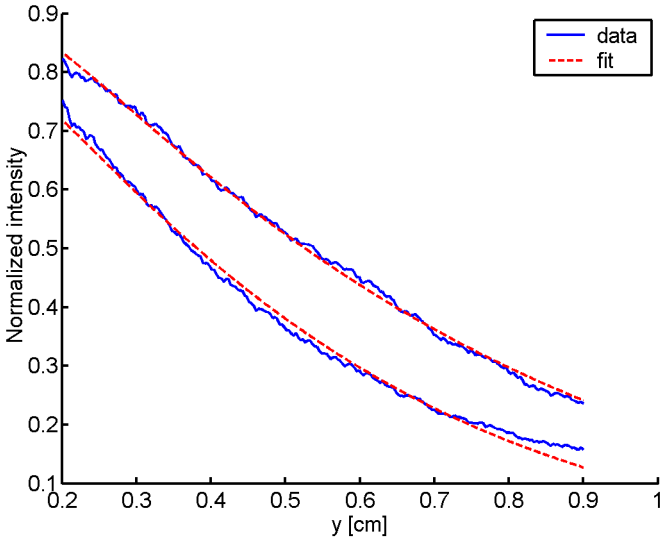



**Figure 5-8.** Fitting results of Equation 5-32 to experimentally measured intensity profiles.

of $\Delta n$ is exponential with time. Note also that Equation 5-32 results from the assumption of a Bragg matched constant grating. Dynamic recording effects, which are present in our experiment, might lead to deviations from this assumption.

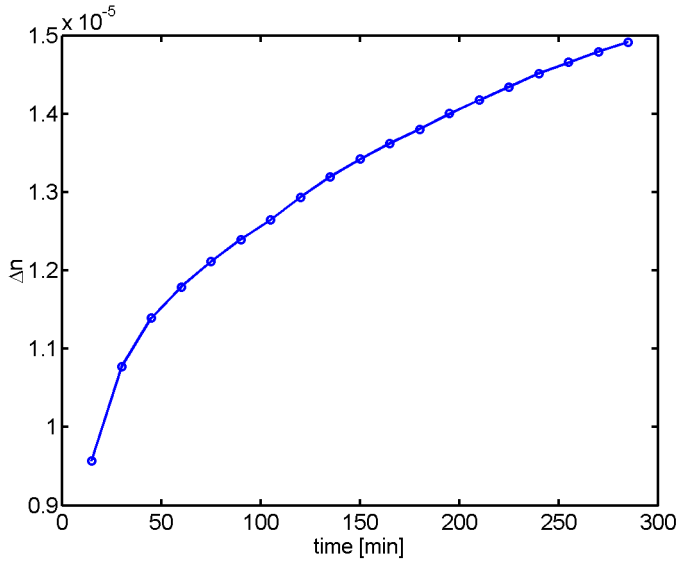



**Figure 5-9.** Refractive index modulation $\Delta n$ (fit parameter to Equation 5-32) vs. recording time. The connecting line is a guide for the eye.

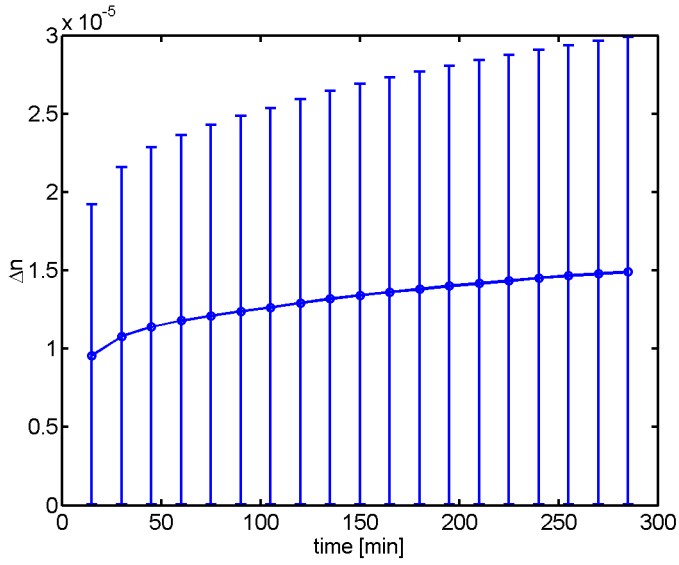



**Figure 5-10.** Refractive index modulation $\Delta n$ (fit parameter to Equation 5-32) vs. recording time. The error bars indicate the 95% confidence intervals. The connecting line is a guide for the eye.

# 5.4 Wavelength detuning of strong gratings in the 90$^o$ geometry

## 5.4.1 Introduction

In this section we discuss the effect of wavelength detuning on the Diffraction Efficiency $DE_P$ of strong gratings in the 90$^o$ geometry. In Section 5.2.2 we derived the coupled mode equations, assuming that both incident fields are perpendicular to the input boundaries. In Section 5.2.3 we solved analytically those equations, in the case of Bragg match, $|\overrightarrow{\Delta k}| = 0$. Unfortunately, no analytical solution exists for the $|\overrightarrow{\Delta k}| \neq 0$ case.

We start this section by presenting an algorithm that allows us to solve numerically the coupled mode equations in the case of $|\overrightarrow{\Delta k}| \neq 0$. We proceed by discussing some of the issues that arise regarding the accuracy of the numerical solution, and derive methods to calculate the required simulation parameters. Finally we present simulation results.

At this point we do not discuss angular detuning. It is possible to re-derive the coupled mode equations to take into account incident fields nonperpendicular to the input boundaries. Nevertheless, the requirement to match the rectangular boundary conditions to the nonrectangular computational grid of our numerical solution increases the computational strain considerably.

## 5.4.2 Numerical solution of the coupled mode equations

### 5.4.2.1 Algorithm

The coupled mode equations can be rewritten in differential form, yielding

$$A_1(x, y+dy) = A_1(x, y) - j\kappa A_2(x, y) e^{-j\overrightarrow{\Delta k}\hat{r}} dy, \quad \textbf{(EQ 5-35)}$$

$$A_2(x+dx, y) = A_2(x, y) - j\kappa A_1(x, y) e^{+j\overrightarrow{\Delta k}\hat{r}} dx. \quad \textbf{(EQ 5-36)}$$

Regarding $\overrightarrow{\Delta k}\hat{r}$, from Equation 5-11 on page 102 we can write

$$\overrightarrow{\Delta k}\hat{r} = 2\pi\frac{n_c\Delta\lambda}{\lambda_0(\lambda_0+\Delta\lambda)}(y-x), \qquad \textbf{(EQ 5-37)}$$

where $\Delta\lambda = \lambda - \lambda_0$. An intuitive way to think about these equations is that, if we know both $A_1$ and $A_2$ for a specific point $(x, y)$, we can "propagate" the solution for $A_1$ along $\hat{y}$, and the solution for $A_2$ along $\hat{x}$.

We know the boundary conditions $A_1(x, 0)$ and $A_2(0, y)$. Since we know both $A_1(0, 0)$ and $A_2(0, 0)$, we can calculate $A_1(0, y)$ and $A_2(x, 0)$, therefore we have the complete solution on the input boundaries. From the points $(x, 0)$ we can calculate $A_1(x, dy)$. From the points $(0, y)$ we can calculate $A_2(dx, y)$. Therefore we have the complete solution for the point $(dx, dy)$.

Now we have the same situation as in the beginning of the algorithm, namely we know $A_1$ along the $x$ axis, $A_2$ along the $y$ axis, and both at the new origin $(dx, dy)$. Therefore, by repeating the same steps, we can calculate $A_1$ and $A_2$ everywhere.

The fact that we can "propagate" the solution through the crystal is the result of the causality of the 90° geometry. Light that propagates along $\hat{x}$ can only be diffracted along $\hat{y}$, and vice versa. A photon that goes through point $(x, y)$ will never return to that point, unlike the reflection geometry. We will see more effects of this causality in Section 5.5, where we will discuss the impulse response of strong gratings in the 90° geometry.

Our algorithm, and for that matter the coupled-mode equations in 90° geometry, is limited to gratings that have the same spatial frequency throughout their extent. Nevertheless, it is perfectly suitable to address gratings whose modulation depth is a function of space. Indeed, in Equation 5-35 and Equation 5-36 we can assume $\kappa = \kappa(x, y)$ without affecting the algorithm. Therefore our algorithm can be used to simulate, say, apodized gratings, but not chirped gratings.

#### 5.4.2.2 Numerical stability considerations

Assume that we are considering detuning in frequency. In that case Equation 5-37 becomes

$$\overrightarrow{\Delta k}\vec{r} = \frac{2\pi}{c} n_c (f - f_0)(x - y), \qquad \textbf{(EQ 5-38)}$$

where $c$ is the speed of light in vacuum.

Assuming that the step of the computational grid is the same in both directions, i.e., $dx = dy$, we have

$$\overrightarrow{\Delta k}\vec{r} = \frac{2\pi}{c} n_c (f - f_0) m dx, \qquad \textbf{(EQ 5-39)}$$

where $m$ an integer. To maintain the accuracy of the simulation we must have $\overrightarrow{\Delta k}\vec{r} \ll 2\pi$, otherwise we will have constructive interference of the incident and diffracted fields, despite finite detuning. Using Equation 5-39, this last condition becomes

$$dx \ll \frac{c}{n_c (f - f_0)}, \qquad \textbf{(EQ 5-40)}$$

therefore the higher the detuning, the more dense the computational grid required.

### 5.4.3 Simulation results

In Figure 5-11 we give color-coded plots of the amplitude of the incident field $A_1$ and the diffracted field $A_2$, vs. $x$ and $y$, for various degrees of detuning[1]. As detuning increases, light beams diffracted from different points of the grating interfere destructively. As a result, for large values of detuning, the incident beam goes through the grating producing minimal diffraction.

In Figure 5-12 we plot the diffraction efficiency $DE_P$ vs. wavelength $\lambda$ for various values of $\Delta n$. Similarly to what was observed in the reflection geometry (see Section 4.3.5 on page 76), we see that increasing the grating strength leads to an

1. The parameters used here are the same as in Section 5.3.1 on page 104.

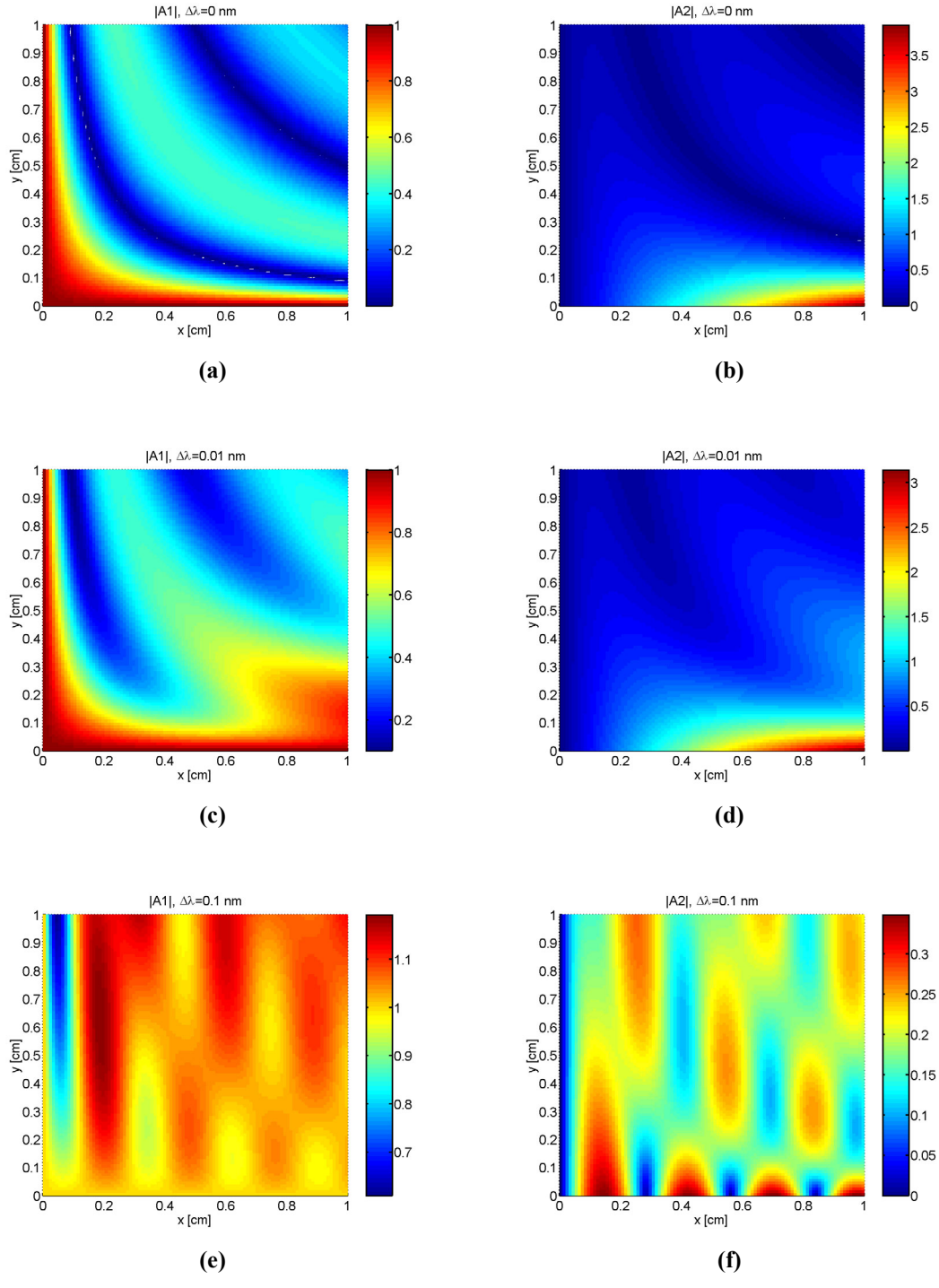



**Figure 5-11.** Color-coded plots of field amplitude for the incident beam $A_1$, and the diffracted beam $A_2$, for various values of detuning $\Delta\lambda$: (a) $A_1$ for $\Delta\lambda = 0$ nm. (b) $A_2$ for $\Delta\lambda = 0$ nm. (c) $A_1$ for $\Delta\lambda = 0.01$ nm. (d) $A_2$ for $\Delta\lambda = 0.01$ nm. (e) $A_1$ for $\Delta\lambda = 0.1$ nm. (f) $A_2$ for $\Delta\lambda = 0.1$ nm. Note that the color code is different for each plot.

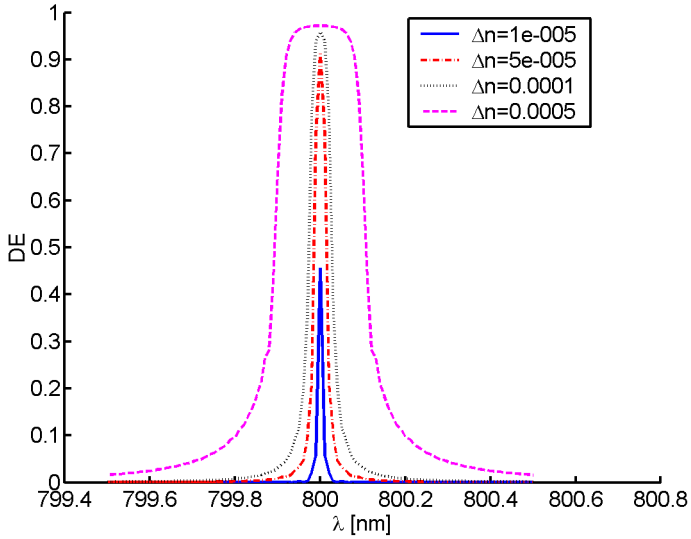



**Figure 5-12.** Diffraction efficiency $DE_P$ vs. wavelength, for various values of refractive index modulation $\Delta n$. Higher values of $\Delta n$ result in larger bandwidth.

increase of the width of the stopband. Again, notice that for reasonable values of $\Delta n$ the resulting bandwidths are in the order of tenths of a nm, and thus allow us to create very fine notch filters.

# 5.5 Impulse response of 90$^o$ volume holograms

## 5.5.1 Introduction

In this section we discuss the impulse response of strong gratings in the 90$^o$ geometry, in order to complete their characterization. Having studied their spectral properties in Section 5.4, we will use their response in the frequency domain to calculate the impulse response. Since there is no analytical solution for the non Bragg-matched case, we will have to approach this problem numerically as well.

We start, in Section 5.5.2, by discussing how we can use the Fast Fourier Transform (FFT) to address problems in continuous time and frequency, and how we can define the minimum sampling requirements. In Section 5.5.3 we combine those requirements with the ones presented in Section 5.4.2.2, regarding the numerical stability of the detuning simulation, as to obtain a complete picture of the computa-

tional requirement for the calculation of the impulse response. Finally, in Section 5.5.4, we present and discuss the results of our simulations.

### 5.5.2 Use of FFT with continuous time and frequency

Since our problem is defined in continuous time and frequency, and we are using a digital computer to simulate it, we need a mapping from continuous time and frequency to discrete time and frequency.

The definition of the Fourier Transform (FT) is [5-10]

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt, \qquad \textbf{(EQ 5-41)}$$

and the definition of the Discreet Fourier Transform (DFT)[1] is

$$X(k) = \sum_{n=1}^{N} x(n)e^{-j2\pi\frac{(k-1)(n-1)}{N}}. \qquad \textbf{(EQ 5-42)}$$

The number of samples $N$ of the DFT is equal to the number of samples of the time series $x(n)$, and is one of the simulation parameters we will have to choose.

According to sampling theory we have

$$z(n) = e^{-j2\pi\frac{n-1}{N}}. \qquad \textbf{(EQ 5-43)}$$

The sampling frequency $f_s$ is mapped to point $-1$ on the unit circle in $z$-space. The value of $z$ corresponding to $\frac{N}{2}+1$ is, according to Equation 5-43, $e^{-j2\pi\frac{\left(\frac{N}{2}+1-1\right)}{N}} = e^{-j\pi} = -1$, so $z\left(\frac{N}{2}+1\right)$ corresponds to $f = f_s$, and $X\left(\frac{N}{2}+1\right)$ to $X(f_s)$.

1. The notation we are using adheres to what would be used in a Matlab implementation, therefore indexes run from 1 to N, as opposed to the more commonly used 0 to N-1 range.

The value of $z$ corresponding to $1$ is $e^{-j2\pi\frac{(1-1)}{N}} = 1$, so $z(1)$ corresponds to $f = 0$. Since the mapping from continuous to discrete frequency is linear, it is easy to extrapolate all intermediate values, and therefore calculate $X(k)$, for $k = 1, \ldots, \frac{N}{2}+1$.

Since we assume our signals to be real, we have for the FT:

$$X(-f) = \int_{-\infty}^{\infty} x(t)e^{j2\pi ft}dt = \left(\int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt\right)^* = X^*(f). \qquad \textbf{(EQ 5-44)}$$

For the DFT we will have equivalently:

$$X(N-k) = \sum_{n=1}^{N} x(n)e^{-j2\pi\frac{(N-k-1)(n-1)}{N}} =$$

$$\sum_{n=1}^{N} x(n)e^{-j2\pi\frac{N(n-1)}{N}}e^{+j2\pi\frac{(k+2-1)(n-1)}{N}} =$$

$$\sum_{n=1}^{N} x(n)e^{+j2\pi\frac{(k+2-1)(n-1)}{N}} \Rightarrow$$

$$X(N-k) = X^*(k+2). \qquad \textbf{(EQ 5-45)}$$

This last equation is valid for $k = 2, \ldots, \frac{N}{2}$, therefore from the values of $X(k)$ for $k = 2, \ldots, \frac{N}{2}$ we can obtain the values for $N, \ldots, \frac{N}{2}+2$ respectively, thus obtaining all the values of $X(k)$ required to calculate $x(k)$ in the discrete time domain.

**Sampling parameters.** Assume we have $N$ samples ($N$ will be the same both in frequency and time). It is recommended to chose $N = 2^m$ for $m$ integer, as most FFT implementations are optimized under this assumption.

Assume that the total duration of the signal is $T_D$. The time between samples is $T_s = \frac{T_D}{N}$. We have by definition $f_s = \frac{1}{T_s} = \frac{N}{T_D}$. The distance between samples in frequency is $\Delta f_s = \frac{f_s}{N} = \frac{1}{N}\frac{N}{T_D} = \frac{1}{T_D}$, so the longer the duration, the finer the sampling in frequency. But we have from the Nyquist theorem:

$$f_s \geq 2f_{max} \Rightarrow$$
$$\frac{N}{T_D} \geq 2f_{max} \Rightarrow$$
$$N \geq 2f_{max}T_D, \qquad \textbf{(EQ 5-46)}$$

so if we want to make the frequency sampling finer or have signals with high frequency components we are forced to use a larger value for $N$.

### 5.5.3 Numerical stability considerations

In this section we combine the results of the previous section with those of Section 5.4.2.2 to derive limits for our simulation parameters. In theory the input pulse is a $\delta$-function, therefore its spectrum extends to infinity. In practice, any short-pulse laser has a limited bandwidth. Let us assume we have a Ti:Saphire pulsed laser, operating at a center wavelength of 800 nm, which produces pulses in the order of femptoseconds, with a bandwidth of about 20 nm, which allows us to calculate $f_{max}$. Also assume that we have a square crystal, of side $L$. Since diffraction in the 90$^o$ geometry is causal, all light will be out of the crystal after time $2\tau$, where $\tau = \frac{Ln_c}{c}$, with $c$ the speed of light in vacuum, and $n_c$ the refractive index of the crystal. So let's assume $T_D = 4\tau$.

Since we have both $f_{max}$ and $T_D$ we can use Equation 5-46 to calculate $N$ and round it upwards to the closest power of two, so we have $N = 2^n$, where $n$ is an

integer. We can now write: $f_{max} = \frac{f_s}{2} = \frac{N}{2T_D} = \frac{N \cdot c}{8 n_c L} = \frac{2^{n-3} c}{n_c L}$. Assume that we use $M$ points per meter in the computational grid of the numerical calculation of the coupled mode equations, the grid step will then be $dx = \frac{1}{M}$. Since $f_{max} = f - f_0$, Equation 5-40 on page 114 can be written as $dx \ll \frac{c}{n_c f_{max}}$, therefore, we finally get $M \gg \frac{2^{n-3}}{L}$, and for our simulations we can choose $M = \frac{2^n}{L}$.

### 5.5.4 Simulation results

Before applying our simulation to the 90$^o$ geometry, we used the analytical results derived in Section 4.3.5.2 on page 78 regarding the reflection geometry, to test our algorithm. The results we obtained in this case were confirmed by results available in literature [5-11].

**Simulation parameters.** We run our simulation using $\lambda_0 = 800$ nm, $n_c = 2.3$, $\Delta n = 5 \times 10^{-5}$, and $L_x = L_y = 1$ cm. In Figure 5-13 we have a color-coded plot of the output field intensity vs. time $t$ and distance from the input face of the crystal $y$. For point $y_0$ the impulse response starts at $t = \frac{y_0}{c} n_c$ and ends at $t' = \frac{y_0 + L_x}{c} n_c$, having a duration of $\tau = \frac{L_x}{c} n_c$, as expected from the causality of the geometry.

In Figure 5-14 we plot the impulse response vs. time for specific points of the output face of the crystal. At $y = 0$ the output is a square pulse, since all the light was diffracted only once, and had to travel different distances. At $y > 0$ we see that the impulse response is limited between $\frac{y}{c} n_c$ and $\frac{y}{c} n_c + \tau$, but is not constant. This is due to the fact that light can reach these points travelling thought different paths in the crystal. The light components that arrive at the output at any given point

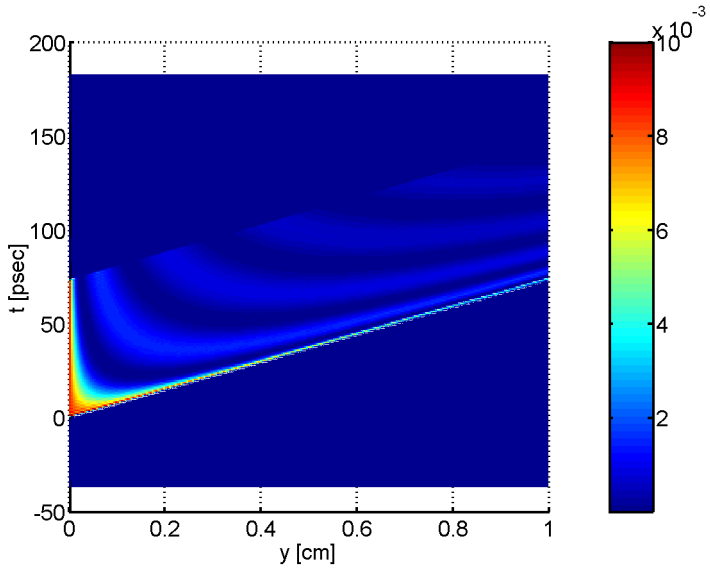



**Figure 5-13.** Color-coded plot of the output field intensity vs. time $t$ and distance from the input face of the crystal $y$.

through different paths have each undergone a different series of diffractions, and therefore have a different phase. The summation of these components is what makes the impulse response vary within the abovementioned time window.

## 5.6 Discussion

In the previous sections we developed theoretical and numerical tools that allow us to investigate the properties of strong volume gratings in the 90$^o$ geometry. In this section we use these results to discuss the use of those strong gratings as filters for WDM optical communications. The main parameters that we are interested in are the diffraction efficiency $DE_P$, the bandwidth, and the dispersion properties of the filters.

As we concluded in Section 5.3.1 on page 104, the $DE_P$ is a monotonic function of $\kappa\sqrt{L_x L_y}$. The high diffraction efficiency required by telecommunications filters can be achieved by increasing any of these three parameters.

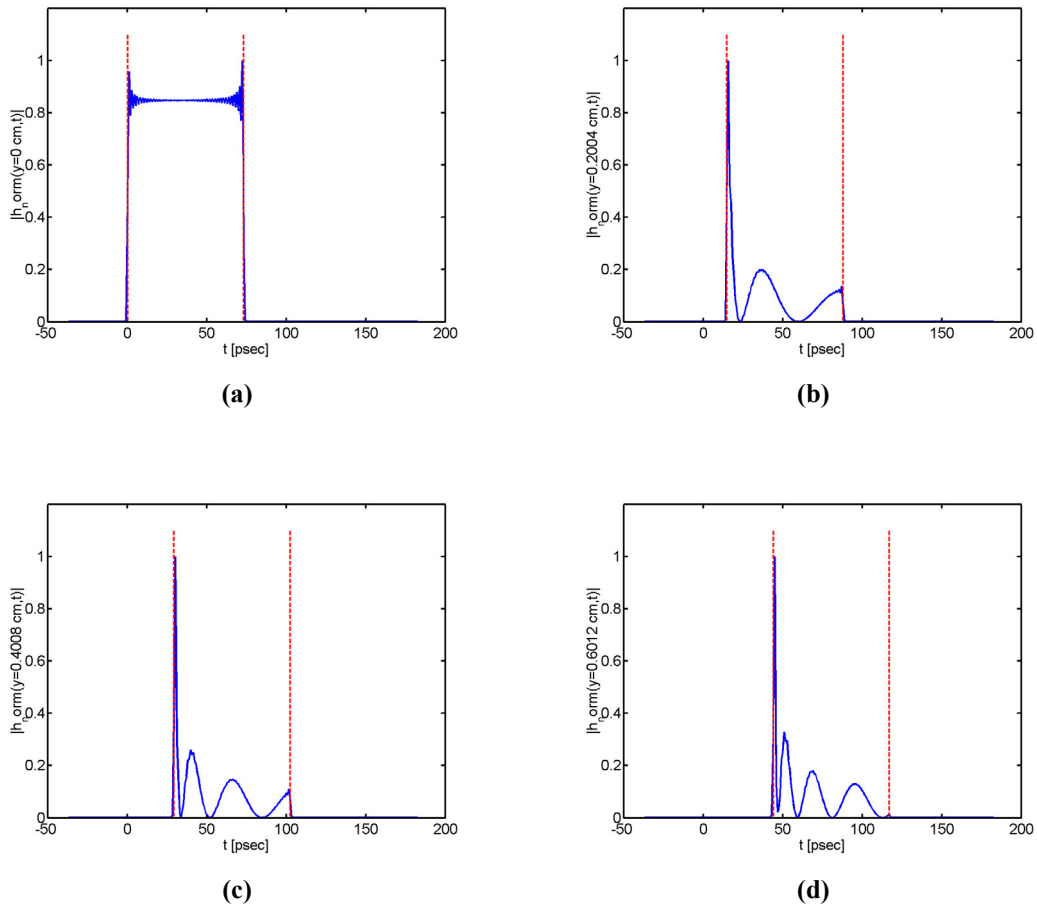



**Figure 5-14.** Impulse response vs. time for specific points of the output face of the crystal: (a) $y = 0$ cm. (b) $y = 0.2004$ cm. (c) $y = 0.4008$ cm. (d) $y = 0.6012$ cm.The dashed vertical lines correspond to $t = \frac{yn_c}{c}$ and $t = \frac{yn_c}{c} + \frac{L_x n_c}{c}$.

The first parameter, the grating coupling strength $\kappa$ depends linearly on the modulation of the refractive index $\Delta n$. For known photorefractive materials, there is a limit to the modulation depth that we can achieve. In addition, $\kappa$ also determines the bandwidth of the filter, but since the maximum $\Delta n$ we can achieve results in relatively small bandwidths, we can say that the recording material properties define the limit on the value of $\kappa$.

The second parameter, the length of the output boundary $L_y$, is limited by two factors. The first one, which also comes into play for $L_x$, is the degree of difficulty

involved with the recording of a uniform grating over a large volume. The second factor has to do with the coupling of the diffracted light to a fiber, or any other waveguide. Even if we assume that we can elongate $L_y$ at will, we will have to collect all (or at least most) of the diffracted light coming out of that surface, and couple it to a fiber. It should be obvious that the longer $L_y$, the more challenging, and less efficient, this coupling will be, possibly balancing the effects of a high $L_y$ on the final $DE_P$.

The third parameter, the length of the input boundary $L_x$, is limited, mainly, by the dispersion properties of the grating. In Section 5.5.4 on page 120 we found that the duration of the impulse response is $\tau = \frac{L_x n_c}{c}$, therefore, increasing $L_x$ would elongate the incoming pulses, which would then overlap at the output.

The abovementioned points demonstrate that the use of strong volume gratings in the 90$^o$ geometry presents considerable challenges, namely meeting the requirement for high $DE_P$, while keeping dispersion down to tolerable levels, and coupling the diffracted light back into a fiber. Collimating the light coming out of the input fiber and Bragg matching it to the grating constitutes an additional hurdle. Future work in this area should address these problems. One possible path would be to focus on the design of dispersive elements that can improve coupling efficiency and counter dispersion. An alternative route would be to manipulate the properties of the recorded grating, e.g. by chirping it, or apodizing it, as to mitigate some of the undesired effects mentioned here.

## 5.7 References

[5-1] M. Born and E. Wolf, *Principles of optics*. 1980, Cambridge, New York: Cambridge University Press.

[5-2] J. Goodman, *Introduction to Fourier optics*. Second ed. 1996: McGraw-Hill.

[5-3] H. J. Coufal, G. Sincerbox, and D. Psaltis, eds. *Topics in applied physics: holographic memories*. 1999, Springer: Berlin, New York, Tokyo.

[5-4] A. Taflove, *Computational electrodynamics*. 1995, Boston: Artech House.

[5-5] Z. Liu, *Optical information storage and processing*, Ph.D. Thesis, California Institute of Technology, 2002.

[5-6] W. H. Liu, D. Psaltis, and G. Barbastathis, *Real-time spectral imaging in three spatial dimensions.* Optics Letters, 2002. **27**(10): pp. 854-856.

[5-7] H. Kogelnik, *Coupled wave theory for thick hologram gratings.* Bell System Technical Journal, 1969. **48**(9): pp. 2909-2947.

[5-8] L. Solymar and D. J. Cooke, *Volume holography and volume gratings*. 1981, London, New York: Academic Press.

[5-9] R. P. Kenan, *Theory of crossed-beam diffraction gratings.* IEEE Journal of Quantum Electronics, 1978. **14**(12): pp. 924-929.

[5-10] A. V. Oppenheim and R. W. Schafer, *Discrete-time signal processing*. 1989: Prentice Hall.

[5-11] K. B. Hill and D. J. Brady, *Impulse responses of strong reflection holograms.* Applied Optics, 1993. **32**(23): pp. 4305-4316.

# 6 Effect of temperature on absorption in $LiNbO_3$:Fe

## 6.1 Introduction

In the previous chapters we explored the use of holographic gratings for a number of applications. One of the most promising holographic recording materials is iron-doped lithium niobate ($LiNbO_3$:Fe). The most severe limitation of $LiNbO_3$:Fe is that stored holograms are erased upon readout.

The recording mechanism in photorefractive crystals is well understood [6-1], and can be briefly described as follows: Inhomogeneous illumination excites electrons from high intensity areas. These electrons migrate and are trapped in low intensity regions. Thus electric space–charge fields build up and modulate the refractive index via the electro-optic effect. Light absorption is the initial and hence a crucial step of this photorefractive process. Homogenous light is used to read out the

recorded grating, eventually redistributing the electrons and erasing the electronic pattern.

Several techniques have been developed to limit this effect [6-2, 6-3, 6-4, 6-5]; one of the most promising among them is thermal fixing [6-6, 6-7], which consists of heating the crystal (typically to 180 $^{o}$C) during recording. As a result, the electronic space–charge pattern is copied to an ionic one that is persistent against illumination at room temperature. Although thermal fixing is applied frequently, only little is known about how the electronic excitation and transport parameters are affected by heating.

In this chapter we examine the effect of temperature on absorption in iron–doped lithium niobate crystals. First we give an outline of the absorption mechanisms (Section 6.2) and discuss theoretical considerations (Section 6.3). Then we present the experimental methods we used (Section 6.4) and the results obtained (Section 6.5). Finally we propose a theoretical model which explains our experimental observations and discuss its implications in terms of possible applications and extensions to different dopants (Section 6.6).

## 6.2 Absorption Mechanisms in $LiNbO_3$:Fe

The introduction of Fe atoms in $LiNbO_3$ crystals generates two absorption bands. Absorption of blue or green light excites electrons from $Fe^{2+}$ to the conduction band, generating $Fe^{3+}$. Furthermore, there is a crystal-field splitting of the $Fe^{2+}$ level, adding a new energy level denoted $(Fe^{2+})^*$ (Figure 6-1) [6-8]. The excitation $Fe^{2+} \rightarrow (Fe^{2+})^*$ takes place at about 1200 nm. In addition, $LiNbO_3$ shows fundamental band-to-band absorption at short wavelengths, starting with a weak tail at about 600 nm, and becoming strong below 400 nm.

The absorption peaks attributed to the abovementioned mechanisms are fairly broad. We assume that thermal vibrations of the crystal lattice induce broadening of the corresponding energy states [6-9, 6-10]. Two important outcomes of this model are: On the energy scale, absorption spectra should have a Gaussian

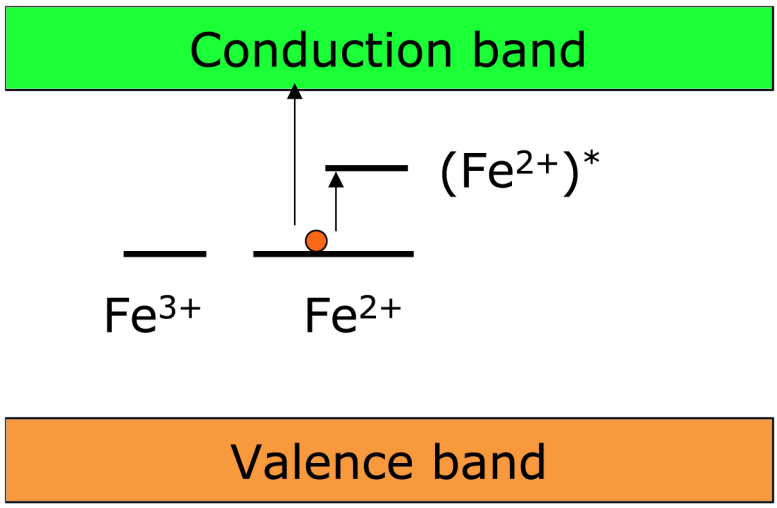



**Figure 6-1.** Absorption mechanisms in $LiNbO_3$:Fe. Two possible transitions are shown, $Fe^{2+} \rightarrow Fe^{3+} + e^-$ and $Fe^{2+} \rightarrow (Fe^{2+})^*$.

shape. In the ground state as well as in the excited state the electrons can have vibrational thermal energy that decreases or increases the photon energy required for an optical transition.

## 6.3 Theoretical considerations

As mentioned in Section 6.2, each absorption band on the $\alpha$ vs. $E$ curve should have a Gaussian shape [6-9, 6-10]:

$$\alpha = \frac{C}{\sigma}\exp\left[-\left(\frac{E-E_0}{\sigma}\right)^2\right], \qquad \textbf{(EQ 6-1)}$$

where $E_0$ is the peak absorption energy, $\sigma$ is the width of the absorption band, and $C$ is a constant proportional to the concentration of atoms responsible for the absorption in that band.

We propose that, for each absorption band, the width $\sigma$ is an increasing function of temperature $T$, while the other parameters, $E_0$ and $C$, are temperature independent. Therefore the change in absorption $\alpha$ will follow:

$$\Delta\alpha = C\left\{\frac{1}{\sigma_1}\exp\left[-\left(\frac{E-E_0}{\sigma_1}\right)^2\right] - \frac{1}{\sigma_2}\exp\left[-\left(\frac{E-E_0}{\sigma_2}\right)^2\right]\right\}, \quad \textbf{(EQ 6-2)}$$

where $\sigma_1 = \sigma(T_1)$ and $\sigma_2 = \sigma(T_2)$.

# 6.4 Experimental methods

## 6.4.1 Experimental setup

We use a Cary 500 spectrometer to measure the optical density of a $LiNbO_3$:Fe crystal. The Fe doping level of our crystal is $5.6 \times 10^{25}$ atoms/m$^3$, and the concentration ratio $c_{Fe^{2+}}/c_{Fe^{3+}}$ is 0.32. The dimensions of the crystal are $1.2 \times 4.9 \times 5.1$ mm$^3$, with the c–axis along the longest dimension. Light propagates along the shortest dimension of the crystal. A polarizer is used to select the polarization state of the incident light. We mount the crystal on a custom-made heatable holder, whose temperature we adjust with a Eurotherm 2000 temperature controller, that provides us 0.1 $^o$C accuracy. The holder keeps the crystal tilted to a 5$^o$ angle to avoid Fabry–Perot effects. We vary the temperature of the crystal in the range from 30 to 150 $^o$C with 10 $^o$C steps. For each temperature we measure the optical density, which is defined as $OD = \log\frac{I_{in}}{I_{out}}$, where $I_{in}$ is the light intensity going in the crystal and $I_{out}$ is the light intensity that makes it through the crystal.

## 6.4.2 Reflection correction

In order to calculate the absorption coefficient $\alpha$ we need to take into account the multiple reflections on both surfaces of the crystal. Assume $d_0$ is the thickness of the crystal and $\theta$ the incidence angle. Then the effective thickness of the crystal is defined as $d = \frac{d_0}{\cos\theta}$. Also assume $R = \left(\frac{n-1}{n+1}\right)^2$ (for $\theta \ll 1$) to be the reflection coefficient, where $n$ is the refractive index of the crystal. When the beam enters or exits the crystal, it maintains $(1-R)$ of its original intensity. When it is reflected, it

maintains $R$ of its original intensity, and each time it travels through the crystal, its intensity is multiplied by a factor of $\exp(-\alpha d)$ (see Figure 6-2). Taking all this into account the output intensity $I_{out}$ will be

$$I_{out} = I_{in}[(1-R)^2\exp(-\alpha d) + R^2(1-R)^2\exp(-3\alpha d) +$$

$$R^4(1-R)^2\exp(-5\alpha d) + \ldots] \Rightarrow$$

$$I_{out} = I_{in}(1-R)^2\exp(-\alpha d)\frac{1}{1-R^2\exp(-2\alpha d)} \Rightarrow$$

$$\exp(-2\alpha d) + 2\left[\frac{(1-R)^2}{2R^2}\frac{I_{in}}{I_{out}}\right]\exp(-\alpha d) = \frac{1}{R^2}$$

Defining $g = \frac{(1-R)^2}{2R^2}\frac{I_{in}}{I_{out}}$ we get:

$$\exp(-2\alpha d) + 2g\exp(-\alpha d) = \frac{1}{R^2} \Rightarrow$$

$$\exp(-2\alpha d) + 2g\exp(-\alpha d) + g^2 = \frac{1}{R^2} + g^2 \Rightarrow$$

$$[\exp(-\alpha d) + g]^2 = \frac{1}{R^2} + g^2 \Rightarrow$$

$$\exp(-\alpha d) = \sqrt{\frac{1}{R^2} + g^2} - g$$

so finally we have

$$\alpha = -\frac{1}{d}\ln\left(\sqrt{\frac{1}{R^2} + g^2} - g\right). \qquad \textbf{(EQ 6-3)}$$

We proceed to calculate the optical absorption coefficient $\alpha$, using Equation 6-3, taking into account the dependence of the refractive index $n$ on wavelength (modified Sellmeier equations) and on temperature [6-11], as well as the thermal expansion of the crystal [6-12, 6-13].

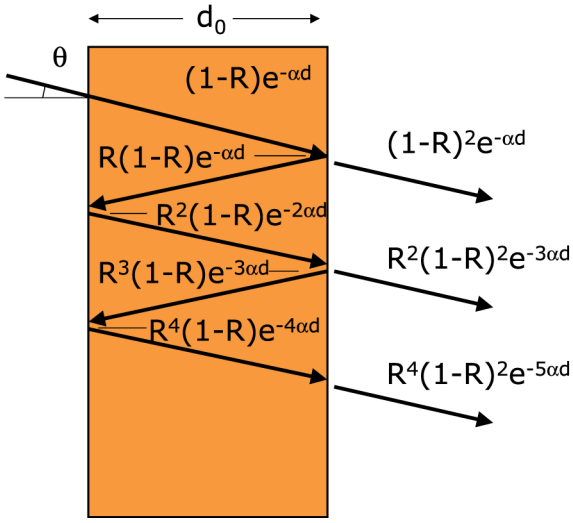



**Figure 6-2.** Reflection correction for calculation of the absorption coefficient $\alpha$ from optical density data.

# 6.5 Experimental results

## 6.5.1 Temperature dependence of absorption

Our initial results consist of curves of optical absorption coefficient $\alpha$ vs. photon energy $E$ for various temperatures (see, e.g., Figure 6-3). In order to obtain a clearer view, as well as to exclude band–edge effects, we proceed to subtract the $\alpha$ vs. $E$ curves pairwise. Thus we obtain the curves of $\Delta\alpha = \alpha(T_1) - \alpha(T_2)$ vs. $E$ shown in Figure 6-4.

We fit the experimentally obtained $\Delta\alpha$ vs. $E$ curves to Equation 6-2 on page 128. For each absorption band we isolate the corresponding photon energy region, and feed the data of all absorption differences $\Delta\alpha$ to our fitting algorithm simultaneously. We obtain as fit parameters the absorption width $\sigma(T)$ for all temperatures, as well as $C$ and $E_0$. A juxtaposition of experimental data and fitting curves can be seen in Figure 6-4.

It makes sense to keep the parameter $C$ constant across all temperatures for each absorption band, since it is related to the total number of absorbing atoms that is

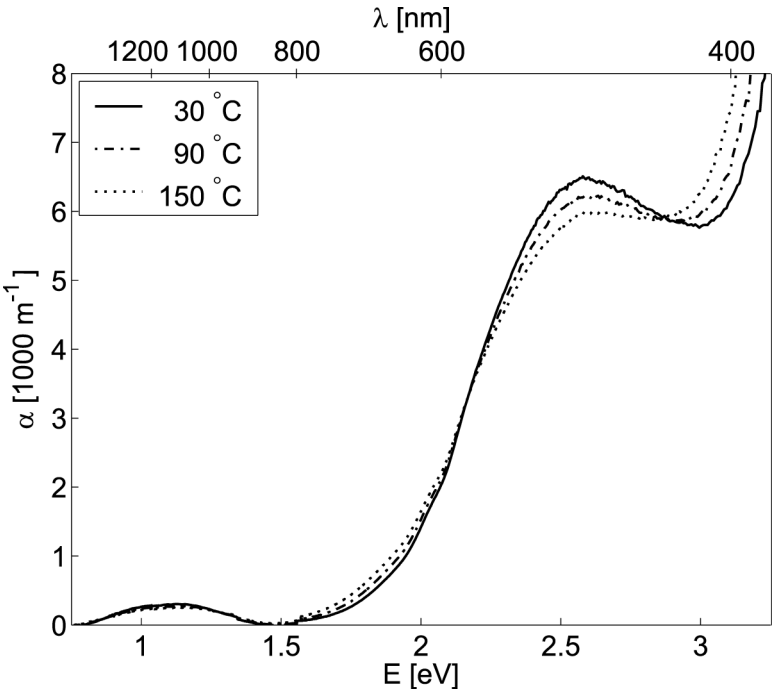



**Figure 6-3.** Absorption coefficient $\alpha$ vs. photon energy $E$ with varying crystal temperature. Ordinary light polarization is used.

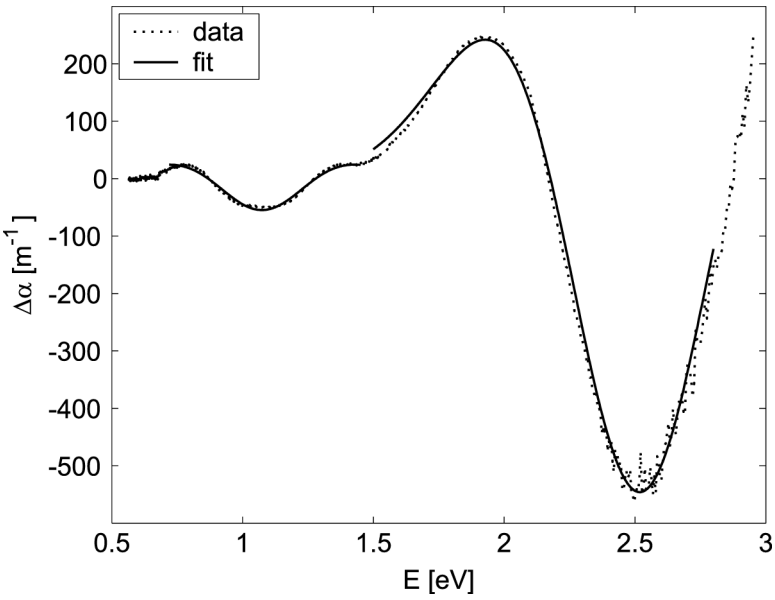



**Figure 6-4.** Absorption coefficient change $\Delta\alpha$ between 30 $^{o}$C and 150 $^{o}$C (dotted line) and fitting curve (solid line). Ordinary light polarization is used.

not changed by heating. And since just a single photorefractive center is involved, the energy position $E_0$ of the absorption band should also remain constant vs. temperature. The resulting values and 95% confidence intervals for $C$ and $E_0$ are shown in Table 6-1.

| | $C$ [$m^{-1}$] | $\delta C$ [$m^{-1}$] | $E_0$ [eV] | $\delta E_0$ [eV] |
|---|---|---|---|---|
| **Extraordinary polarization, high photon energy** | 1486 | 219 | 2.6264 | 0.0012 |
| **Ordinary polarization, high photon energy** | 1874 | 191 | 2.5179 | 0.0007 |
| **Ordinary polarization, low photon energy** | 199 | 61 | 1.0740 | 0.0005 |

**Table 6-1.** Fitted values and 95% confidence intervals for the parameters $C$ and $E_0$ of Equation 6-2.

The resulting fitting parameters for $\sigma$, the absorption band width, are shown in Figure 6-5 for all three absorption bands. Error bars indicate the 95% confidence intervals. As predicted by our model, the absorption width increases with temperature.

## 6.5.2 Model validation

The model we propose describes well the experimental data (Figure 6-4). Still, we would like to check our model from the learning theory point of view. In order to do that, we attempt to fit our experimental data to the equations detailed in Table 6-2. Since the model we propose is a difference of Gaussians, and each Gaussian has three parameters, our model should have at least four parameters, i.e., in the case where the two Gaussians share all parameters but one. Therefore, for the first two

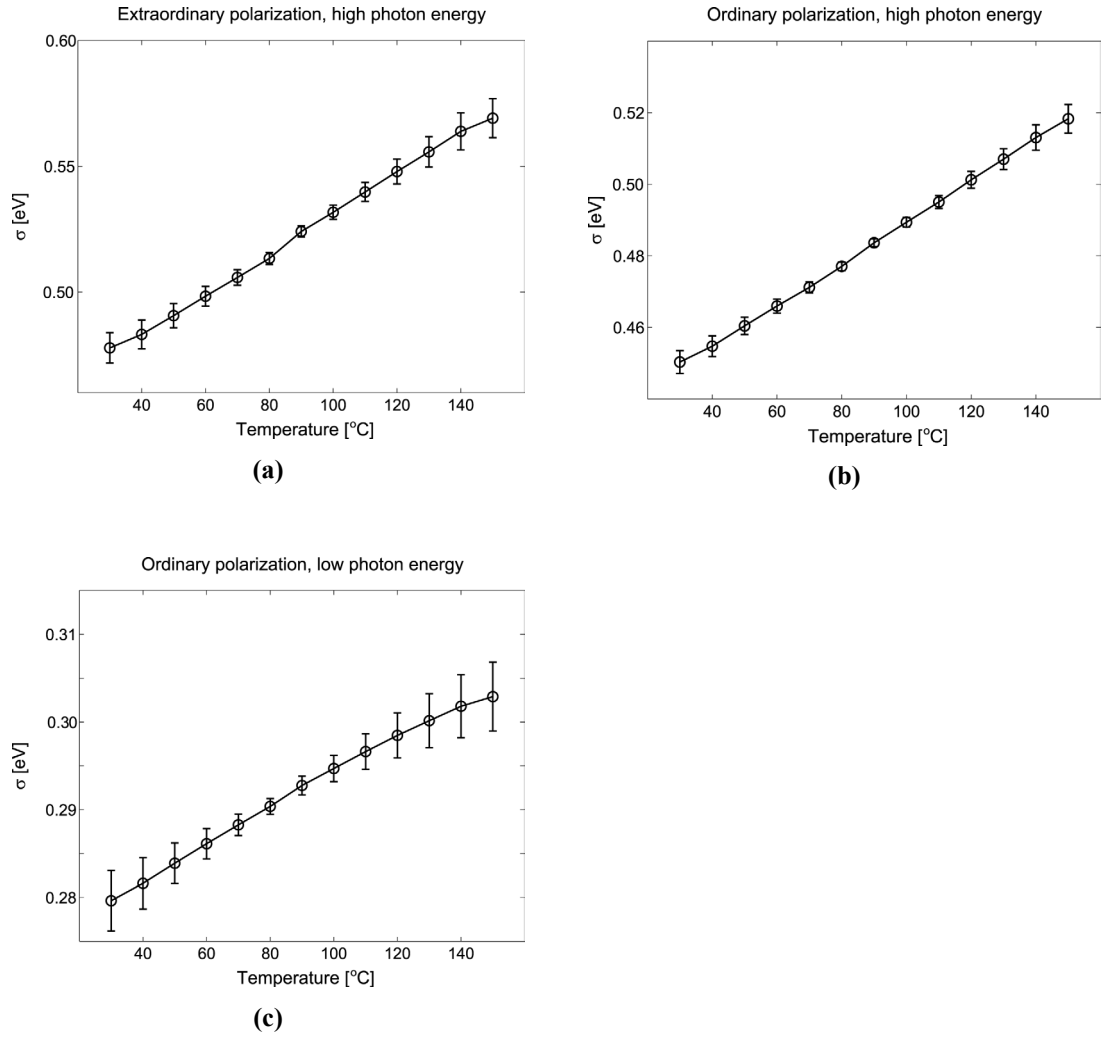



**Figure 6-5.** Absorption width σ (fit parameters to Equation 6-2) for different light polarization states and photon energy levels. (a) Extraordinary light polarization, high photon energy. (b) Ordinary light polarization, high photon energy. (c) Ordinary light polarization, low photon energy. Error bars indicate 95% confidence intervals and the connecting lines are guides to the eye.

models, having two and three fit parameters respectively, we arbitrarily choose to use polynomials.

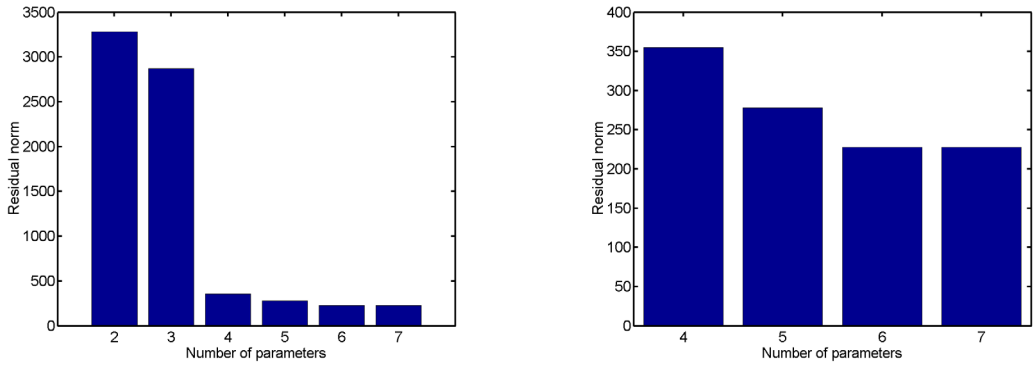
| **No. of parameters** | **Model** |
|---|---|
| 2 | $\Delta\alpha = c_1 + c_2 x$ |
| 3 | $\Delta\alpha = c_1 + c_2 x + c_3 x^2$ |

**Table 6-2.** Equations used to fit the experimental data for model validation.

| No. of parameters | Model |
|---|---|
| 4 | $\Delta\alpha = C\left\{\frac{1}{\sigma_1}\exp\left[-\left(\frac{E-E_0}{\sigma_1}\right)^2\right]-\frac{1}{\sigma_2}\exp\left[-\left(\frac{E-E_0}{\sigma_2}\right)^2\right]\right\}$ |
| 5 | $\Delta\alpha = C\left\{\frac{1}{\sigma_1}\exp\left[-\left(\frac{E-E_0}{\sigma_1}\right)^n\right]-\frac{1}{\sigma_2}\exp\left[-\left(\frac{E-E_0}{\sigma_2}\right)^n\right]\right\}$ |
| 6 | $\Delta\alpha = \frac{C_1}{\sigma_1}\exp\left[-\left(\frac{E-E_0}{\sigma_1}\right)^n\right]-\frac{C_2}{\sigma_2}\exp\left[-\left(\frac{E-E_0}{\sigma_2}\right)^n\right]$ |
| 7 | $\Delta\alpha = \frac{C_1}{\sigma_1}\exp\left[-\left(\frac{E-E_0}{\sigma_1}\right)^{n_1}\right]-\frac{C_2}{\sigma_2}\exp\left[-\left(\frac{E-E_0}{\sigma_2}\right)^{n_2}\right]$ |

**Table 6-2.** Equations used to fit the experimental data for model validation.

We proceed to fit our data using the equations in Table 6-2, and collect the residual norm for each case, i.e., the square root of the squared difference of the experimental data and the corresponding fit (see Figure 6-6). There is a substantial drop of the residual norm as we switch from three to four parameters. As we increase the number of parameters further, we observe a much slower decline.




**(a)** **(b)**

**Figure 6-6.** Residual fitting norm vs. number of fitting parameters. (a) Considerable drop from three to four parameters. (b) Moderate drop for increases of the number of parameters beyond four.

Experimental error in the measurement of the crystal thickness $d$, the refractive index $n$ and its optical density $OD$ result in uncertainty of a few percent for the

values of $\Delta\alpha$ used for fitting (see Equation 6-3 on page 129). The fitting error we obtain using four parameters is comparable to the experimental uncertainty. This precludes us from assessing that the slow decline of the residual error using more than four parameters bears convincing evidence to four being the right number of parameters for our model. On the other hand, it provides a strong indication that our theoretical model is an adequate description of the experimental data.

In order to be able to proceed to a more thorough validation of our model, we would need to reduce the experimental error. Our error analysis indicates that the main source of experimental error is the uncertainty about the thickness of the crystal $d$, which comes as no surprise given that the absorption coefficient $\alpha$ depends logarithmically on all other parameters (see Equation 6-3 on page 129). Therefore, a more accurate technique to measure the crystal thickness would be required to allow us to perform model validation.

### 6.5.3 Polarization dependence

In Figure 6-5 we present fit parameters for different light polarizations and photon energy levels. High photon energy levels (around 2.5 eV) correspond to the $Fe^{2+} \rightarrow Fe^{3+} + e^-$ transition, and low photon energy levels (about 1 eV) to the $Fe^{2+} \rightarrow (Fe^{2+})^*$ transition. No results are presented for extraordinary light polarization and low photon energy, since the $Fe^{2+} \rightarrow (Fe^{2+})^*$ transition is very weak for this light polarization [6-8].

We notice that the absorption width corresponding to the $Fe^{2+} \rightarrow Fe^{3+} + e^-$ transition ($\sigma(T = 30\ ^oC) \cong 0.46$ eV) is larger than the width corresponding to the $Fe^{2+} \rightarrow (Fe^{2+})^*$ transition ($\sigma(T = 30\ ^oC) \cong 0.28$ eV). For the high photon energy absorption band, the width of the band $\sigma$, and its variation vs. temperature, depend rather weakly on light polarization. The absorption band for extraordinary light polarization is slightly broader (about 10%) than that for ordinarily polarized light.

### 6.5.4 Absorption increase in the near infrared

In Figure 6-7 we show the absorption coefficient $\alpha$ at 780 nm vs. temperature $T$. The data shown in this figure was obtained directly from measurement, using ordinarily polarized light. The absorption coefficient $\alpha$ more than doubles as a result of moderate heating (to 150 $^{o}$C). This significant increase of absorption is very promising for a corresponding increase in holographic recording sensitivity in the infrared region.

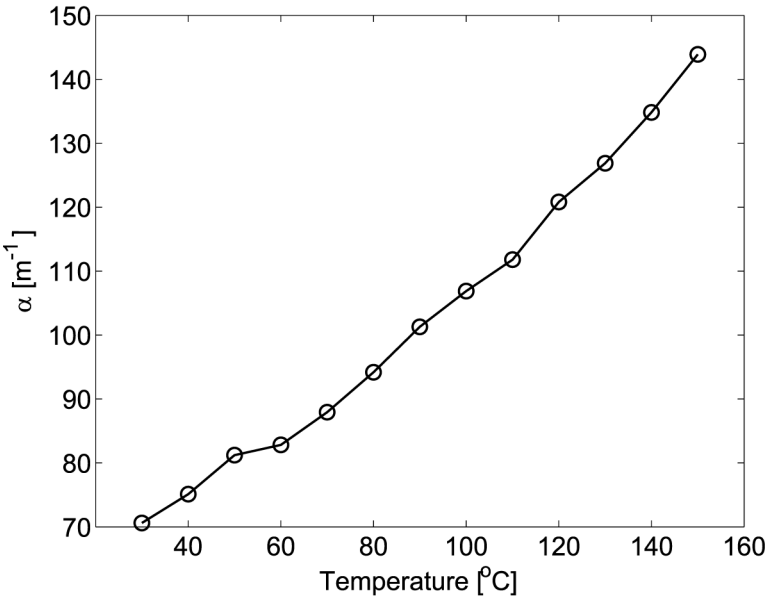



**Figure 6-7.** Absorption coefficient $\alpha$ at 780 nm vs. temperature for ordinary light polarization.

## 6.6 Discussion

The measurements and corresponding fits that are presented in previous sections show that the width of each absorption band depends on temperature. Thus our initial assumption, that the line-broadening arises chiefly from thermal processes, is confirmed. However, it can not be excluded that a portion of the broadening comes from, e.g., iron centers that are embedded into different environments of intrinsic crystal defects. But this effect is evidently not the main reason for the line broadening.

Our experiments reveal that the $Fe^{2+}$ absorption bands in $LiNbO_3$ are significantly broadened (up to about 0.07 eV for ordinary light polarization and 0.09 eV for extraordinary light polarization) by moderate heating (up to 150 $^{o}$C). The width of absorption bands increases linearly with temperature both in the visible and infrared regions.

In Equation 6-1 the absorption coefficient $\alpha$ due to a specific transition is modeled as a Gaussian whose only temperature dependent parameter is the width $\sigma$. The area under the Gaussian function, which is proportional to the number of atoms participating in the corresponding transition, should remain constant with temperature, as should the photon energy corresponding to maximum absorption $E_0$.

Our experimental results confirm this theoretical model. Indeed, we obtain good fits of experimental data to Equation 6-2, assuming constant values for $C$ and $E_0$, while $\sigma$ increases linearly with temperature.

The dependence of absorption on temperature is of great practical importance. Thermal fixing techniques that are widely used to provide holograms with long lifetimes require recording at high temperature. Knowledge of the behavior of absorption at such high temperatures is essential in enabling a better understanding of the mechanisms involved. Our results also indicate that by increasing the temperature of $LiNbO_3$:Fe, spectral sensitivity can be broadened, thus most probably facilitating improved holographic recording in the infrared region.

As we mentioned at the beginning of this section, the physical mechanisms leading to absorption broadening are very fundamental. $LiNbO_3$ crystals with different dopants and even different photorefractive crystals should exhibit similar behaviors. In all these cases temperature tuning of absorption and therefore sensitivity should be possible, enabling holographic recording in parts of the spectrum which are inaccessible at room temperature.

# 6.7 References

[6-1] K. Buse, *Light-induced charge transport processes in photorefractive crystals I: Models and experimental methods.* Applied Physics B - Lasers and Optics, 1997. **64**(3): pp. 273-291.

[6-2] K. Buse, A. Adibi, and D. Psaltis, *Non-volatile holographic storage in doubly doped lithium niobate crystals.* Nature, 1998. **393**(6686): pp. 665-668.

[6-3] J. J. P. Drolet, E. Chuang, G. Barbastathis, and D. Psaltis, *Compact, integrated dynamic holographic memory with refreshed holograms.* Optics Letters, 1997. **22**(8): pp. 552-554.

[6-4] E. Chuang and D. Psaltis, *Storage of 1000 holograms with use of a dual-wavelength method.* Applied Optics, 1997. **36**(32): pp. 8445-8454.

[6-5] J. A. Ma, T. Chang, J. Hong, R. Neurgaonkar, G. Barbastathis, and D. Psaltis, *Electrical fixing of 1000 angle-multiplexed holograms in SBN:75.* Optics Letters, 1997. **22**(14): pp. 1116-1118.

[6-6] J. J. Amodei and D. L. Staebler, *Holographic pattern fixing in electro-optic crystals.* Applied Physics Letters, 1971. **18**: p. 540.

[6-7] A. Yariv, S. Orlov, and G. Rakuljic, *Holographic storage dynamics in lithium niobate: Theory and experiment.* Journal of the Optical Society of America B-Optical Physics, 1996. **13**(11): pp. 2513-2523.

[6-8] H. Kurz, E. Kratzig, W. Keune, H. Engelmann, U. Gonser, B. Dischler, and A. Raeuber, *Photorefractive centers in $LiNbO_3$, studied by optical-, Moessbauer- and EPR-Methods.* Appl. Phys., 1977. **12**: p. 355.

[6-9] J. M. Spaeth, J. R. Niklas, and R. H. Bartram, *Structural analysis of point defects in solids*. 1992: Springer.

[6-10] H. Henry and C. P. Slichter, *Moments in the degeneracy of optical spectra*, in *Physics of Color Centers*, W.B. Fowler, Editor. 1968, Academic: New York.

[6-11] D. S. Smith and H. D. Riccius, *Refractive indices of lithium niobate.* Optics Communications, 1976. **17**: pp. 332-335.

[6-12] R. T. Smith and F. S. Welsh, *temperature dependence of the elastic, piezoelectric, and dielectric constants of lithium tantalate and lithium niobate.* Journal of Applied Physics, 1971. **42**(6): p. 2219.

[6-13] U. Schlarb and K. Betzler, *Refractive indeces of lithium niobate as a function of temperature, wavelength, and composition: A generalized fit.* Physical Review B, 1993. **48**(21): pp. 15613-15620.

# 7 Measurement of quadratic electro-optic coefficient

## 7.1 Introduction

One possible method of holographic grating tuning is the application of external electric fields [7-1, 7-2], which are used to modulate the refractive index via the electro-optic effect [7-3, 7-4]. The application of a homogeneous electric field affects the dc part of the refractive index. Nevertheless the quadratic electro-optic coefficient will affect existing gratings even with the application of a homogeneous external electric field.

The quadratic electro-optic coefficient has been measured, e.g., for paraelectric KTN [7-5]. The method described in [7-5] can be extended in a straightforward way to other cubic 432 crystals that do not have a linear electro-optic effect. The measurement of the *quadratic* electro-optic coefficient in the presence of a *linear* elec-

tro-optic coefficient is, however, considerably more challenging, since the effect of the former is usually several orders of magnitude smaller than the effect of the later for reasonable applied fields: Stated simply, any quadratic effect in the measurement system will appear at the output as quadratic through the linear electro-optic coefficient. The quadratic electro-optic effect of the crystal, being so much smaller than its linear counterpart, is virtually indistinguishable, and can be easily eclipsed by the quadratic effects of the system.

In this chapter we begin by reviewing an experimental setup used to measure the linear electro-optic coefficient. We then proceed to explain which are the limitations that prevent the use of such a setup for the measurement of the quadratic electro-optic coefficient in materials that have also linear electro-optic coefficients, highlighting the challenges presented. Finally, we present the design of a measurement system capable of capturing the quadratic electro-optic coefficient even in the presence of a much stronger linear electro-optic coefficient, as well as results for the case of manganese-doped lithium niobate ($LiNbO_3$:Mn).

# 7.2 Measurement of the linear electro-optic coefficient

## 7.2.1 Introduction

The application of an electric field to a crystal results in a change of the refractive index. This is known as the electro-optic effect. The first order approximation of this effect is the linear electro-optic effect, which is described by

$$d\left(\frac{1}{n_{ij}^2}\right) = r_{ijk}E_k, \qquad \textbf{(EQ 7-1)}$$

where $n_{ij}$ is the $i,j$ component of the impermeability tensor and $E_k$ is the $k$-component of the applied field [7-4]. The corresponding linear electro-optic coefficient is defined as $r_{ijk}$. The linear electro-optic effect of a crystal is characterized by a 3×3×3 tensor, which, due to symmetry, has 18 independent elements. In the fol-

lowing sections we will describe the theory behind an experiment to determine the linear electro-optic coefficients, and we will give experimental details and results.

### 7.2.2 Theory

We place the crystal, whose linear electro-optic coefficient we want to measure, in one arm of a Mach-Zehnder interferometer. Assume that the interferometer shown in Figure 7-1 induces a phase difference of $\phi$ between its two arms. The intensity $I$ measured by the photodetector versus the phase difference $\phi$ is given by

$$I = I_0 + I_1 \cos\phi . \qquad \textbf{(EQ 7-2)}$$

Application of a field $E$ on the crystal will induce a change of the refractive index $n$

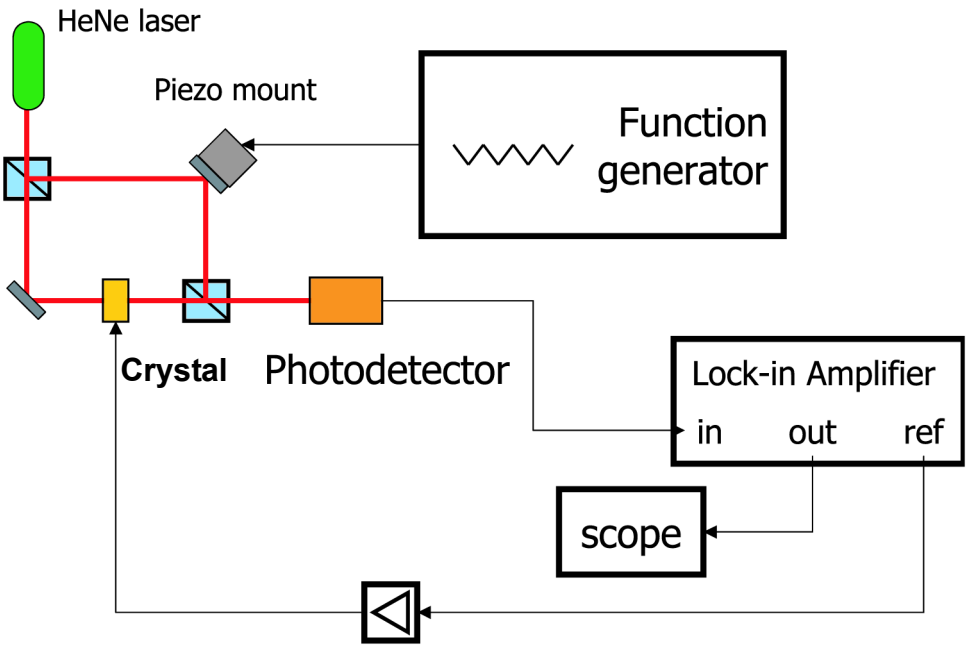



**Figure 7-1.** Schematic diagram of the experimental setup used for the measurement of the linear electro-optic coefficient of a crystal.

through the electro-optic effect, as described by Equation 7-1. Fixing the indexes $i$, $j$, and $k$, by applying the field on a specific direction and measuring the index change for a specific polarization, we can write [7-4]:

$$dn = -\frac{1}{2} n^3 r E . \qquad \textbf{(EQ 7-3)}$$

The resulting phase change is

$$d\phi = -\frac{1}{2}kLn^3rE, \quad \textbf{(EQ 7-4)}$$

where $L$ is the thickness of the crystal along the direction of light propagation, and $k = \frac{2\pi}{\lambda}$ the wavenumber, with $\lambda$ the wavelength in free space.

Assume that we apply a sinusoidal field to the crystal:

$$E(t) = E_0 \cos \omega t. \quad \textbf{(EQ 7-5)}$$

The phase difference between the two arms of the interferometer can be written as $\phi = \phi_0 + d\phi$, where $d\phi$ is given by Equation 7-4. For reasonable values of $E$, i.e., in the order of a few thousand V/m, we expect $d\phi \ll 1$. We can now expand Equation 7-2 to its Taylor series, obtaining:

$$I(\phi_0 + d\phi) = I_0 + I_1 \cos\phi_0 - I_1(d\phi)\sin\phi_0 - \frac{I_1}{2}(d\phi)^2\cos\phi_0 + \frac{I_1}{3!}(d\phi)^3\sin\phi_0 + \ldots \quad \textbf{(EQ 7-6)}$$

Given Equation 7-4, the third term of Equation 7-6 will be proportional to the linear electro-optic coefficient $r$, and the fourth and higher terms will be proportional to $r^2$, $r^3$, etc., and, since $r \ll 1$ and $d\phi \ll 1$, can be ignored. Therefore, using Equations 7-4 and 7-5, Equation 7-6 can be simplified to:

$$I(\phi_0 + d\phi) = I_0 + I_1 \cos\phi_0 + \frac{1}{2}I_1 kLn^3 rE_0 \cos(\omega t)\sin\phi_0 \quad \textbf{(EQ 7-7)}$$

It becomes obvious that the output intensity $I$ will have a frequency component $\omega$ of amplitude $\frac{1}{2}I_1 kLn^3 rE_0 \sin\phi_0$. We can use a Lock-In Amplifier to measure the amplitude of this frequency component, as shown on Figure 7-1. Since we apply a triangular signal on the piezo actuator which controls one of the mirrors of the interferometer, $\sin\phi_0$ will take values in the range $[-1, 1]$, therefore the amplitude of the trace on the scope $V_{\mathrm{LIA}}$ will be proportional to $\frac{1}{2}I_1 kLn^3 rE_0$. Since Lock-In Amplifiers output the RMS value of the frequency component under measurement we will have for the peak-to-peak (p-p) measured value:

$$\sqrt{2}V_{\text{LIA,p-p}} = cI_1 kLn^3 rE_0, \quad \textbf{(EQ 7-8)}$$

where $c$ is a constant depending on the efficiency of the photodiode.

If we apply no field on the crystal, Equation 7-6 becomes $I(\phi_0 + d\phi) = I_0 + I_1 \cos\phi_0$. So by measuring the output of the photodiode without applying any field, we can obtain the voltage value proportional to $I_1$, which we define as:

$$V_{\text{PD,p-p}} = 2cI_1. \quad \textbf{(EQ 7-9)}$$

Finally, the amplitude of the applied field $E_0$ can be calculated from the applied voltage $V_{\text{app}}$ and the dimension of the crystal along the direction of the applied field, $L_c$:

$$E_0 = \frac{V_{\text{app,p-p}}}{2L_c}. \quad \textbf{(EQ 7-10)}$$

Taking into account Equation 7-8, Equation 7-9, and Equation 7-10 we finally obtain

$$r = \frac{2\sqrt{2}V_{\text{LIA,pp}}}{V_{\text{PD,p-p}}V_{\text{app,p-p}}}\frac{1}{\pi n^3}\frac{\lambda L_c}{L}, \quad \textbf{(EQ 7-11)}$$

where we have used $k = \frac{2\pi}{\lambda}$.

### 7.2.3 Experimental methods and results

In this section we give details about the experimental setup, present the measured values, and compare them to existing results from literature. Our sample is a Mn doped Lithium Niobate crystal, with doping level 0.5 wt%. The dimensions of the crystal are 3.98 × 0.84 × 5 mm$^3$, with the c-axis lying along the longest dimension. The refractive index of the crystal is $n = 2.2$. As a light source we use a HeNe laser, with $\lambda = 632\ \text{nm}$. The light propagates along the dimension of length $L = 0.84\ \text{mm}$ and is polarized along the c-axis of the crystal.

We apply a triangular waveform to the piezo, whose frequency is 0.02 Hz and whose peak-to-peak amplitude is about 100 V after amplification. With no applied field we measure the peak-to-peak value of the photodetector $V_{\text{PD,p-p}} = 377\ \text{mV}$.

We then apply a field to the crystal along the c-axis, at frequency $f = 9.963\ \text{kHz}$ and amplitude $V_{\text{app,p-p}} = 154\ \text{V}$. The dimension of the crystal along the direction of the applied field is $L_c = 5\ \text{mm}$. We use a Stanford Research Systems SR830 DSP Lock-In Amplifier to measure the effect of the applied field at frequency $f$. The measured output of the Lock-In Amplifier is $V_{\text{LIA,pp}} = 5.265\ \text{mV}$.

Given the light polarization and the direction of the applied field, the linear electro-optic coefficient that we measure is $r_{33}$. Plugging the experimentally measured values in Equation 7-11 for $LiNbO_3$:Mn, we obtain $r_{33} = 28.9\times10^{-12}\ \frac{\text{m}}{\text{Volt}}$. Values found in literature range from $r_{33} = 28.8\times10^{-12}\ \frac{\text{m}}{\text{Volt}}$ to $r_{33} = 36.7\times10^{-12}\ \frac{\text{m}}{\text{Volt}}$ [7-6, 7-7], and measurements we have performed using different methods have convinced us that the value of $r_{33}$ should lie at the lower end of that range. Therefore our method produces results that are consistent with our expectations.

# 7.3 Measurement of the quadratic electro-optic coefficient

## 7.3.1 Introduction

The electro-optic effect can be described more accurately by adding a quadratic term, so that Equation 7-1 on page 141 becomes

$$d\left(\frac{1}{n_{ij}^2}\right) = r_{ijk}E_k + s_{ijkl}E_kE_l. \qquad \textbf{(EQ 7-12)}$$

The quadratic electro-optic effect of a crystal is characterized by a 3 × 3 × 3 × 3 tensor, which, due to symmetry, has 36 independent elements. In the following sections we will give the theoretical background that will allow us to analyze the

challenges pertaining the measurement of the quadratic electro-optic coefficient in the presence of a much stronger linear electro-optic effect. We will devise a measurement technique that circumvents these difficulties, and present the corresponding experimental details and results.

## 7.3.2 Theory

In the presence of a quadratic electro-optic coefficient $s$ Equation 7-3 on page 142 becomes

$$dn = -\frac{1}{2}n^3(rE + sE^2), \qquad \textbf{(EQ 7-13)}$$

and Equation 7-4 becomes

$$d\phi = -\frac{1}{2}kLn^3(rE + sE^2), \qquad \textbf{(EQ 7-14)}$$

where $d\phi \ll 1$.

The application of a sinusoidal field, as described by Equation 7-5, would lead in this case to the output intensity $I$ having both a frequency component $\omega$ and a frequency component $2\omega$. Unfortunately, we cannot attribute the $2\omega$ component exclusively to the quadratic electro-optic effect. This is clearly illustrated in Figure 7-2, where it is shown how the linear electro-optic effect can produce a $2\omega$ output, if the phase difference $\phi_0$ is not $(2m+1)\frac{\pi}{2}$, $m \in \aleph$.

Alternatively, from Equation 7-6 on page 143, we see that the term $-\frac{I_1}{2}(d\phi)^2\cos\phi_0$ will generate a $2\omega$ component even if we assume the quadratic electro-optic coefficient $s$ to be zero. By examining Equation 7-6 more carefully, we notice that, assuming $\phi_0 = \frac{\pi}{2}$, we obtain

$$I\left(\frac{\pi}{2} + d\phi\right) = I_0 - I_1 d\phi + \frac{I_1}{3!}(d\phi)^3 + \ldots \qquad \textbf{(EQ 7-15)}$$

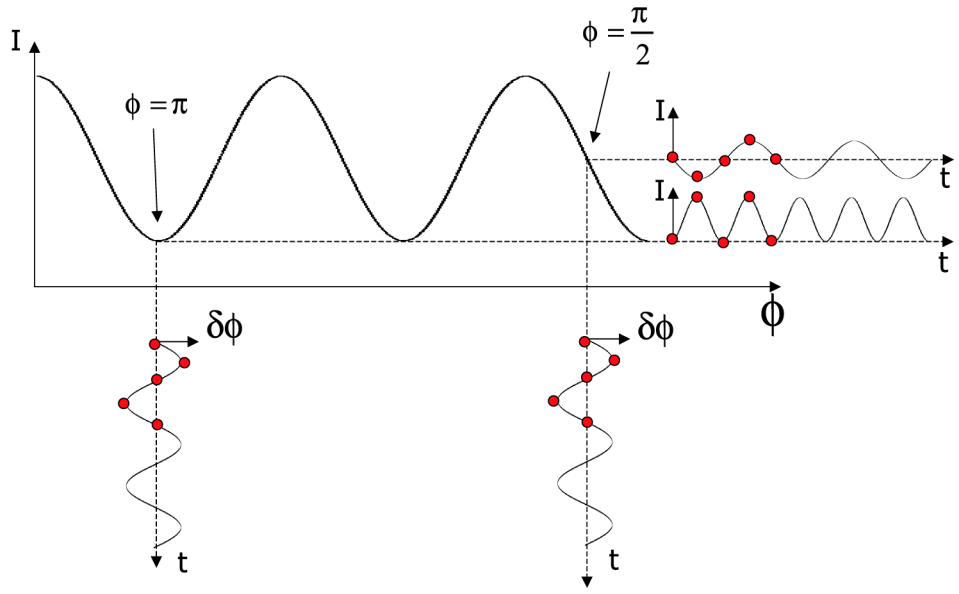



**Figure 7-2.** This plot demonstrates graphically how the same small phase perturbation can lead to intensity perturbations of different frequencies, depending on the center phase.

Therefore the measured intensity $I$ is an odd function of $d\phi$, and as a consequence the linear electro-optic effect will not produce any $2\omega$ component. In addition to that, since both $r$ and $s$ are much smaller than unity, for reasonable values of the applied field $E$, we can ignore $(d\phi)^3$ and terms of higher order, thus obtaining

$$I\left(\frac{\pi}{2}+d\phi\right) = I_0 - I_1 d\phi . \qquad \textbf{(EQ 7-16)}$$

So by locking $\phi_0$ to $\frac{\pi}{2}$ we should, in principle, be able to measure the quadratic electro-optic coefficient $s$.

In the next section we will present a setup that allows us to lock the phase difference between the two arms of the Mach-Zehnder interferometer to either $m\pi$, or $(2m+1)\frac{\pi}{2}$, $m \in \aleph$. The sections after that do not depend on this setup, which is presented here for the sake of completeness.

### 7.3.3 Phase locker setup

Active phase locking is a well known technique that is used widely to record strong holograms in photorefractive materials [7-8, 7-9]. In this section we present an overview of this technique in the context of the formalism we have developed.

In Section 7.2.2 we arrived at Equation 7-6 on page 143 by assuming that a small perturbation of the phase difference is induced through the application of an electric field, via the electro-optic effect. We will now assume that no field is applied to the crystal (for that matter a crystal does not even need be present in the path of the interferometer), but instead that a small sinusoidal perturbation is applied on the piezo on which one of the mirrors is mounted (see Figure 7-3), leading to $d\phi = d\phi_0 \cos\omega t$. Additionally, we assume that the phase perturbation $d\phi_0$ is small enough, so that we can ignore higher order terms, thus obtaining

$$I(\phi_0 + d\phi) = I_0 + I_1 \cos\phi_0 - I_1 (d\phi) \sin\phi_0 - \frac{I_1}{2}(d\phi)^2 \cos\phi_0 . \qquad \textbf{(EQ 7-17)}$$

If we set up the Lock-In-Amplifier to detect the second harmonic $2\omega$, it should be evident from Equation 7-17 that it will provide an output proportional to $-I_1 \cos\phi_0$. If we integrate this output, add it to the sinusoidal excitation and feed it to the piezo, the dc component of the phase $\phi_0$ will change until it takes a value such that $\cos\phi_0 = 0$. At this point the output of the Lock-In Amplifier will become zero, therefore the output of the integrator will remain constant and $\phi_0$ will stop changing. The solution of $\cos\phi_0 = 0$ is $\phi_0 = (2m+1)\frac{\pi}{2}$, $m \in \aleph$. It is easy to verify that, depending on the phase setting of the Lock-In Amplifier, either $\phi_0 = (4m+1)\frac{\pi}{2}$, or $\phi_0 = (4m+3)\frac{\pi}{2}$, $m \in \aleph$, will be stable equilibrium points. In any case, using this technique, we are able to lock $\phi_0$ in a position where the intensity measured by the photodetector $I$ is locally an odd function of $d\phi$.

If we set up the Lock-In-Amplifier to detect the first harmonic $\omega$ we can show with a similar derivation that the phase $\phi_0$ will be locked at $\phi_0 = 2m\pi$ or $\phi_0 = (2m+1)\pi$, $m \in \aleph$, where $I$ is locally an even function of $d\phi$.

We should also mention that by using a frequency $\omega_{LP}$ for the phase-locking setup and a different frequency $\omega_{eo}$ for the field applied on the crystal, we are able to lock the phase of the setup without affecting any measurements performed using a different Lock-In Amplifier to characterize the electro-optic effect.

### 7.3.4 Harmonic limited measurement setup

Based on the discussion in Section 7.3.2, it would seem that by adding a phase-locker to the setup used to measure the linear electro-optic coefficient and by using the Lock-In Amplifier to detect the second harmonic of the applied electric field, we should be able to measure the quadratic electro-optic coefficient. The schematic diagram of such a setup is shown in Figure 7-3. Indeed, using the phase locker we

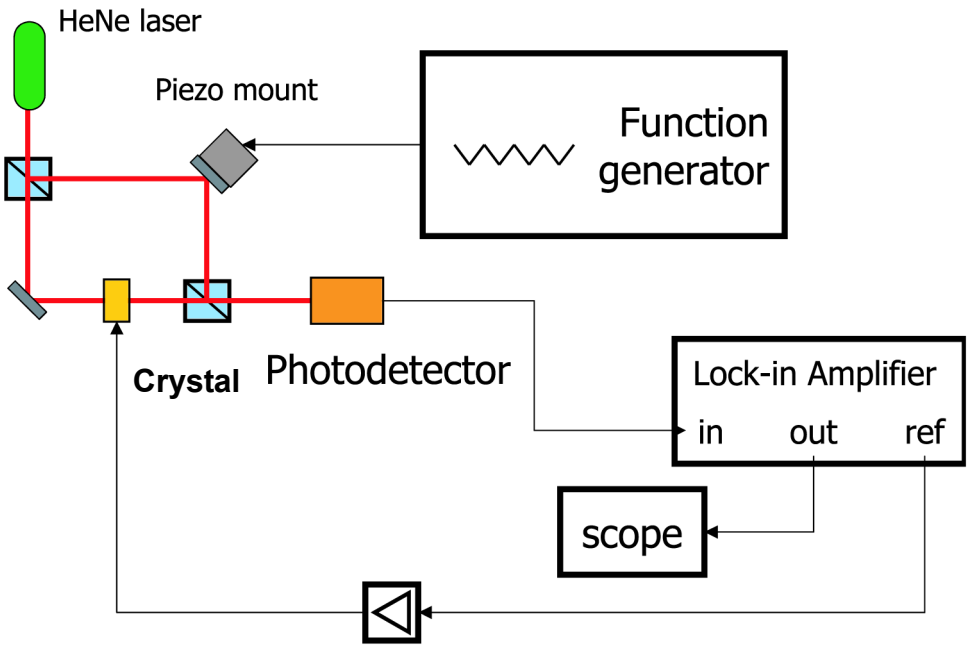



**Figure 7-3.** Schematic diagram of the experimental setup intended to measure the quadratic electro-optic coefficient of a crystal.

are able to lock the phase difference between the two arms of the interferometer to $\phi_0 = \frac{\pi}{2}$, thus the intensity monitored by the photodiode is described by Equation 7-16. Plugging Equation 7-14 into Equation 7-16 we get

$$I\left(\frac{\pi}{2} + d\phi\right) = I_0 + \frac{I_1}{2} kLn^3(rE + sE^2). \quad \textbf{(EQ 7-18)}$$

Assuming the applied field is $E(t) = E_0 \cos\omega t$, we can easily conclude that the $2\omega$ component of the intensity detected by the Lock-In Amplifier will be $\frac{I_1}{4} kLn^3 sE_0^2$. It would seem that we can use this frequency component to measure the quadratic electro-optic coefficient $s$, if it were not for one important detail: All function generators, no matter how precise, generate harmonics. Therefore a more realistic assumption for the applied field would be

$$E(t) = E_1 \cos\omega t + E_2 \cos 2\omega t. \quad \textbf{(EQ 7-19)}$$

The ratio of first to second harmonics is typically $\frac{E_1}{E_2} \approx 10^4$ for the output of a digitally synthesized waveform generator, such as the reference generator of our Lock-In Amplifier. This ratio is decreased when the reference signal is amplified through a high voltage amplifier, before being applied to the crystal. As a result of Equation 7-19 the $2\omega$ component will be $\frac{I_1}{4} kLn^3 sE_1^2 + \frac{I_1}{2} kLn^3 rE_2$, so instead of measuring $s$ we would be measuring $s + 2\frac{E_2}{E_1^2} r$. For typical parameter values the second term could easily eclipse the quadratic electro-optic coefficient that we want to measure.

More generally, any nonlinearity along the loop formed by the Lock-In Amplifier, high voltage amplifier, crystal, and photodetector, will contribute to the $2\omega$ component through the linear electro-optic coefficient $r$. Though the quadratic electro-optic coefficient $s$ will be manifested through the main driving force, the square of

the first harmonic of the applied field $E_1^2$, it could still have a significantly smaller contribution to the $2\omega$ term, thus rendering this technique inadequate for its measurement.

### 7.3.5 Quadratic electro-optic measurement setup

As we saw in Section 7.3.4, any quadratic effect along the measurement loop will allow the linear electro-optic effect to produce a $2\omega$ component at the output of the measurement system. Since, for reasonable values of the applied field $E$, the effect of the linear electro-optic coefficient $r$ is expected to be larger than the effect of the quadratic electro-optic coefficient $s$ by several orders of magnitude, any nonlinearity of the measurement system will interfere with the measurement of $s$. A simple solution to this problem is to use the quadratic electro-optic effect as a nonlinearity that will allow us to mix two different signals applied on the crystal. The experimental setup for this approach is shown in Figure 7-4 and operates as follows:

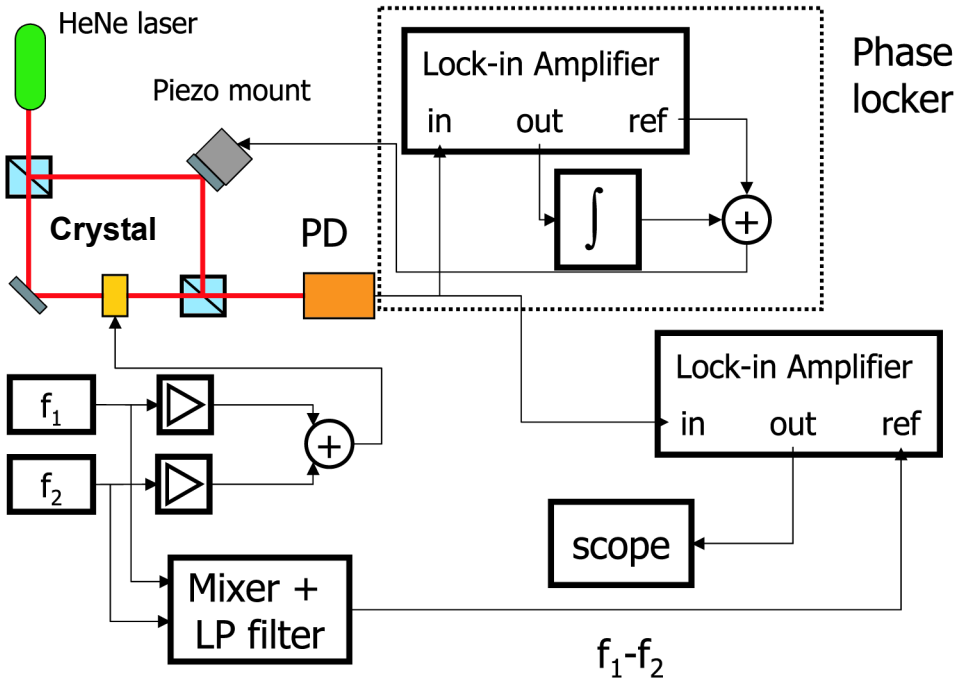



**Figure 7-4.** Schematic diagram of the experimental setup used for the measurement of the quadratic electro-optic coefficient of a crystal.

Two frequencies, $\omega_1$ and $\omega_2$ with $\omega_1 > \omega_2$ are generated and amplified independently. They are then applied in series to our crystal sample. Assuming that they each have a second harmonic, we will have for the total applied field on the crystal:

$$E = E_1\cos\omega_1 t + E_2\cos(\omega_2 t + \theta_2) + E_3\cos(2\omega_1 t + \theta_3) + E_4\cos(2\omega_2 t + \theta_4), \quad \textbf{(EQ 7-20)}$$

where we assume that the second harmonics are considerably smaller than the first ones, i.e., $\frac{E_1}{E_3} \cong \frac{E_2}{E_4} \cong 10^4$.

The two signals composing $E$ are also mixed and low-pass filtered as to produce a signal of frequency $\omega_1 - \omega_2$, which is fed to the Lock-In Amplifier as a reference signal.

Squaring Equation 7-20 generates a number terms of different frequencies. Still the only component of frequency $\omega_1 - \omega_2$ is the term $E_1E_2\cos[(\omega_1 - \omega_2)t + \theta_2]$. Note that this frequency component is produced by the squared applied field $E^2$, and therefore will appear at the output only through the quadratic electro-optic effect. The linear electro-optic effect does not generate any signal at this frequency. So, going back to Equation 7-15 on page 146 and using Equation 7-14, we get

$$I_{\omega_1-\omega_2} = -\frac{I_1}{2}kLn^3sE_1E_2. \quad \textbf{(EQ 7-21)}$$

At this point we are again faced with the question of how to connect these results with experimentally measurable quantities. Again, very similarly to the derivations at the end of Section 7.2.2, we have

$$\sqrt{2}V_{\text{LIA,p-p}} = cI_1kLn^3sE_1E_2, \quad \textbf{(EQ 7-22)}$$

where $c$ is a constant depending on the efficiency of the photodiode, and

$$E_1E_2 = \frac{V_{1,\text{ app,p-p}}}{2L_c}\frac{V_{2,\text{ app,p-p}}}{2L_c}. \quad \textbf{(EQ 7-23)}$$

Putting together Equation 7-21, Equation 7-22, Equation 7-23 and Equation 7-9 on page 144 we finally get

$$s = \frac{8\sqrt{2}V_{\text{LIA}}}{V_{\text{PD,p-p}}V_{1,\text{ app,p-p}}V_{2,\text{ app,p-p}}}\frac{1}{\pi n^3}\frac{\lambda L_c^2}{L}. \qquad \textbf{(EQ 7-24)}$$

### 7.3.6 Experimental methods and results

Many of the experimental details for this measurement, such as the sample and the light source, remain the same as those mentioned in Section 7.2.3 on page 144. For the phase locker, we apply a sinusoidal modulation of $f_p = 778.8\ \text{Hz}$ on the piezo actuator. We lock onto the second harmonic and achieve stable equilibrium, which we successfully test by applying small perturbations to the system.

We then apply two independently generated and amplified sinusoidal fields on the crystal, with frequencies $f_1 = 8.003596\ \text{kHz}$ and $f_2 = 7\ \text{kHz}$. We also mix and low-pass filter the output of the two generators using a custom-made circuit, in order to produce a signal of frequency $f_1 - f_2$, which we then feed as the reference to a second Lock-In Amplifier. In order to improve the accuracy of the measurement we vary the peak-to-peak values of the applied voltages $V_{1,\text{ app,p-p}}$ and $V_{2,\text{ app,p-p}}$ measure the output of the Lock-In Amplifier $V_{\text{LIA}}$ at $f_1 - f_2$. As expected, the output of the amplifier is proportional to the product of the applied voltages, as can be seen on Figure 7-5. By performing a linear fit to the acquired data we obtain a measurement for $\frac{V_{\text{LIA}}}{V_{1,\text{ app,p-p}}V_{2,\text{ app,p-p}}} = 9.4\times10^{-10}\ \frac{1}{\text{V}}$.

Given the light polarization and the direction of the applied fields the quadratic electro-optic coefficient that we measure is $s_{33}$. Plugging the value for $\frac{V_{\text{LIA}}}{V_{1,\text{ app,p-p}}V_{2,\text{ app,p-p}}}$ into Equation 7-24, we finally get $s_{33} = 9.7\times10^{-18}\left(\frac{\text{m}}{\text{V}}\right)^2$.

## 7.4 Discussion

In this chapter we developed a method to measure the *quadratic* electro-optic coefficients of a crystal that has much larger corresponding *linear* electro-optic coefficients. We discussed how the presence of a linear electro-optic effect, combined

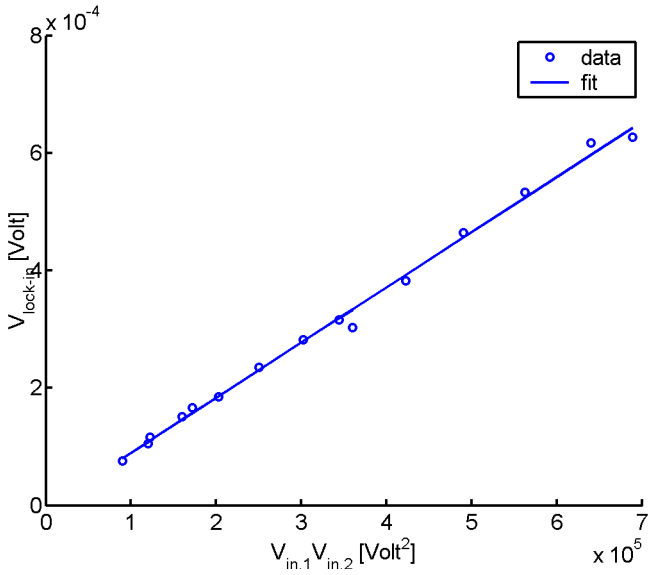



**Figure 7-5.** Output of the Lock-In Amplifier vs. the product of applied voltages on the crystal. The circles represent data points and the solid line a linear fit.

with quadratic effects in the measurement loop, can affect our measurement, and we proposed a technique to overcome this hurdle.

This technique was used to measure $s_{13}$, and resulted in a value of $|s_{13}| = 4\times10^{-19}$ $m^2/V^2$ [7-10]. The same coefficient $s_{13}$ was measured using two different methods: The first one consists of recording holographically a grating in the crystal, and then applying a high field, and monitoring its effect on the Bragg angle. The second also consists of applying a high field on a recorded grating, but in this case the value of $s_{13}$ is inferred from the effect of the field on the diffraction efficiency. Both methods yield essentially the same value, namely, $|s_{13}| = 7\times10^{-20}$ $m^2/V^2$ [7-10], which is smaller than the one measured interferometrically. This discrepancy could be the result of some remnant quadratic effect in the measurement loop of the interferometric setup. Therefore, at this point, we can only state that our method provides an upper bound for the quadratic electro-optic coefficient, until further investigations lead us to the root of the discrepancy.

# 7.5 References

[7-1] J. V. Alvarez-Bravo, R. Muller, and L. Arizmendi, *Electric-field multiplexing of volume holograms in $LiNbO_3$.* Europhysics Letters, 1995. 31(8): pp. 443-448.

[7-2] B. Srinivasan and R. K. Jain, *First demonstration of thermally poled electro-optically tunable fiber Bragg gratings*. IEEE Photonics Technology Letters, 2000. 12(2): pp. 170-172.

[7-3] A. Yariv and P. Yeh, *Optical waves in crystals*. 1984, New York: John Wiley & Sons.

[7-4] J. F. Nye, *Physical properties of crystals, their representation by tensors and matrices*. 1985, New York: Oxford University Press.

[7-5] K. Buse, F. Havermeyer, L. Glabasnia, K. Schlomp, and E. Kratzig, *Quadratic polarization-optic coefficients of cubic $KTa_{1-x}Nb_xO_3$ crystals.* Optics Communications, 1996. **131**(4-6): pp. 339-342.

[7-6] I. P. Kaminow, E. H. Turner, R. L. Barns, and J. L. Bernstein, *Crystallographic and electro-optic properties of cleaved $LiNbO_3$.* Journal of Applied Physics, 1980. **51**(8): pp. 4379-4384.

[7-7] R. Holmes, Y. Kim, and C. Brandle, *Evaluation of Crystals of $LiNbO_3$ doped with MGO or $TiO_2$ for electro-optic devices.* Ferroelectrics, 1983. **51**(1-2): pp. 41-45.

[7-8] P. A. M. Dossantos, L. Cescato, and J. Frejlich, *Interference-term real-time measurement for self-stabilized 2-wave mixing in photorefractive crystals.* Optics Letters, 1988. **13**(11): pp. 1014-1016.

[7-9] J. Frejlich, L. Cescato, and G. F. Mendes, *Analysis of an active stabilization system for a holographic setup.* Applied Optics, 1988. **27**(10): pp. 1967-1976.

[7-10] M. Luennemann, Personal Communication, June 2002.