Optical Microsystems in Silicon Based on a Fabry-Perot Resonance Cavity

Application for spectral analysis of visible light

José Higino G. Correia
OPTICAL MICROSYSTEMS IN SILICON
BASED ON A FABRY-PEROT
RESONANCE CAVITY

Application for spectral analysis of visible light
OPTICAL MICROSYSTEMS IN SILICON
BASED ON A FABRY-PEROT
RESONANCE CAVITY

Application for spectral analysis of visible light

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus
Prof.ir. K.F. Wakker,
in het openbaar te verdedigen
ten overstaan van een commissie
door het College voor Promoties aangewezen,
op dinsdag 11 mei 1999 te 13:30 uur

door

José Higino Gomes CORREIA

Engenheiro Físico, Universidade de Coimbra, Portugal
geboren te Carmona, Angola
Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. J.H. Huijsing

Samenstelling promotiecommissie:

De Rector Magnificus
Prof. dr. ir. J.H Huijsing
Dr. ir. R.F. Wolffenbuttel
Prof. dr. ir. P.P.L. Regtien
Prof. dr. C. Couto
Prof. dr. C.I.M. Beenakker
Prof. dr. W. Mokwa
Dr. M. Bartek

Technische Universiteit Delft, voorzitter
Technische Universiteit Delft, promotor
Technische Universiteit Delft, toegewezen promotor
Universiteit Twente
University of Minho, Portugal
Technische Universiteit Delft
RWTH Aachen, Germany
Technische Universiteit Delft

Published and distributed by:

Delft University Press
P.O. Box 98
2600 MG Delft
The Netherlands
Phone: (+31)-15-2783254
Fax: (+31)-15-2781661
E-mail: DUP@DUP.TUDelft.NL


Copyright © 1999 by J. H. G. Correia

All rights reserved. No part of this publication may be reproduced or
distributed in any form or by any means, or stored in database or retrieval
system, without the prior written permission of the publisher or the author.

PRINTED IN THE NETHERLANDS
In memory of my father
# Table of Contents

## 1 Introduction

1.1 Spectral analysis of visible light .................................. 1

1.1.1 Wavelength-selecting elements .................................. 3

1.1.2 Scanning-type vs. array-type configuration ...................... 3

1.1.3 Macrospectrometers: state of the art .......................... 4

1.1.4 Miniaturized spectrometers: advantages, applications ......... 8

1.1.5 Microspectrometers: state of the art .......................... 8

1.2 Motivation and objectives ......................................... 14

1.2.1 Scanning-type microspectrometer ............................... 15

1.2.2 Array-type microspectrometer ................................. 16

1.2.3 A CMOS optical microsystem .................................. 16

1.3 Organization of the thesis ........................................ 17

References ....................................................................... 18

## 2 Optical theory of a spectrometer

2.1 Introduction .................................................................. 21

2.2 Basic terms .................................................................. 22

2.3 Optical spectral analysis .............................................. 24

2.4 Spectrometer description .............................................. 26

2.4.1 Entrance slit and collimating optics ............................ 27

2.4.2 Wavelength-selecting element ................................... 28

2.4.3 Output optics ......................................................... 30

2.4.4 Photodetectors ....................................................... 30

2.4.5 Scanning-type vs. array-type approach ....................... 31

2.4.6 Spectroscopic terms .............................................. 32
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>Basic theory</td>
<td>37</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Maxwell equations: free space</td>
<td>37</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Maxwell equations: isotropic medium, no charges</td>
<td>38</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Monochromatic electromagnetic waves</td>
<td>38</td>
</tr>
<tr>
<td>2.5.4</td>
<td>Optical admittance</td>
<td>39</td>
</tr>
<tr>
<td>2.5.5</td>
<td>The Poynting vector</td>
<td>40</td>
</tr>
<tr>
<td>2.6</td>
<td>Interaction of a electromagnetic wave with a multilayer stack</td>
<td>40</td>
</tr>
<tr>
<td>2.6.1</td>
<td>Ideal boundaries</td>
<td>41</td>
</tr>
<tr>
<td>2.6.2</td>
<td>A simple boundary</td>
<td>43</td>
</tr>
<tr>
<td>2.6.3</td>
<td>Optical properties of a thin film</td>
<td>46</td>
</tr>
<tr>
<td>2.6.4</td>
<td>Optical properties of a multilayer stack</td>
<td>50</td>
</tr>
<tr>
<td>2.6.5</td>
<td>Quarter-wave and half-wave optical thicknesses</td>
<td>52</td>
</tr>
<tr>
<td>2.6.6</td>
<td>Non-ideal boundaries</td>
<td>53</td>
</tr>
<tr>
<td>2.7</td>
<td>Ideal plane Fabry-Perot resonator</td>
<td>54</td>
</tr>
<tr>
<td>2.7.1</td>
<td>The principle</td>
<td>54</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Resonator modes</td>
<td>55</td>
</tr>
<tr>
<td>2.7.3</td>
<td>Resonator modes in the visible spectral range</td>
<td>58</td>
</tr>
<tr>
<td>2.8</td>
<td>Non-ideal plane Fabry-Perot resonator</td>
<td>59</td>
</tr>
<tr>
<td>2.8.1</td>
<td>Spectral response</td>
<td>59</td>
</tr>
<tr>
<td>2.8.2</td>
<td>FWHM and finesse</td>
<td>61</td>
</tr>
<tr>
<td>2.8.3</td>
<td>Free spectral range</td>
<td>62</td>
</tr>
<tr>
<td>2.8.4</td>
<td>Resolving power</td>
<td>62</td>
</tr>
<tr>
<td>2.8.5</td>
<td>Resonator medium losses</td>
<td>63</td>
</tr>
<tr>
<td>2.8.6</td>
<td>Lossy mirrors: reflection, finite size, flatness</td>
<td>63</td>
</tr>
<tr>
<td>2.8.7</td>
<td>Angle of incidence and non-parallelism</td>
<td>64</td>
</tr>
<tr>
<td>2.8.8</td>
<td>Matrix method vs. Airy function for transmission</td>
<td>65</td>
</tr>
<tr>
<td>2.9</td>
<td>Mirrors</td>
<td>65</td>
</tr>
<tr>
<td>2.9.1</td>
<td>Dielectric HL layers succession</td>
<td>65</td>
</tr>
<tr>
<td>2.9.2</td>
<td>Metallic mirrors</td>
<td>66</td>
</tr>
<tr>
<td>2.9.3</td>
<td>Operation at visible spectral range</td>
<td>67</td>
</tr>
<tr>
<td>2.9.4</td>
<td>Absorption in metallic thin-films</td>
<td>68</td>
</tr>
</tbody>
</table>
2.10 Fabry-Perot based spectrometer .......................... 68
  2.10.1 Scanning type .......................................................... 68
  2.10.2 Array type ................................................................. 69
  2.10.3 Spectrum reconstruction ........................................... 70

References ................................................................. 73

3 Scanning-type microspectrometer 75

3.1 Introduction .............................................................. 75
3.2 Diaphragm design ......................................................... 77
  3.2.1 Selection of the diaphragm geometry ......................... 77
  3.2.2 Diaphragm materials ............................................... 81
3.3 Optical design ........................................................... 81
  3.3.1 Optical properties of the materials ......................... 81
  3.3.2 Optical simulations ............................................... 83
3.4 Load deflection: analysis of static response ........ 87
  3.4.1 Input data ............................................................... 89
  3.4.2 FEM techniques ...................................................... 91
  3.4.3 FEM simulations vs. load deflection results ........... 91
  3.4.4 Influence of internal stress ................................... 94
  3.4.5 Influence of corner shape ..................................... 94
3.5 Load deflection: analysis of dynamic response .... 96
  3.5.1 Slip flow ............................................................... 98
  3.5.2 Diaphragm response time ....................................... 99
  3.5.3 Pull-in voltage ..................................................... 103
3.6 Fabrication ............................................................... 104
  3.6.1 Fabrication sequence .............................................. 105
  3.6.2 Assembly of the FPMI ............................................. 108
3.7 Experimental results ................................................. 109
  3.7.1 Spectral response .................................................. 110
  3.7.2 Stiction and hysteresis .......................................... 112
Table of Contents

3.8 Conclusions ......................................................... 113
References ................................................................. 114

4 Array-type microspectrometer ........................... 117

4.1 Introduction .............................................................. 117
4.2 Motivation ................................................................. 117
4.3 Design ......................................................... 118
  4.3.1 Selection of materials .............................................. 118
  4.3.2 Optical simulations of the etalon structure .................... 119
  4.3.3 Phototransducers .................................................... 123
  4.3.4 The complete structure ............................................. 124
  4.3.5 Optical simulation of the complete structure .................... 125
4.4 Fabrication .............................................................. 127
4.5 Experimental results .................................................. 130
4.6 Conclusions .............................................................. 133
References ................................................................. 135

5 A CMOS optical microsystem ......................... 137

5.1 Introduction .............................................................. 137
5.2 Microinstrumentation system ................................. 138
  5.2.1 Concept ............................................................ 138
  5.2.2 Microinstrumentation bus interface ............................. 140
  5.2.3 Possible configurations for optical microsystems ......... 144
Table of Contents

5.3 A CMOS on-chip integrated optical microsystem 145
   5.3.1 Monolithic vs. hybrid integration .......................... 145
   5.3.2 A CMOS single-chip microspectrometer ................... 146
   5.3.3 Fabry-Perot etalon ........................................ 147
   5.3.4 The photodiodes ........................................... 147
   5.3.5 Readout electronics ....................................... 148
   5.3.6 Optical microsystem bus interface ......................... 149

5.4 Fabrication ..................................................... 150

5.5 Experimental results .......................................... 153

5.6 Conclusions .................................................... 156

References ......................................................... 157

6 Discussion and conclusions 159

6.1 Introduction .................................................... 159

6.2 Scanning-type vs. array-type microspectrometer ............ 159

6.3 Performance comparison of scanning-type devices ............ 160

6.4 Performance comparison of array-type devices .............. 161

6.5 Surface roughness ............................................. 162

6.6 Conclusions .................................................... 164
   6.6.1 Scanning-type and array-type microspectrometer ....... 164
   6.6.2 Single-chip CMOS optical microsystem ................... 165
   6.6.3 Applications .............................................. 165

6.7 Future work .................................................... 166

References ......................................................... 169
Table of Contents

Summary 171
Samenvatting 173
Resumo 175
List of Publications 177
Acknowledgements 181
About the Author 183
1.1 Spectral analysis of visible light

Light is an electromagnetic phenomenon described by the same theoretical principles that govern all forms of electromagnetic radiation. The range of optical wavelengths contain three bands: ultraviolet (10 to 390 nm), visible (390 to 760 nm) and infrared (760 nm to 1 mm) [1].

Visible light covers only a small part of the entire electromagnetic spectrum (see Fig. 1-1), however it is the only part which can be perceived by the human eye. The eye has three kinds of light receivers: red, green and blue. Based on the spectral distribution of intensity of the incident wave, a certain color (primary or a three-stimulus mixture) is grasped mentally. Color analysis and control play an important role in our lives (art, science, technology, business, industry, etc.).

The information supplied by the human eye, working together with the visual cortex of the brain, can only be used to describe visible light in a very qualitative way. Therefore, exclusive use of vision (optimized for color difference) to quantify light and color is quite imprecise and subjective. This limited capability is sufficient in everyday life, but not in many industrial applications. Accurate measurements can be performed
with spectral analysis: a process, in which the entire radiation band of interest is split into its spectral components and the intensity of each component is measured.

Fig. 1-1  *Optical frequencies and wavelengths [1].*

Optical spectra are often analyzed to determine chemical or optical characteristics of materials. Many applications require spectral analysis at visible wavelengths (~390 nm to ~760 nm). If the analyzed object is an active source of visible radiation, lots of valuable information about physical processes can be extracted from spectral characteristics of the emitted light. Typical applications are: atomic spectroscopy (study of visible spectral lines due to emission by isotopes), astronomy and astrophysics (study of extra-terrestrial sources in the solar system, nebulae, aurorae and galaxies). Some objects, however, are not active sources of visible radiation. In this case, an external calibrated light source is the outcome. During its interaction with the object usually some of the spectral components are absorbed or reflected in different way than others. Spectral analysis of the reflected and/or transmitted light can therefore give information about the object composition and processes involved. Examples are: inspection of product defects in the industry by means of color determination, biochemical analysis (e.g. human blood, muscle
tissue, DNA fibres), light scattering by materials (liquids, plastics, polymers and gels) [2].

### 1.1.1 Wavelength-selecting elements

Spectral analysis of visible light is based on either a colorimetric or on a spectrometric principle. A colorimetric device responds to a certain optical spectrum, i.e., to a color, by generating signals that correspond to the chromaticity coordinates in a standardised colorimetric system. For measuring the intensity of every spectral component in the radiation, a colorimetric sensor is not suitable; one must use a spectrometric device.

To perform spectral analysis, one must use a wavelength-selective detector. Its principal part is an element which performs spatial or time decomposition of the incident radiation into narrowband components. A dispersive element (a prism, grating or an interferometer) is commonly used for this purpose.

A prism or grating decomposes the incident optical radiation spatially into different wavelength components. The method is based on wavelength-dependent angular deflection and results in a specific refraction (prism) or diffraction (grating) angle for each wavelength. Prisms exhibit a non-linear relationship between wavelength and refraction angle and are not often used in state-of-the-art equipment. The grating consists of a flat plate with many parallel lines or grooves by which the light beam is diffracted. The diffracted waves interfere with each other and form an interference pattern with a characteristic angular distribution of the spectral components.

In an interferometer, the wavelength selection is performed by changing the path-length difference between two or more beams which interfere with each other. From all different types of interferometers (Michelson, Fabry-Perot, Mach-Zehnder, etc...) the Fabry-Perot interference filter is most frequently used.

### 1.1.2 Scanning-type vs. array-type configuration

The individual spectral components can be measured in two ways. A movable configuration can be used with one detector and one wavelength-selecting element (WSE). The wavelength for detection is selected by mechanical scanning of the WSE or the detector (e.g. scanning
monochromator). The second option is to use an array of detectors or an array of WSEs (no movable parts are required).

1.1.3 Macrospectrometers: state of the art

A wide range of spectrometers for the spectral analysis of visible light, based on the wavelength-selecting elements presented in the previous section are commercially available. A few examples, ranging from high-accuracy laboratory grade equipment to miniaturized hand-held systems, are presented in this section. These devices have in common that they are fabricated in the conventional way, i.e., they consist of discrete optical components and fine mechanics have been used for their assembly.

The parameter that is commonly used to compare different systems is the resolving power (also called resolution), $R$. It is defined as $R = \frac{\lambda}{\Delta \lambda}$, with $\Delta \lambda$ the smallest wavelength difference distinguishable at a specific wavelength, $\lambda$.

A very high-resolution grating spectrometer was developed by Lindblom et al. [3] (Fig. 1-2). Wavelength scanning is accomplished through variation of the refractive index of the gas filling the monochromator body, which necessitates control of the temperature within ±0.01°C. With this instrument a very high resolution can be obtained ($R = 2 \times 10^6$).

The optical path length available (total dispersion of the system is equal to the sum of the dispersion of the four individual gratings) also determines the high resolution. Thus, the very high $R$ implies large dimensions. Small values of $R$ are a limitation of microspectrometers.

The dimensions of the device are in the order of a meter and it weighs approximately 600 kg, which makes it quite unsuitable for measurements in the field.
1.1 Spectral analysis of visible light

Fig. 1-2  Optical arrangement of a very high-resolution grating spectrometer. The light enters through slit S, passes through window W, is collimated by mirror C, diffracted successively by gratings G1-G4, and focussed by mirror F onto detector IMD. The optics is mounted inside a pressure chamber PR [3].

The MS257 monochromator/spectrograph [4] with a detector system is one of the most popular laboratory spectrometers. The operation of this instrument is also based on a dispersive element - a multiple grating. The MS257 has a lateral entrance port (incident light) and two exit ports (axial port and lateral port) that can be used for photodiode arrays or CCD detector heads. To facilitate optical alignment, the lateral output port can be used to give an in-line configuration, but use of the axial port is preferred to prevent additional light losses due to the extra turning mirror. Two concave mirrors and one plane diffraction grating are used. The first mirror collimates the incident light and the second one focuses the light dispersed by the grating. The features of this instrument are: total system automation, low stray light, no re-diffracted light, grating positioning accuracy is better than 0.15 nm and wavelength repeatability is 0.06 nm. The wavelength resolution, $\Delta \lambda$, is 0.1 nm for a usable wavelength range between 170 nm to 24 $\mu$m, thus $R > 1700$. The size dimensions are 0.5 m in length, 0.18 m in height and 0.3 m in width (Fig. 1-3).
Fig. 1-3  The MS257 monochromator/spectrograph with some optional accessories [4].

The variable free spectral range Fabry-Perot interferometer [5] is composed of: an interferometer head (see Fig.1-4, weight:11.2 kg, length 0.43 m), a mirror set, a detector and an interferometer controller. A wavelength resolution, $\Delta \lambda$, of 0.006 nm at wavelength $\lambda=532$ nm is demonstrated, thus $R=(532/6).10^3$. A mirror parallelism of $\lambda/100$ can be achieved. The final alignment and spacing are precisely controlled by three piezoelectric transducers with excellent long-term stability and low thermal expansion.

Modern interferometer mirrors can be polished such that deviations from the perfect plane surface are less than 2.5 nm ($\lambda/200$ at $\lambda=500$ nm). Sometimes, instrumental stability of $\lambda/5000$ or better is required over a period of days or weeks. This implies maintaining the parallelism and spacing of the Fabry-Perot plates better than to 0.1 nm (at $\lambda=500$ nm) [5].
1.1 Spectral analysis of visible light

Fig. 1-4  A spectrum analyzer based on a tunable Fabry-Perot interferometer [5].

Finally, as an example of a portable system, a commercially available instrument Mechelle 4500 (grating-based) using a CCD camera for fast parallel readout is presented [6]. The key features are: 200-670 nm wavelength range, for a constant spectral resolution, R=6000. The size is 0.4 m x 0.22 m x 0.2 m and the weight is less than 10 kg (see Fig. 1-5).

Fig. 1-5  A portable spectrograph. The CCD camera is integrated with the mechanics and optics to give a compact design with high stability [6].
1.1.4 Miniaturized spectrometers: advantages, applications

Conventional benchtop spectrometers involve a complex system of lenses and moving parts (or a high-density photodetector array), and are thus bulky and expensive. By applying micromachining techniques one can produce a spectrometer with drastically reduced size and costs (in high-volume production). In addition, a number of these systems can be combined for improved spectral range and/or resolution. Moreover, an integrated optical device has a number of advantages over a conventional optical system, such as a simplified assembly, stable alignment, and compact size. The dimensional advantage of a miniaturized spectrometer is, in many applications, of higher importance than its reduced resolution.

Examination of products in a manufacturing line by means of laboratory analysis is time consuming, increases costs and sometimes stops production. Therefore, it has become important to perform on-line measurements in the process line in order to correct any process problems in real time. A small spectrometer has huge potential to serve the needs of future automated optical inspection systems.

Identification of the composition of gases and liquids, chemical analysis by optical absorption, emission line characterization, colorimetry and biochemical analysis are some of the applications where a miniaturized spectrometer can be useful.

1.1.5 Microspectrometers: state of the art

The feasibility of realizing miniaturized spectrometers in silicon has been demonstrated by applying micromachining technology (both bulk and surface micromachining). Grating and Fabry-Perot based systems were shown [7-16].

A low-cost microspectrometer based on a micromachined grating mounted above a CCD imaging device was presented by Yee et al. [8](see Fig. 1-6). The measured resolution is R=69.8 at 632.8 nm. The spacing between the grating and detector determines the achievable R.

T. Kwa et al. [9] developed an integrated silicon (grating-based) spectrometer (Fig. 1-7). The spectral resolution obtained, R=10 for λ between 380 nm and 730 nm, is lower compared to what can be achieved with an interferometer-based spectrometer. This device is suitable for integration in silicon using wafer-to-wafer bonding of two wafers. The top
wafer contains a transmission grating fabricated using aluminum interconnect lines to disperse light into its spectral components and electronics components to detect these. The bottom wafer is a passive mirror which is used to extend the optical path (determined by the lateral dimensions). The multiple reflections (three times) at the lower wafer are critical, because of scattering due to surface roughness.

Fig. 1-6  *The spectrometer device intended for biochemical analysis based on a grating above a CCD array detector* [8].

Fig. 1-7  *Integrated silicon grating based spectrometer* [9].

Jerman *et al.* [11] (Fig. 1-8) designed a Fabry-Perot microinterferometer for application in optical fiber communications at 1.55 μm in order to
select a particular wavelength channel. Two silicon wafers, of which one contains an electrostatically deflectable membrane, were micromachined and bonded together at the wafer level. Both wafers have highly reflective dielectric mirrors and arrays of metallic electrodes deposited on the interior surfaces. The membrane consists of a silicon mesa surrounded by a corrugated suspension. The space that is enclosed between the two mirrors is the optical resonance cavity. Electrostatic actuation makes it possible to tune between different wavelengths. The reflectivity of the two mirrors and the parallelism between them determine the resolving power of the microinterferometer. Such a device can operate in the near infrared region, at wavelengths greater than 1150 nm, where the silicon substrates are transparent. Voltages between 0-70 V are required to tune the device over the entire operation range, which is about 37 nm.

Some of the limitations of this device are the high voltage its requires for operation, the undesired absorption in the substrate, planarity control, slow response time (inertia of the movable silicon part of the device) and the usable spectral range is limited to the infrared part of the spectrum.

![Diagram](image)

Fig. 1-8  *Fabry-Perot interferometer based spectrometer [11]*.

**J. Patterson et al.** [7] (Fig.1-9) developed a Fabry-Perot optical filter, composed of a movable cantilever top mirror and a silicon-based bottom mirror. The device consists of a thermally grown silicon dioxide (SiO₂) cantilever structure fabricated on a (111) silicon substrate. The structure was etched in a 10% KOH solution to laterally undercut the SiO₂. A 20 nm thick gold layer is sputtered over the entire wafer to form the top mirror as well as the electrical contact. The initial air gap spacing is 6.2 μm. Voltages between 0-53 V are required for an operating range of 60 nm.
1.1 Spectral analysis of visible light

Finally, Raley et al. [9] (Fig. 1-10) developed a microinterferometer in bulk micromachining for the visible part of the spectrum. Like the design of Jerman et al., the device has a limited spectral range (about 50 nm) and requires a high operating voltage (100 V).

![Fabric-Perot optical filter based on a cantilever structure](image)

Fig. 1-9 Fabric-Perot optical filter based on a cantilever structure [7].

![Characteristics of the microinterferometer](image)

Fig. 1-10 Fabric-Perot interferometer-based spectrometer for visible wavelengths [9].
K. Aratani et al. [14] (Fig. 1-11) have demonstrated a surface micromachined movable mirror/membrane (diaphragms 40 μm square) consisting of a silicon nitride layer sandwiched between two polysilicon layers.

Fig. 1-11 Surface micromachined array of Fabry-Perot interferometers [14].

Application of a 5 V bias between the membrane and substrate results in a 20% change in reflectance at λ=780 nm. The relatively low voltage operation is due to the long suspension beams, which also result in a low fill factor when used in an array.

Tran et al. [15] (Fig. 1-12 and Fig. 1-13) developed a tunable Fabry-Perot microinterferometer for infrared light (1.517 μm). The front mirror was made of three and a half periods of quarter wavelength Si-SiO₂ layers and is located at the center of the membrane released from the substrate. The back mirror is similar to the front mirror and was deposited directly on a double-sided polished silicon substrate. Underneath the back mirror an aluminum layer was evaporated, which defines the aperture of the device and serves as the bottom electrode. The membrane was made of a 1.3 μm thick sputtered TiW layer sandwiched between a 1.1 μm CVD oxide and a 1.36 μm Si/SiO₂ Bragg mirror. The length of the optical cavity is about 3 μm. The holes shown in Fig. 1-13 were etched to facilitate the release etching of the sacrificial layer between the substrate and membrane and to relieve material stress. Polyimide was used as the sacrificial layer. The complete device requires 19 layers. A full operating range of 60 nm requires an operating voltage of 65 V.
1.1 Spectral analysis of visible light

---

Fig. 1-12 Summary of the micromachining process of the IR device. a) shows the schematic cross section of the device before polyimide release etching and b) shows the completed device [15].

---

Fig. 1-13 A Fabry-Perot wavelength-tunable optical filter realized in surface micromachining technology [15].
Introduction

Surface micromachining has a demonstrated capability for the fabrication of small-sized structures and, consequently, a huge potential for manufacturing arrays. Moreover, small air gaps of controllable thickness can be fabricated and therefore relatively small voltages are required to adjust the cavity height with great precision. Squeeze film damping could affect the response time of the device, but the hole pattern required for sacrificial release decreases this effect.

1.2 Motivation and objectives

Among many other applications, micromachining in silicon has created new opportunities for optical and optoelectronic systems. The main potential lies in the miniaturization of optical systems and in the extension of their functionality. This is mainly due to the fact that movable micromechanical structures, microactuators and microoptical elements can be monolithically integrated on the same substrate. Standard processing technologies for the fabrication of VLSI ICs (very large-scale integration of integrated circuits) are merged with sensor processing [7]. The general requirement for such an on-chip co-integrated system is that sensor processing does not interfere with proper functioning of already integrated microelectronic devices and vice-versa.

The usefulness of microstructures is sometimes disputed, because of the lack of applications. The most severe limitation of microactuator is the fact that it cannot deliver macroforces or macrotorques to act on the macroscopic world. However, movable structures are attractive for optical applications. Small mechanical displacement can be sufficient to interrupt a light beam without delivering force. Compared with macroscale opto-mechanical devices, the micro-optomechanical devices are simpler, smaller, lighter, faster (higher resonance frequencies), more robust and less expensive. Also, the compatibility with VLSI technology makes it possible to integrate micro-optomechanical systems with driving circuits or signal-conditioning circuits on the same chip, thus improving the functionality of the devices.

Therefore, the main goal of the research described in this thesis is to realize an integrated microspectrometer in silicon with high optical quality and acceptable resolution (for operation in the visible part of spectrum).
1.2 Motivation and objectives

Two concepts were investigated: scanning-type and array-type microspectrometer, both based on a Fabry-Perot interference filter. After a careful comparative study of both approaches, an optical microsystem based on an array of Fabry-Perot etalons was implemented in CMOS technology.

1.2.1 Scanning-type microspectrometer

The structure of the scanning-type microspectrometer investigated in this thesis is schematically shown in Fig. 1-14. It is a bulk-micromachined Fabry-Perot microinterferometer composed of two bonded wafers. This micro-mechanical device features simple fabrication, allows operation in the visible spectral range [16] and is integrable with photodetectors and electronics in silicon. Electrostatic actuation is necessary to control cavity spacing and mirror parallelism.

![Diagram of a scanning-type microspectrometer](Image)

Fig. 1-14 A tunable bulk-micromachined Fabry-Perot microinterferometer for the visible spectral range [16].
1.2.2 Array-type microspectrometer

The array-type microspectrometer is based on an array of Fabry-Perot optical thin-film filters (etalons) of different resonance width and is shown in Fig. 1-15. 16-channel microspectrometers were fabricated in both standard bipolar and CMOS technology. Each of the channels is sensitive in a single peak. This single-chip Fabry-Perot array-type microspectrometer suitable for the visible spectral range does not require electrostatic actuation [17].

The impinging spectrum is filtered in the Fabry-Perot resonator and the selected spectral component is measured in transmission by means of an underlying integrated photodiode array [18].

Fig. 1-15 An array of Fabry-Perot etalons.

1.2.3 A CMOS optical microsystem

A CMOS optical microsystem was designed which includes: a photodiode for each etalon, analog readout-electronics for each photodiode, a multiplexer with decoder for the photodiodes array, a light-to-frequency converter and a local/external sensor bus interface (see Fig. 1-16) [19] [20]. We fabricated the complete chip using a standard CMOS process followed by a post-processing module for deposition of the Fabry-Perot filters.
1.3 Organization of the thesis

This chapter introduces the subject of optical spectral analysis and presents the motivation and objectives of this thesis. Chapter 2 describes the basic optical theory associated with the Fabry-Perot optical resonance cavity. Moreover, the basic operational limits of a Fabry-Perot based microspectrometer are analyzed. Chapter 3 concerns with the design, fabrication and experimental results of the tunable bulk-micromachined Fabry-Perot microinterferometer (scanning-type device). It also presents the electro-mechanical theory describing device operation. Chapter 4 presents a microspectrometer based on an array of Fabry-Perot etalons and respective photodetectors (array-type device). Chapter 5 presents a fully integrated optical microsystem for the spectral analysis of visible light, based on an array of Fabry-Perot etalons. The on-chip co-integrated microsystem, readout electronics and bus interface are described. Chapter 6 compares the two concepts investigated and the respective advantages and drawbacks are discussed. Finally, we draw conclusions and give recommendations for future work.
References


1.3 Organization of the thesis


2.1 Introduction

This chapter presents the theory required to describe the operation of an optical spectrometer based on a plane Fabry-Perot resonator. First, the basic terms (e.g. light, visible light, wavelength, spectrometer, etc.) are defined and the field of optical spectral analysis is introduced. Subsequently, the general building blocks of a spectrometer, with emphasis on a wavelength-selecting element, are described and some additional spectroscopic terms are defined.

The general electromagnetic theory of light (Maxwell equations) is the base of the theoretical analysis. First, an interaction of a monochromatic wave of both normal and oblique incidence with a simple boundary is analyzed. The results are then used to derive general equations for reflectance and transmittance of a thin film deposited on a substrate. Finally, the matrix method is introduced, which makes it possible to derive the expression describing the interaction between a monochromatic wave and an assembly of thin films. In addition, the notions of equivalent layer, optical admittance, absentee layer and quarter-wave optical thickness are presented.
Optical theory of a spectrometer

The subsequent part of this chapter concentrates on a plane Fabry-Perot resonator. Basic properties are derived from a highly idealized model (the concept of a linear system with feedback is used). Furthermore, some phenomena resulting from the resonator non-idealities are analyzed. The spectral transmission, FWHM, finesse, resolving power and free spectral range are derived. The study takes into account the losses in the mirrors and cavity medium due to absorption or scattering. Also, the effect of non-perpendicular incidence, the finite size of the mirrors and the non-parallelism between them are treated. Two possible mirror types (dielectric and metallic) are compared. Finally, a Fabry-Perot based spectrometer is outlined. Two possible types are investigated: scanning-type and array-type. At the end, the signal processing required for spectrum reconstruction is discussed.

2.2 Basic terms

Light, visible light and wavelength

Light is an electromagnetic phenomenon described by the same theoretical principles that govern all forms of electromagnetic radiation. A complete mathematical description of the electromagnetic character of light was made by James Maxwell (1831-1879). The corpuscular nature of light (photons) is complementary to the classical wave description of light and manifests itself whenever light is emitted or absorbed. However, for description of light propagation, the electromagnetic wave theory is more suitable. Light propagates in the form of waves. In free space, light waves travel with a constant speed of $c_0 \approx 3 \times 10^8$ m s$^{-1}$, independent of the movement of the system considered. In other media the speed is different and depends on the frequency of the radiation. Light is generally described in terms of wavelength. Wavelength is the distance between any two successive points of a periodic wave for which the oscillation has the same phase. Light occupies a band of the electromagnetic spectrum that extends from the infrared through the visible to the ultraviolet (Fig. 2-1). The range of optical wavelengths stretches from 10 nm to 1 mm. This range includes the visible part of the spectrum (390 nm to 760 nm) perceivable by the human eye.
2.2 Basic terms

Frequency

![Frequency spectrum diagram]

Wavelength

Light

Fig. 2-1  *The electromagnetic spectrum [1].*

**Spectroscopy**

Spectroscopy is the study of the absorption and emission of light and other radiation, as related to the wavelength of the radiation. This science deals with the sources, measurements, analysis and use of spectra. With spectroscopic methods, it is possible to analyze the composition of a small amount of material with an accuracy and speed that cannot be achieved by chemical techniques.

**Quantities and units in Radiometry and Spectrometry**

Radiometry is the science and technology involved in the measurement of electromagnetic radiant energy (or simply, the measurement of optical radiation). Radiometric quantities used are [2]:

- **radiant energy**: the total energy emitted from a radiating source (in J);
- **radiant power or flux**: the radiant energy per unit time (in Js⁻¹ or W);
- **irradiance**: the total radiant flux emitted by a source divided by the surface area of that element (Wm⁻²);
- **radiant intensity**: the total radiant flux emitted by a source per unit solid angle in a given direction (in Wsr⁻¹).
Optical theory of a spectrometer

- *radiance*: the radiant intensity of a source divided by the area of the source (in W sr\(^{-1}\)m\(^{-2}\)).

When radiant energy, or any related quantity, is measured in terms of its monochromatic components, it becomes a function of wavelength. Therefore, the designations for these quantities must be preceded by the adjective *spectral*, as in spectral irradiance. The symbol itself, for each quantity, is followed by the symbol for wavelength, (\(\lambda\)).

**Optical spectrometer**

For an analysis of a spectrum usually four components are required: a source of light, a dispersive element, a detector and a recorder. The dispersive element is used to transmit different wavelengths, the detector to sense where light appears after dispersion, and a recorder to make a spectrogram: a temporary or permanent copy of the spectrum. An instrument intended to accept light, to separate it into its component wavelengths and to detect the spectrum is called an **optical spectrometer**.

A **spectrograph** is a spectrometer that also records spectra. A **spectroscope** is a combination of disperser and associated optics for visual observation of spectra. A monochromator consists of a disperser and the associated optical components arranged to accept light from an external source and to transmit only a small range of wavelengths.

A **spectrophotometer** is a spectrometer used to record quantitatively the amount of light emitted or absorbed by a sample of material according to the wavelength.

### 2.3 Optical spectral analysis

The identification of the spectral components present in an incident electromagnetic radiation through a measurement of the intensity of every spectral component is called spectral analysis. This operation determines the intensity as a function of the wavelength, \(I=f(\lambda)\), for a specific spectrum.

If the light wave description is known as a function of time (i.e. \(U=f(t)\)), it would be possible to use only mathematical methods to perform decomposition into the spectral domain, i.e. \(U=f(\lambda)\). However, currently there are no methods available which allow direct measurement of light
wave as a function of time. This is due to the extremely high frequencies associated with visible light. Nevertheless, the development of mathematical tools (e.g. Fourier analysis), which facilitates the description of periodic functions, are very useful and helped the foundation of the modern approach to physical optics. Furthermore, a new field of optical signal processing based on modern mathematical theories and adaptive optical elements has emerged [1].

**Fourier analysis**

The Fourier theory states that a Fourier series, a sum of sinusoidal functions, can be used to describe any periodic function and the Fourier transform, an integral transform, can be used to describe non-periodic functions. Moreover, the Fourier theory allows the representation of a function in terms of its frequency or in its temporal characteristics and it using these, one can easily switch from one representation to the other. This theory can be used to describe an arbitrary wavefront in terms of combinations of plane waves. The representation in a Fourier series of a periodic function \( f(t) \), i.e. \( f(t) = f(t+T) \), where \( T = 2\pi/\omega \), is:

\[
 f(t) = \frac{a_0}{2} + \sum_{i=1}^{\infty} a_i \cos(i\omega t) + \sum_{i=1}^{\infty} b_i \sin(i\omega t) \quad (2-1)
\]

Each harmonic \( i \) of the fundamental frequency \( \omega \) is multiplied by a coefficient \( a_i \) (expansion in terms of cosine function) or \( b_i \) (expansion in terms of sine function).

A non-periodic function can be represented by the integral (Fourier transform):

\[
 F\{f(t)\} = F(\omega) = \int_{-\infty}^{\infty} f(\tau) e^{-i\omega \tau} d\tau , \quad (2-2)
\]

which transforms \( f(t) \) from a temporal representation in to the frequency representation \( F(\omega) \). The inverse transform can also be performed:

\[
 F^{-1}\{F(\omega)\} = f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega \quad (2-3)
\]
Optical theory of a spectrometer

Sometimes it is necessary to compare functions (e.g. interference of two waves). The correlation function is a useful integral for comparing the similarity between two functions:

\[ h(\tau) = a(t) \otimes b(t) = \int_{-\infty}^{\infty} a(t)b(t - \tau)dt \]  

(2-4)

It can be thought of as the calculation of the area of overlap of two functions as one of the functions slides over the other.

Another very useful integral in linear systems is the convolution. This smoothing operation can round any sharp peaks. The convolution integral of \( a(t) \) and \( b(t) \) is defined as:

\[ g(\tau) = a(t) * b(t) = \int_{-\infty}^{\infty} a(t)b(\tau - t)dt \]  

(2-5)

2.4 Spectrometer description

A spectrometer identifies the wavelength of the spectral components present in the incident radiation and measures the intensities. Its principal components are: a wavelength-selecting element (WSE) that selects a narrow spectral band and a detector that detects light passing through its filter [3].

To optimize light input into a spectrometer it is often advisable to place some optical elements, to match the source with the instrument. A general spectrometer system consists of the following components: input optics (an entrance slit and a collimating lens), a WSE, output optics (a focusing lens) and a photodetector (see Fig.2-2).
2.4 Spectrometer description

![Diagram of spectrometer components: entrance slit, collimating lens, wavelength selecting element, focusing lens, photodetector.](image)

**Fig. 2-2  General construction of a spectrometer.**

### 2.4.1 Entrance slit and collimating optics

The task of these elements is to match the light source to the input slit of the instrument. The entrance slit is illuminated by an external source whose spectrum is to be analyzed. The slit defines the external source, to make sure that the measurement only depends on its wavelength and to eliminate the influence of wavelength-independent variables, such as the shape of the source, the divergence of the light beam, and the position of the source relative to the optical axis in the spectrometer [4] [5].

The collimating lens transforms the divergent light beam coming from the entrance slit into a parallel one. Together, the entrance slit and the collimating lens form a collimator which is to produce a parallel beam. Obviously, the collimator becomes superfluous when the light source itself already produces a beam of high parallelism. This condition is satisfied by lasers and sources that can be considered to be at infinite distance [4] [5]. Sometimes, the collimation is combined with operation of the WSE. To a certain degree this is true for a Fabry-Perot device, the filtering action of which also causes a parallel beam. However, is not the case for a grating or prism element. Therefore, the collimator mainly serves to maximize the throughput of the illumination in the focal plane if the WSE is a Fabry-Perot device.
2.4.2 Wavelength-selecting element

The principal part of the spectrometer is the WSE. Basically three types of dispersive elements are used: prisms, gratings or interferometers. These WSEs are based on different methods for dispersing spectra: refraction, diffraction and interference, respectively.

Prisms

Historically, glass prisms were the first to be used to break up or disperse light into its components (colors). The path of a light ray bends (refracts) when it passes from one transparent medium to another, e.g. from air to glass. Different colors (wavelengths) of light are bent through different angles; hence a ray leaves a prism in a direction that depends on its color. The relationship between the wavelength and the refraction angle is non-linear.

![Diagram of refraction through a prism](image)

**Fig. 2-3** Refraction of light through a prism having index \( n_2 \), immersed in a medium having refractive index \( n_1 \). The angles \( i \) and \( r \) are the angles of incidence and refraction, respectively.

Gratings

Diffraction gratings are composed of closely spaced transmitting slits on a flat surface (transmission gratings) or alternate reflecting and non-reflecting grooves on a flat or curved surface (reflection gratings). A typical grating might be ruled on an aluminum coating on a piece of glass and consists of equally spaced parallel lines, from as few as 100 lines per millimeter of width up to as many as 6000 lines per millimeter. Holographic gratings are chosen when groove density should be at least 1200 gr/mm or more for use in near UV, visible and near IR. Ruled
2.4 Spectrometer description

gratings are applied in IR above 1.2 µm with low groove density, e.g., less than 600 gr/mm.

![Diffraction grating diagram](image)

Fig. 2-4  *Basic arrangement of a diffraction grating for visible light.*

**Interferometers**

A third class of elements for dispersing spectra consists of interferometers. These divide a wave front by semi-transparent surfaces, allow the beams thus produced to travel along different paths, and then recombine the beams. These beams interfere constructively only if their path difference is 1, 2, 3,... times the wavelength. After redirection and recombination for both of which the same or another beam splitter is used, the total intensity can be detected. This device acts as a filter that transmits certain wavelengths and reflects the others back to the light source. Fig. 2-5 illustrates four very important types of interferometers: a) Mach-Zender, b) Michelson, c) Sagnac and d) Fabry-Perot cavity [6]. The first three interferometers have complex structures (more than one optical path).

A more compelling reason for using interference filters is the greatly increased potential grasp of energy over the others dispersive elements (diffraction grating and prism), [7]. In the visible region the ratio of the energy collected by a Fabry-Perot filter and a diffraction grating is between 76 and 760 [7] (where we assume that each element uses an ideal system to make maximum use of its energy-gathering power). Therefore, a smaller Fabry-Perot filter has a very significant increase in energy grasp over a diffraction grating.
2.4.3 Output optics

The task of this component is to match the output of the WSE to the photodetector. Commonly, a focusing lens is used, which projects the images at the focal plane of the photodetector. The main purpose is to maximize optical throughput and thus the signal supplied by the detector at a given optical intensity.

2.4.4 Photodetectors

The photodetector measures the intensity of the spectral components in the transmitted light with respect to a designated wavelength. Depending on the spectral range that is analyzed and the signal intensity, different types of detectors can be used: e.g. photomultipliers, image intensifiers, bolometers and photoconductive cells. For the visible light the most common detectors are: photographic films, photodiodes and Charge Coupled Devices (CCDs). In the past a strip of photographic film or a photographic plate recorded the pattern of spectral lines (it also served to store the data). Presently, photodiode readouts are employed, using a current-voltage amplifier followed by an analog-digital converter. This electronic form of data-acquisition enables digital recording and storage of the data.
2.4 Spectrometer description

2.4.5 Scanning-type vs. array-type approach

When a spectroscopic measurement is performed, each of the spectral components must be measured individually. Two basic approaches can be used:

a) sequential measurement using scanning techniques.

b) parallel measurement based on an array of detectors.

Scanning-type approach

In the scanning-type spectrometer, measurement is performed in a time sequence. Two configurations are possible and both are used in practice. When the relative position of the input light beam is considered to be fixed, then, at a certain position of the WSE relative to the detector, the first spectral component is measured. Subsequently, the relative position of the WSE with respect to the detector is slightly changed and the next spectral components are measured. Rotation of the WSE over a well-defined angle enables scanning and recording of the entire spectrum. Alternatively the photodetector can be moved to scan across all spectral components (see Fig.2-6).

![Diagram showing scanning type: a) movable WSE, b) movable detector.](image-url)
Array-type approach

In the array-type approach, the measurement is performed in parallel. This means that a large array of detectors is required. Each of them measures one spectral component. The number of detectors or channels used can be a limiting factor for the resolving power. In contrast to the previous approach no mechanical scanning is required and, therefore, array-type systems are suitable for high-speed measurements. In practice, one WSE and an array of detectors is used in commercially available systems (see Fig. 2-7). As will be shown in chapter 4, combining an array of WSEs and detector is a very interesting option in a microspectrometer.

![Array-type diagram](image)

**Fig. 2-7** Array type: a) array of WSEs and detectors. b) array of detectors.

2.4.6 Spectroscopic terms

Coherent waves

Non-coherent light arises because there are unpredictable fluctuations in the light source or in the medium through which light propagates. Also, non-coherent light can result from scattering on rough surfaces. An example of coherent light is the monochromatic wave \( u(r,t) = \text{Re}\{U(r)\exp(i2\pi vt)\} \), for which the complex amplitude \( U(r) \) is a deterministic complex function, e.g., \( U(r) = A \exp(-ikr)/r \) in the case of a spherical wave [1]. The dependence of the wave function on time and position is perfectly periodic and predictable. On the other hand, for
non-coherent light, the dependence of the wave function on time and position cannot be totally predicted [1].

If interference is to occur between waves, these must be coherent. Practical waves are only coherent during a certain period. As light travels a certain distance during that time interval, practical specifications are expressed in the **coherence length**. The coherence length (in meters) of an optical source is determined by \( L_c = c/\Delta v \), where \( c \) is the speed of light in meters per second and \( \Delta v \) is the linewidth of the source in hertz.

Light waves in the useful range of a given filter typically have a coherence length that is larger than the optical thickness of the filter, and the various partial waves can interfere with each other. In section 2.4.2 we recombined two waves and assumed that the two waves are coherent. We could do this because the coherence length of most light is several wavelengths long. For a continuous wave (a laser beam), the coherence length is several meters. On the other hand white light from a thermal source, which is composed of the superposition of emissions from a very large number of atoms radiating independently and at different frequencies and phases, does not have this coherence. In a typical thin-film configuration, where the total thickness of the stack is usually a small number of waves thick, all the waves reflected from the interfaces can be added coherently, i.e., the phases of the waves have to be taken into account when these amplitudes are added [8].

**Spectral responsivity**

A light sensor is defined by its wavelength-dependent response, which is called its spectral responsivity, \( R(\lambda) \). This generally refers to the electrical signal generated by a photodetector \( s(\lambda) \) when irradiated with a known radiant flux of a specific wavelength \( \Phi(\lambda) \) and is determined by means of the relationship:

\[
R(\lambda) = \frac{s(\lambda)}{\Phi(\lambda)}
\]  

(2-6)

The output signal of the detector can be in amperes, volts, counts/seconds, etc.

**Resolving power**

The power of an optical instrument to resolve adjacent spectral components is conveniently examined by considering its operation on two
narrow spectral components of equal intensity (see Fig.2-8). The minimum separation, $\Delta \lambda$, when the components are said to be just resolved, is called the resolving power and is given by:

$$R = \frac{\lambda}{\Delta \lambda}$$  

(2-7)

This criterion for resolution is due to Lord Rayleigh and was first used for prism and grating-based instruments. Lord Rayleigh suggested that two components of equal intensity may be considered to be resolved when the peak of each component coincides with the first diffraction minimum of the other. The ratio from the intensity at the mid-point of the resultant spectrum to the peak intensity is then $8/\pi^2=0.811$.

![Fig. 2-8 An illustration of the Raleigh criterion for resolution of two spectral components.](image)

Stray light

WSEs are used to isolate a particular spectral component from all other wavelengths, but any practical WSE will transmit residual out-of-band wavelengths or the reflected components reflect back via the instrument housing. This is known as stray light. Generally, the collimator is used to minimize stray light and the inner walls of the instrument are designed anti-reflective.

Spectrometer light gathering or étendue

Fig.2-9 shows a light source, an optical path, a collecting lens and a spectral filter. The effective areas of the different elements are indicated
together with the solid angles that they subtend [6]. The spectral radiance of the source is given by the function \( S(\lambda) \) - that is, power per unit bandwidth, per unit area of source, per unit solid angle radiated isotropically. The peak transmission of the spectral filter at the centre of its band pass is \( T_\lambda \).

The light from area A (light source) is collected by the lens, transmitted through aperture D and incident on the entry of the optical filter. The light flux passing through the system, expressed as radiant power per unit bandwidth is: \( \text{RP}(\lambda) = S(\lambda)A\Omega T_\lambda \). \( A\Omega \) gives the light-gathering power of the instrument and has come to be known as the étendue, \( U \). In case of an air medium the condition to be met by the collection lens and the input aperture is respectively: \( A\Omega = L\Omega_1 = D\Omega_2 \).

![Diagram showing light gathering of a spectroscopic instrument showing source, collecting lens and spectral filter](image)

**Fig. 2-9** *Light gathering of a spectroscopic instrument showing source, collecting lens and spectral filter [6].*

The light-gathering power (or étendue) of a Fabry-Perot interferometer used in a basic geometry of a spectrometer is shown in Fig. 2-10. The étendue is the product of the source area observed and the collection solid angle. The detector has an aperture diameter \( a \), and an usable plate diameter \( D \) defines the etalon aperture. When correctly focused, only light from the source area defined by the diameter \( (af_1/f) \) is detected. The étendue is independent of the lens geometry and can only be increased by making \( D \) larger; i.e., by employing larger etalon plates. Obviously, the collected light is also proportional to the source brightness [6].
Optical theory of a spectrometer

Fig. 2-10 The geometry of a basic Fabry-Perot spectrometer [6].

Extinction or contrast factor

In the study of emission or relatively narrow line spectra a useful parameter of a spectrometer is the extinction or contrast factor [6]. This is defined (measured with a quasi-monochromatic source) as the ratio of the intensity at the peak (I_{\text{max}}) to the baseline or wing of the curve - I_{\text{min}} (see Fig. 2-11). For a Fabry-Perot interferometer I_{\text{min}} would be the intensity level half-way between successive orders. Also, using the reflectance value, we can calculate R, the contrast factor:

\[ CF = \frac{I_{\text{max}}}{I_{\text{min}}} = \left(\frac{1 + R}{1 - R}\right)^2 \]  \hspace{1cm} (2-8)

Fig. 2-11 Transmission peaks are equally spaced as a function of frequency.
2.5 Basic theory

Some basic theory is necessary in order to make calculations on the properties of multilayer thin-film coatings. The following analysis is a condensed discussion. For a full rigorous treatment, of the electromagnetic field equations, the reader is referred to [1] and [7]. To describe light in free space, six scalar functions of position and time are required. The Maxwell equations combine and relate all these functions in a set of partial differential equations.

2.5.1 Maxwell equations: free space

An electromagnetic field is described by two related vector fields: the electric field \( \mathbf{E}(\mathbf{r},t) \) and the magnetic field \( \mathbf{H}(\mathbf{r},t) \). Both are vector functions of position and time and they satisfy the Maxwell equations in free space [1]:

\[
\nabla \times \mathbf{H} = \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{2-9}
\]

\[
\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{2-10}
\]

\[
\nabla \cdot \mathbf{E} = 0 \tag{2-11}
\]

\[
\nabla \cdot \mathbf{H} = 0 \tag{2-12}
\]

where the constants \( \varepsilon_0 = 8.85 \times 10^{-12} \text{ F m}^{-1} \) and \( \mu_0 = 4\pi \times 10^{-7} \text{ H m}^{-1} \) are, respectively, the electric permittivity and the magnetic permeability of free space; and \( \nabla \) and \( \nabla \times \) are the divergence and the curl operations. A necessary condition for \( \mathbf{E}(\mathbf{r},t) \) and \( \mathbf{H}(\mathbf{r},t) \) to satisfy the Maxwell conditions is that each of their components satisfies the wave equation:

\[
\nabla^2 u - \frac{\varepsilon_0}{c_0^2} \frac{\partial^2 u}{\partial t^2} = 0 \tag{2-13}
\]

where \( c_0 = 1/(\varepsilon_0 \mu_0)^{1/2} \) = 3x10^8 m s^{-1} is the speed of light in free space, while the scalar function \( u \) represents any of the three components of \( \mathbf{E}(\mathbf{r},t) \) and \( \mathbf{H}(\mathbf{r},t) \), \( (E_x, E_y, E_z, H_x, H_y, H_z) \).
2.5.2 Maxwell equations: isotropic medium, no charges

In a isotropic medium without free electric charges or currents, two more vectors field need to be defined - the electric flux density $\mathbf{D}(r,t)$ and the magnetic flux density $\mathbf{B}(r,t)$. The Maxwell equations relate the four fields $\mathbf{E}(r,t)$, $\mathbf{H}(r,t)$, $\mathbf{D}(r,t)$ and $\mathbf{B}(r,t)$, by:

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}$$  \hspace{1cm} (2-14)

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$  \hspace{1cm} (2-15)

$$\nabla \cdot \mathbf{D} = 0$$  \hspace{1cm} (2-16)

$$\nabla \cdot \mathbf{B} = 0$$  \hspace{1cm} (2-17)

2.5.3 Monochromatic electromagnetic waves

In a linear and homogeneous medium one refers to the electromagnetic wave as being monochromatic if all components of the electric and magnetic fields are harmonic functions of time of the same frequency $\omega$ [1]. These components can be expressed in terms of their complex amplitudes:

$$\mathbf{E}(r,t) = \text{Re}\{\mathbf{E}(r)\exp(i\omega t)\}$$  \hspace{1cm} (2-18)

$$\mathbf{H}(r,t) = \text{Re}\{\mathbf{H}(r)\exp(i\omega t)\}$$  \hspace{1cm} (2-19)

where $\mathbf{E}(r)$ and $\mathbf{H}(r)$ are the complex amplitudes of the electric and magnetic fields, respectively (or only $\mathbf{E}$ and $\mathbf{H}$), and $\omega = 2\pi v$ is the angular frequency and $v$ is the frequency.

The Maxwell equations for monochromatic light in a linear, homogeneous, isotropic, nondispersive, source-free medium (with medium equations: $\mathbf{D} = \varepsilon \mathbf{E}$ and $\mathbf{B} = \mu_0 \mathbf{E}$, where $\mathbf{D}$ and $\mathbf{B}$ are the complex amplitudes of
electric flux density and magnetic flux density respectively) and substituting \((\partial/\partial t) = i\omega\):

\[
\nabla \times \mathbf{H} = i\omega \varepsilon \mathbf{E} \tag{2-20}
\]

\[
\nabla \times \mathbf{E} = -i\omega \mu_0 \mathbf{H} \tag{2-21}
\]

\[
\nabla \cdot \mathbf{E} = 0 \tag{2-22}
\]

\[
\nabla \cdot \mathbf{H} = 0 \tag{2-23}
\]

Since the components of \(\mathbf{E}(r,t)\) and \(\mathbf{H}(r,t)\) satisfy the wave equation (2-13) (with light waves travel in a medium of refractive index \(n\), \(\nu_m = c_0/n\) and \(n = (\varepsilon/\varepsilon_0)^{1/2}\)), the components of \(\mathbf{E}\) and \(\mathbf{H}\) must satisfy the Helmholtz equation:

\[
\nabla^2 U + k^2 U = 0, \text{ with } k = \omega(\varepsilon/\mu_0)^{1/2} = nk_0 \tag{2-24}
\]

where the scalar function \(U = U(r)\) represents any of the six components of the vectors \(\mathbf{E}\) and \(\mathbf{H}\), and \(k_0 = \omega/c_0\).

### 2.5.4 Optical admittance

The complex refractive index of a material, \(N\), has two components: \(n\), the refractive index and \(k\), the extinction coefficient:

\[
N = n - ki \tag{2-25}
\]

\[
n^2 - k^2 = \varepsilon_r/\mu_r \text{ and } 2nk = \mu_r \sigma/\omega \varepsilon_0 \tag{2-26}
\]

where \(\varepsilon_r\) and \(\mu_r\) are the permittivity and permeability of the material, \(\sigma\) the dielectric conductivity of the material, \(\varepsilon_0\) the permittivity of free space and \(\omega\) the angular frequency of the electromagnetic wave.

The optical admittance of a medium is \(y = N/c_0\mu\), with \(\mu = \mu_r \mu_0\) as the permeability of the medium. In free space, the optical admittance is given by \(y_{fs} = (\varepsilon_0/\mu_0)^{1/2}\). At optical wavelengths \(\mu_r\) is unity, so that we can write: \(y = Ny_{fs}\). The wave propagation direction is given by the unit vector.
Optical theory of a spectrometer

\[ s = \alpha i + \beta j + \gamma k, \] where \( i, j, \) and \( k \) are unit vectors along the \( x, y, \) and \( z \) axes, respectively. From this we can write:

\[ \frac{N}{v_m\mu} (s \times E) = H \quad (2-27) \]

and

\[ H = y(s \times E) = Ny_{fs}(s \times E) \quad (2-28) \]

**2.5.5 The Poynting vector**

The flow of electromagnetic power is governed by the vector \( P = ExH \), known as the Poynting vector. The direction of power flow is along the direction of the Poynting vector, i.e., is orthogonal to both \( E \) and \( H \). The value of the Poynting vector oscillates at twice the frequency of the wave. This is defined as the intensity. For a harmonic wave we can derive an expression for the intensity:

\[ I = 0.5\text{Re}(ExH^*) \quad (2-29) \]

where * denotes complex conjugate. The scalar value of \( I \) is \((E \) and \( H \) are perpendicular):

\[ I = 0.5\text{Re}(EH) \quad (2-30) \]

With equation (2-28) results for the intensity vector:

\[ I = 0.5\text{Re}(yEE^*s) = 0.5ny_{fs}EE^*s \quad (2-31) \]

**2.6 Interaction of a electromagnetic wave with a multilayer stack**

In this section we derive the equations which describe the interaction of a monochromatic wave with a multilayer stack in a linear and homogeneous medium. This assembly of thin films can be used as a WSE. Thin-film interference coatings typically rely on the difference in refractive index of two or more materials used to produce interference effects within a multilayer structure. A simple boundary and a thin film deposited on a
2.6 Interaction of a electromagnetic wave with a multilayer stack

substrate are first studied in order to analyze the optical behavior of a multilayer structure. The complete basic thin-film optical theory is available in literature [7], [8].

2.6.1 Ideal boundaries

When a monochromatic wave interacts with a linear and homogeneous thin film (perfect surfaces, no diffuse light at either side, see Fig.2-12), three processes occur: transmission, absorption and reflection.

![Diagram showing transmission, absorption, reflection](image)

Fig. 2-12 Ideal case where a monochromatic wave is incident on a linear, homogeneous material giving rise to: transmission, absorption and reflection.

Transmittance

Transmission is the term used to describe the process by which incident radiant flux leaves a surface or medium from a side other than the incident side, usually the opposite one. The spectral transmittance $\tau(\lambda)$ of a medium is the ratio of the transmitted spectral flux $\Phi_{\lambda,t}$ to the incident spectral flux $\Phi_{\lambda,i}$, or [9]

$$\tau(\lambda) = \frac{\Phi_{\lambda,t}}{\Phi_{\lambda,i}}, \quad (2-32)$$
the transmittance $\tau$ is the ratio of the transmitted flux $\Phi_t$ to the incident flux $\Phi_i$, or [9]

$$\tau = \frac{\int_0^\infty \tau(\lambda) \Phi_{\lambda,i} d\lambda}{\int_0^\infty \Phi_{\lambda,i} d\lambda}$$  \hspace{1cm} (2-33)

**Absorptance**

Absorption is the process by which incident radiant flux is converted to another form of energy, usually heat, inside the element. Absorptance is the fraction of incident flux that is absorbed. The absorptance $\alpha$ of an element is defined by $\alpha = \Phi_{\lambda,a} / \Phi_{\lambda,i}$. Similarly, the spectral absorptance $\alpha(\lambda)$ is the ratio of spectral power absorbed $\Phi_{\lambda,a}$ to the incident spectral power $\Phi_{\lambda,i}$, or [9]

$$\alpha = \frac{\int_0^\infty \alpha(\lambda) \Phi_{\lambda,i} d\lambda}{\int_0^\infty \Phi_{\lambda,i} d\lambda}$$  \hspace{1cm} (2-34)

**Reflectance**

Reflection is the process where a fraction of the radiant flux incident on a surface is returned to the same hemisphere. The reflection can be specular (in the mirror direction), diffuse (scattered into the entire hemisphere), or a combination of both. Reflectance $\rho$ is the ratio of the radiant flux reflected $\Phi_r$ to the incident radiant flux $\Phi_i$, or [9]

$$\rho = \frac{\Phi_r}{\Phi_i}$$  \hspace{1cm} (2-35)
2.6 Interaction of a electromagnetic wave with a multilayer stack

Spectral reflectance at a specified wavelength $\lambda$ is similarly defined as [9]

$$\rho(\lambda) = \frac{\Phi_{kr}}{\Phi_{ki}} \quad (2-36)$$

2.6.2 A simple boundary

In a homogeneous medium, all components of the fields $E$, $H$, $D$ and $B$, are continuous functions of position. A single-boundary, absorption-free medium, is shown in Fig.2-13 (incident medium and exit medium). At the boundary between the two dielectric media and in the absence of free electric charges and currents, the tangential components of $E$ and $H$ must be continuous. The incident beam is split into a reflected beam and a transmitted beam.

Normal incidence

For the study of normal incidence we will consider the incident wave to be a plane-polarized harmonic wave. Fig.2-13 shows the directions of the electric and magnetic vectors for the incidence, transmission and reflectance.

---

Fig. 2-13  
a) Plane wavefront incidence on a single surface.  
b) Normal incidence, $E$ and $H$ directions for incidence, reflection and transmittance [7].
Optical theory of a spectrometer

Applying the boundary conditions and considering only the amplitudes without any phase changes, comes for the electric vector continuous across the boundary:

\[ E_i + E_r = E_i, \]  
(2.37)

the magnetic vector continuous across the boundary uses: \( y_0 = n_0 y_{fs} \) and \( y_1 = n_1 y_{fs} \):

\[ y_0 E_i + y_0 E_r = y_1 E_i \]  
(2.38)

Eliminating \( E_r \), \( y_1 (E_i + E_r) = y_0 (E_i - E_r) \) makes it possible to derive \( E_r/E_i \) and \( E_r/E_l \), also called the amplitude reflection and transmission coefficients. These are denoted by [7]:

\[ \frac{E_r}{E_i} = \rho = \frac{y_0 - y_1}{y_0 + y_1} = \frac{n_0 - n_1}{n_0 + n_1} \]  
(2.39)

\[ \frac{E_l}{E_i} = \tau = \frac{2y_0}{y_0 + y_1} = \frac{2n_0}{n_0 + n_1} \]  
(2.40)

For \( n_0 > n_1 \), there will be no phase shift between incident and reflected beams at the interface, but if \( n_0 < n_1 \), there will be a phase change of \( \pi \) because the value of \( \rho \) is negative. When we use equation (2.31) and \( E_r = \rho E_i \) and \( E_l = \tau E_i \) results in the net intensity [7]:

\[ \text{net intensity} = 0.5yE_iE_i^* (1 - \rho^2) = 0.5yE_iE_i^* (y_1/y_0) \tau^2, \]  
(2.41)

the intensity of the incident beam, the reflected beam and the transmitted beam are: \( I_i = 0.5y_0E_iE_i^* \), \( I_r = \rho^2 I_i \) and \( I_l = (y_1/y_0) \tau^2 I_i \), respectively. The reflectance \( R \) and transmittance \( T \) are:

\[ T = \frac{I_t}{I_i} = \frac{y_1 \tau^2}{y_0} = \frac{4y_0 y_1}{(y_0 + y_1)^2} = \frac{4n_0 n_1}{(n_0 + n_1)^2} \]  
(2.42)

\[ R = \frac{I_r}{I_i} = \rho^2 = \left( \frac{y_0 - y_1}{y_0 + y_1} \right)^2 = \left( \frac{n_0 - n_1}{n_0 + n_1} \right)^2 \]  
(2.43)
2.6 Interaction of a electromagnetic wave with a multilayer stack

Oblique incidence

In case of non-normal incidence the wave is split into two plane-polarized components, one with the electric vector in the plane of incidence, known as p-polarized light (transverse magnetic field, TM) and one with the electric vector normal to the plane of incidence, known as s-polarized light (transverse electric field, TE). The propagation of each of these two waves can be treated independently of the other. Before, for normal incidence we defined the optical admittance, \( y = nyf_s \). We now define a modified optical admittance, \( \eta \), which connects \( E \) and \( H \): \( \eta = H/E \). Consider a homogeneous wave (energy flows normal to the boundaries with \( E \) and \( H \) parallel to them). The optical admittance for p-waves and s-waves, at oblique incidence, are [7]:

\[
\eta_p = \frac{Ny_{fs}}{\cos \theta} \tag{2-44}
\]

\[
\eta_s = Ny_{fs} \cos \theta, \tag{2-45}
\]

and in both cases we can write for \( R \) and \( T \):

\[
T = \frac{4\eta_0 \eta_1}{(\eta_0 + \eta_1)^2} \tag{2-46}
\]

\[
R = \left( \frac{\eta_0 - \eta_1}{\eta_0 + \eta_1} \right)^2 \tag{2-47}
\]

Normal incidence in absorbing media

The incident medium is absorption-free. In case of transmission into another absorption-free, transparent medium. We have [7]:

\[
R = \rho \rho^* = \left( \frac{N_0 - N_1}{N_0 + N_1} \right) \left( \frac{N_0 - N_1}{N_0 + N_1} \right)^* = \left( \frac{y_0 - y_1}{y_0 + y_1} \right) \left( \frac{y_0 - y_1}{y_0 + y_1} \right)^* \tag{2-48}
\]

\[
T = \frac{n_1}{n_0} \tau \tau^* = \frac{4n_0 n_1}{(N_0 + N_1)(N_0 + N_1)^*} = \frac{4(Re y_0)(Re y_1)}{(y_0 + y_1)(y_0 + y_1)^*} \tag{2-49}
\]

where \( N_0 = n_0, N_1 = n_1 - ik_1, y_0 = N_0 y_{fs} \) and \( y_1 = N_1 y_{fs} \)
Oblique incidence in absorbing media

Similarly to absorption-free media we can write for $R$ and $T$ [7]:

$$R = \left( \frac{\eta_0 - \eta_1}{\eta_0 + \eta_1} \right) \left( \frac{\eta_0 - \eta_1}{\eta_0 + \eta_1} \right)^*$$  \hspace{1cm} (2-50)

$$T = \frac{4Re(\eta_0)Re(\eta_1)}{(\eta_0 - \eta_1)(\eta_0 + \eta_1)^*},$$  \hspace{1cm} (2-51)

and the absorption, $A$, is obtained by $A=1-(R+T)$.

2.6.3 Optical properties of a thin film

A structure is composed of a thin film deposited on a substrate as shown in Fig. 2-14. The incident light coming from the incident medium suffers multiple reflections and refractions between the two interfaces separating the three media (incident medium 0, thin film 1, exit medium 2). If the amplitude of the incident wave is $E_0$, then the total reflected amplitude is given by the summation of all the contributions of the separate rays reflected into incident medium [10], [11], [12]:

$$E_{refl} = r_{01}E_0 + r_{12}t_{01}t_{10}E_0e^{2j\delta} + r_{10}(r_{12})^2t_{01}t_{10}E_0e^{4j\delta} + (r_{10})^2(r_{12})^3t_{01}t_{10}E_0e^{6j\delta} + ... = r_{01}E_0 + r_{12}t_{01}t_{10}E_0e^{2j\delta}[1 + r_{10}r_{12}e^{2j\delta} + (r_{10}r_{12})^2e^{4j\delta} + ...]$$

Fig. 2-14 Propagation of an electromagnetic wave through a thin film deposited on a substrate.
The formula for the reflected amplitude of light results in:

\[ \rho = \frac{E_{\text{refl}}}{E_0} = \left( r_{01} + \frac{r_{12} \cdot t_{01} \cdot t_{10} \cdot e^{2j\delta}}{1 - (r_{10} \cdot r_{12} \cdot e^{2j\delta})} \right) \frac{r_{01} + r_{12} \cdot e^{2j\delta}}{1 + r_{10} \cdot r_{12} \cdot e^{2j\delta}}, \]  \hspace{1cm} (2-52)

where

\[ \delta = \frac{2\pi}{\lambda_0} \cdot n_1 \cdot d \cdot \cos \phi_1, \]  \hspace{1cm} (2-53)

is the change phase of the light beams on traversing the film (\(\lambda_0\) is the wavelength of light in vacuum). Similar to the way in which the reflected amplitude is calculated, we calculate the transmitted amplitude:

\[ E_{\text{transm}} = t_{01} t_{12} E_0 e^{j\delta} + r_{10} t_{12} t_{01} t_{12} E_0 e^{2j\delta} + (r_{10} t_{12})^2 t_{01} t_{12} E_0 e^{4j\delta} e^{j\delta} + (r_{10} t_{12})^3 t_{01} t_{12} E_0 e^{6j\delta} e^{j\delta} + \ldots = t_{01} E_0 + r_{10} t_{12} t_{01} t_{12} E_0 e^{2j\delta} (1 + r_{10} t_{12} e^{2j\delta} + (r_{10} t_{12})^2 e^{4j\delta} + \ldots) \]

\[ \tau = \frac{E_{\text{transm}}}{E_0} = \frac{t_{01} \cdot t_{12} \cdot e^{j\delta}}{1 + (r_{10} \cdot r_{12} \cdot e^{2j\delta})}, \]  \hspace{1cm} (2-54)

The reflectivity and transmissivity in all systems are:

\[ R = \rho \rho^* = \rho^2 \text{ and } T = (n_2/n_0) \tau \tau^* = \tau^2 \]  \hspace{1cm} (2-55)

Assuming all materials are ideal, namely, isotropic, homogeneous and non-absorbant, yields [10], [11]:

\[ R = \frac{t_{01}^2 + r_{12}^2 + 2 \cdot r_{01} \cdot r_{12} \cdot \cos 2\delta}{1 + r_{01}^2 \cdot r_{12}^2 + 2 \cdot r_{01} \cdot r_{12} \cdot \cos 2\delta} \]  \hspace{1cm} (2-56)

\[ T = \frac{t_{01}^2 \cdot t_{12}^2}{1 + r_{01}^2 \cdot r_{12}^2 + 2 \cdot r_{01} \cdot r_{12} \cdot \cos 2\delta} \cdot \frac{n_2 \cdot \cos \phi_2}{n_0 \cdot \cos \phi_0}, \]  \hspace{1cm} (2-57)
the thickness of the film is $d$ and when $\delta$ is replaced by $\delta+\pi$, i.e. when $d$ is replaced by $d+\Delta d$, where $\Delta d$:

$$\Delta d = \frac{\lambda_0}{2n_1 \cos \phi_1} \quad (2-58)$$

The reflectivity and transmissivity of dielectric thin films which differ in thickness by an integer multiple of $\lambda_0/(2n_1\cos \phi_1)$ are the same. For the optical thickness of the film $D$ (i.e. $D=n_1d$), the reflection coefficient has a maximum or a minimum, which can be determined from the condition $[10][11]$:

$$\frac{dR}{dD} = 0 \quad (2-59)$$

The condition is fulfilled when $\sin 2\delta = 0$, i.e. when

$$D = \frac{p\lambda_0}{4 \cdot \cos \phi_1}, \text{ with } (p = 0,1,2,...) \quad (2-60)$$

Two solutions can be distinguished:

1) if $p$ is odd, then $\cos 2\delta = -1$ and equation (2-56) yields:

$$R = \left( \frac{r_{01} - r_{12}}{1 - (r_{01} \cdot r_{12})} \right)^2, \quad (2-61)$$

and for normal incidence, one has

$$r_{01} = \left( \frac{n_0 - n_1}{n_0 + n_1} \right) \quad (2-62)$$

$$r_{12} = \left( \frac{n_1 - n_2}{n_1 + n_2} \right), \quad (2-63)$$
2.6 Interaction of a electromagnetic wave with a multilayer stack

finally becomes:

\[
R = \left( \frac{n_0 \cdot n_2 - n_1}{n_0 \cdot n_2 + n_1} \right)^2 \tag{2-64}
\]

2) if p is even, then \( \cos 2\delta = 1 \) and yields:

\[
R = \left( \frac{r_{01} + r_{12}}{1 + (r_{01} \cdot r_{12})} \right)^2 \tag{2-65}
\]

In particular for normal incidence, one obtains:

\[
R = \left( \frac{n_0 - n_2}{n_0 + n_2} \right)^2, \tag{2-66}
\]

and thus the reflectivity is independent of \( n_1 \). In there is oblique incidence, \( n_1 \) should be replaced by \( n_1 \cos \phi_i \) (\( i = 1,2,3 \)) in all the formulas. A film with an optical thickness, \( p\lambda_0/(2 \cos \phi_i) \) with \( p = 1,2,3,... \) has no influence on the intensity of the reflected (or transmitted) light. Such layers are referred to as \( \textit{absentee} \) layers [10] [11] [12].

To determine the nature of these extreme values, we start with \( D = p\lambda_0/(4 \cos \phi_i) \) and \( p = 1,2,3,... \), and when

\[
\left( \frac{d^2R}{dD^2} \right) > 0 \tag{2-67}
\]

\[
\left( \frac{d^2R}{dD^2} \right) < 0, \tag{2-68}
\]

this means \((-1)^p r_{01} r_{12} [1+r_0 r_2 + r_{12}^2 - r_{01}^2 - r_{12}^2] < 0 \) for the first condition, and for the second condition it yields: \((-1)^p r_{01} r_{12} [1+r_0 r_2 + r_{12}^2 - r_{01}^2 - r_{12}^2] > 0 \). For normal incidence one finds [11]:

maximum, if \((-1)^p (n_0 - n_1)(n_1 - n_2) > 0 \)

minimum, if \((-1)^p (n_0 - n_1)(n_1 - n_2) < 0 \).
Optical theory of a spectrometer

Considering the first medium is air \((n_0 = 1)\), we conclude that whether the reflectivity of a film having an optical thickness of any of the values \(\lambda_0/4\), \(3\lambda_0/4\), \(5\lambda_0/4\),...etc, is a maximum or a minimum depends on whether the refractive index of the film is greater or smaller than that of the backing medium (substrate). For a film whose optical thickness has any of the values \(\lambda_0/2\), \(2\lambda_0/2\), \(3\lambda_0/2\),...etc, the opposite is true.

Therefore, a film whose optical thickness is \(\lambda_0/4\) and whose refractive index is low enough may be used as an antireflectant film [11]. Coating the substrate with high refractive index materials allows high reflectivities. The expressions derived can no longer describe a multilayer structure, and a different approach must be chosen.

2.6.4 Optical properties of a multilayer stack

This section presents a general technique to calculate the transmittance and reflectance of a multilayer stack. It is assumed that the structure is composed of an ideal stack of homogeneous, flat and parallel layers [12], [13], [14].

The tangential components of the electromagnetic wave electric field \(E\) and magnetic field strength \(H\) must to be continuous across each boundary of the multilayered structure in order to satisfy the Maxwell equations. The reflectance \((R)\) of a wave propagating from one medium (refractive index, \(n_1\)) to another medium (refractive index \(n_2\)) is defined as the ratio of reflected power to incident power and is given by:

\[
R = \left(\frac{n_1 - n_2}{n_1 + n_2}\right)^2
\] (2-69)

The transmittance at the interface is 1-R. Reflection of all incident waves occurs at the interface between the two media. The multiple stack is composed of \(q\) layers, as shown in Fig.2-15. As will be shown each layer is defined by its characteristic matrix. A matrix multiplication technique (matrix method) is used to calculate the optical characteristics of the multiple film stack. This solution is only valid for the case of normal incidence.
2.6 Interaction of a electromagnetic wave with a multilayer stack

Fig. 2-15 Schematic structure of a stack represented as an assembly of layers.

Each layer has the refractive index \( N \), which can be complex, \( N = n - ki \). The extinction coefficient, \( k \), represents the attenuation in a lossy medium. The optical admittance of the layer is \( y_j = y_{fs} \cdot N_j \), where \( y_{fs} \) is the optical admittance in vacuum. When a plane wave with wavelength \( \lambda_0 \) is propagating into the cavity at normal incidence, the characteristic matrix \( M_j \) of an optically homogeneous layer with refractive index and thickness \( z_j \) is [7], [13]:

\[
M_j = \begin{bmatrix}
\cos \alpha & \frac{i}{y_j} \cdot \sin \alpha \\
\frac{i}{y_j} \cdot \sin \alpha & \cos \alpha
\end{bmatrix},
\]

(2-70)

where \( k_0 = 2\pi/\lambda_0 \) and \( \alpha = (k_0 \cdot N_j \cdot z_j) \)

The final characteristic matrix \( M \) of the film stack is [14]:

\[
M = M_q \cdot M_{q-1} \cdot \ldots \cdot M_2 \cdot M_1
\]

(2-71)
Optical theory of a spectrometer

The reflectance and the transmittance of the total structure can be calculated as:

\[ R = \frac{|y_{in} \cdot B - C|^2}{|y_{in} \cdot B + C|^2} \quad (2-72) \]

\[ T = \frac{4y_{in} \cdot \text{Re}(y_{out})}{|y_{in} \cdot B + C|^2}, \quad (2-73) \]

where

\[ \begin{bmatrix} B \\ C \end{bmatrix} = M \cdot \begin{bmatrix} 1 \\ y_{out} \end{bmatrix}, \quad (2-74) \]

\( y_{in}, y_{out} \) are the admittances of the medium from which the light enters and exits respectively. The absorbance can be calculated from:

\[ A = 1 - (R + T) \quad (2-75) \]

The optical thickness optimization of a layer is treated in next section.

2.6.5 Quarter-wave and half-wave optical thicknesses

It is easy to calculate the admittance of stacks of quarter-waves and half-waves, therefore, the designs are often specified in terms of quarter-waves with respect to a reference wavelength \( \lambda_0 \).

If the optical thickness is an integral number of quarter-waves (for a reference wavelength \( \lambda_0 \) at which the coating is designed), the characteristic matrix of a thin film is described by the following equation (2-70), derived from [11] [8]:

\[ \pm \begin{bmatrix} 0 & i \\ y_1 & 1 \end{bmatrix} \quad \text{with } \alpha = m.(\pi/4) \quad \text{and } m \text{ odd} \quad (2-76) \]

For \( m \text{ odd} \), \( \sin \alpha = \pm 1 \) and \( \cos \alpha = 0 \). This matrix is easy to handle. Such QWOT (quarter-wave optical thickness) structures are represented by
2.6 Interaction of a electromagnetic wave with a multilayer stack

capital letters H, M, L. Thus, H refers to a material which has a higher refractive index, M has the intermediate one and L the lowest. For an optical thickness of an integral number of half-waves, i.e. when \( m \) is even \((\sin \alpha = 0 \text{ and } \cos \alpha = \pm 1)\), one obtains the unity matrix, which does not have an effect on the transmittance and reflectance (absentee layer) [14]:

\[
\pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

(2-77)

The design of QWOT layers and HWOT layers is refined by means of computer simulations that take into account the dependence on the wavelength or the optical properties of the available materials.

2.6.6 Non-ideal boundaries

When radiant flux is incident upon a surface (non-ideal interface), three processes occur: transmission, absorption and reflection.

Fig. 2-16 shows an ideal case where the transmitted and reflected components are either specular or perfectly diffuse. So far only specular transmission and reflection were assumed.

Fig. 2-17 shows the transmission and reflection for more practical surfaces.

Fig. 2-16 Ideal case: transmitted and reflected components are either specular or perfectly diffuse [9].
Optical theory of a spectrometer

The terminology used in this thesis assigns terms like transmissivity, absorptivity and reflectivity for the properties of a pure material. For the characteristics of a specimen or sample the terms used are transmittance, absorptance and reflectance, respectively.

Fig. 2-17 Non-ideal case: transmission and reflection for surfaces [9].

2.7 Ideal plane Fabry-Perot resonator

2.7.1 The principle

An optical resonator traps and confines light of certain wavelengths and operates like an optical transmission system incorporating feedback: light circulates or is repeatedly reflected within the system without escaping. The simplest resonator (Fig.2-18) is composed of two parallel mirrors between which light is reflected and transmitted with little losses (known as the Fabry-Perot resonator). The wavelength selectivity of a Fabry-Perot optical resonator makes it useful as an wavelength-selecting element in a spectrometer. Previous research efforts on micromachined Fabry-Perot devices demonstrate that simplicity in microfabrication, a vertical
structure with only one optical path and reasonable optical performance was achieved. These are strong arguments in favour of using the Fabry-Perot optical resonator to be used as a WSE in the development of a microspectrometer for the visible and near-infrared range of the spectrum.

The fabrication of an integrated microspectrometer in a standard IC process is limited by technological constraints: both lateral and vertical dimensions are limited and shapes are restricted to more or less flat planes. Therefore, the plane plates Fabry-Perot configuration is more suitable for the selected technology than the spherical plates Fabry-Perot configuration.

![Diagram of Fabry-Perot resonator](image)

Fig. 2-18 *Fabry-Perot resonator with two plane mirrors.*

In an ideal Fabry-Perot resonator the two mirrors are assumed to be lossless and perfectly parallel. Also, no absorption in the resonator medium is assumed.

### 2.7.2 Resonator modes

A monochromatic wave of frequency $v$ has a wave function

$$u(r,t) = \text{Re}\{U(r)\exp(j2\pi vt)\},$$  \hspace{1cm} (2-78)

which represents the transverse component of the electric field. The complex amplitude $U(r)$ satisfies the Helmholtz equation (2-24), where $k = 2\pi v/v_m$ is the wave number and $v_m$ is the speed of light in the medium. The modes of a resonator are the basic solutions of the Helmholtz equation. For the planar-mirror resonator, the transverse components of
the electric field vanish at the surfaces of the mirrors, so that \( U(r) = 0 \) at the planes \( z = 0 \) and \( z = d \). The result is the standing wave [1]:

\[
U(r) = A \sin kz, \quad (2-79)
\]

where \( A \) is a constant which satisfies the Helmholtz equation and vanishes at \( z = 0 \) and \( z = d \) if \( k \) satisfies the condition \( kd = q\pi \), where \( q \) is an integer. This restricts \( k \) to the values \( k_q = q\pi/d \), so that the modes have complex amplitudes \( U(r) = A_q \sin k_q z \), where \( A_q \) are constants. Negative values of \( q \) do not constitute independent modes, since the function \( \sin \) is odd. Therefore, the modes of the resonator are the standing waves \( A_q \sin k_q z \), where the positive integer \( q = 1,2,3,... \) is called the mode number (see Fig. 2-19).

![Diagram](image)

**Fig. 2-19** Field distributions of the modes of a planar-mirror interferometer [1].

An arbitrary wave inside the resonator can be written as a superposition of the resonator modes, \( U(r) = \sum q A_q \sin k_q z \). It follows that the frequency \( v = v_m k/2\pi \) is restricted to the discrete values:

\[
v_q = qv_m/2d, \text{ with } q = 1,2,3,... \quad (2-80)
\]

which are the resonance frequencies of the resonator. The resonance wavelengths are, of course, \( \lambda_q = 2d/q \), whereas \( v_m = c_0/n \) is the speed of light in the cavity medium. The analysis of spectra may be complicated by the overlap of different modes. Reduction of the number of modes corresponds to a decrease in the optical path.
The phase shift imparted by a single round trip of propagation \((2d)\) inside the resonator is \([1]\):

\[
\varphi = k2d = q2\pi, \text{ with } q = 1, 2, \ldots \quad (2-81)
\]

This leads to the relation \(kd = q\pi\), where \(k\) is the wave number, which was defined before as \(k = 2\pi v/c\). The resonator works like a feedback system: the output of the system is fed back in phase with the input (see Fig. 2-20).

![Block diagram representing the optical feedback system.](image)

The plane wave of complex amplitude \(U_0\) is reflected from mirror 2 and propagates back to mirror 1 where it is again reflected (its amplitude then becomes \(U_1\)).

An original monochromatic wave \(U_0\) at point P travels to the right along the axis of the resonator (see Fig. 2-21).

![Ideal resonator](image)

**Fig. 2-21** A wave reflects back and forth between the resonator mirrors, resulting a phase shift each round trip \([1]\).
The wave is reflected from mirror 2 and propagates back to mirror 1, where it is again reflected. Its amplitude at P becomes $U_1$. $U_1$ creates $U_2$ and infinite partial waves $U_0, U_1, U_2, U_3, \ldots$ result as well. Their magnitudes are identical because there is no loss is associated with the reflection and propagation. Therefore, the total wave $U$ is the sum of an infinite number of phasors of equal magnitude [1]:

$$ U = U_0 + U_1 + U_2 + \ldots $$  \hspace{1cm} (2-82)

2.7.3 Resonator modes in the visible spectral range

Fig. 2-22 shows the transmitted wavelength as a function of the air gap between the upper and lower mirror. The shaded band represents the visible range of transmitted wavelengths. It is clear from the figure that for first mode ($q = 1$) the required gap setting ranges from 200 to 390 nm if only a single peak is allowed. This means that the first mode is the only range that has no overlap with another mode in the visible range of the spectrum. The air gap between the mirrors has to be of the order of a half wavelength of visible light. These gap dimensions are compatible with what micromachining techniques can offer [15].

![Graph showing transmitted wavelengths vs. air gap](image)

**Fig. 2-22** Relationship between the transmitted wavelengths and the air gap as a function of the modes for a Fabry-Perot resonator [15].
2.8 Non-ideal plane Fabry-Perot resonator

In section 2.7.1 the plane Fabry-Perot resonator was treated as an optical feedback system. This is appropriate when one wants to understand its operation. Real systems will consist of a multilayer stack. The matrix method (from section 2.6.4) can be applied to find the spectral functions of transmittance and reflectance. Professional software packages are commercially available for designing and optimizing a multilayer stack of thin films based on matrix method (in the next chapters a professional program, TFCalc, will be used in optical simulations).

A non-ideal Fabry-Perot resonator, composed of two semi-transparent lossy mirrors, separated by a distance $d$, is analyzed. An amplitude attenuation factor, $r$, is introduced (for mirror reflections and absorption in the resonator medium). For a full rigorous treatment, of the non-ideal Fabry-Perot resonator, the reader is referred to [1] and [6]. First, the spectral response and transmittance of the resonator are derived.

2.8.1 Spectral response

A plane wave of complex amplitude $U_i$ and intensity $I_i$ which enters a planar-mirror resonator undergoes multiple reflections and transmissions, as illustrated in Fig.2-23.

![Diagram of a non-ideal Fabry-Perot resonator](image)

Fig. 2-23 Transmission of a plane wave across a planar-mirror Fabry-Perot resonator [1].
Optical theory of a spectrometer

In a resonator with losses, the excursion of the wave $U_0$ between the two mirrors results in an infinite sum of phasors. The phase difference after reflection at both mirrors is:

$$\varphi = k2d = \frac{4\pi\nu d}{v_m} \quad (2-83)$$

Due to non-perfect mirror reflections and absorption in the medium an effective mirror reflectance is represented by $R$. This implies an intensity attenuation factor $R^2$. Considering $U_1 = hU_0$, where $h = Re^{j\varphi}$ and the phasor $U_2$ are related to $U_1$ by this same complex factor, $h$, as all consecutive phasors [1], yields:

$$U = U_0 + U_1 + U_2 + \ldots = U_0 + hU_0 + h^2U_0 + \ldots \quad (2-84)$$

$$U = U_0 \cdot (1 + h + h^2 + \ldots) = \frac{U_0}{1 - h} \quad (2-85)$$

The intensity in the resonator can be expressed as:

$$I = |U|^2 = \frac{U_0^2}{|1 - Re^{-j\varphi}|^2} = \frac{l_0}{(1 + R^2 - 2R \cos \varphi)} = \frac{l_0}{(1 - R)^2 + 4R \sin^2 \left(\frac{\varphi}{2}\right)} \quad (2-86)$$

This equation is called the Airy transmission coefficient, which is generally rewritten as [1], [6]:

$$I = \frac{l_{\text{max}}}{1 + \left(\frac{2R}{\pi}\right)^2 \sin^2 \left(\frac{\varphi}{2}\right)} \quad (2-87)$$

where $l_{\text{max}} = l_0/(1-R)^2$; and $l_0 = U_0^2$ is the intensity of the initial wave. $F$ is a parameter known as finesse of the resonator:

$$F = \frac{\pi \cdot R^{1/2}}{1 - R} \quad (2-88)$$
From equation (2-87) follows that the spectral response of the Fabry-Perot resonator becomes:

\[
I = \frac{I_{\text{max}}}{1 + \left(2 \frac{F}{\pi} \right)^2 \cdot \sin^2 \left(\frac{\pi \nu}{\nu_F}\right)},
\]  

(2-89)

where \(\nu_F = \nu_m/2d\) represents the constant frequency difference between adjacent resonance frequencies; \(\nu_q = q\nu_m/2d\), \(q = 1, 2, ...,\) are the resonance frequencies of the resonator. Similar to equation (2-80) we have in terms of wavelength:

\[
q\lambda = 2d
\]

(2-90)

If the complex amplitude and intensity of the transmitted wave are \(U_t\) and \(I_t\), respectively, and \(R_1\), \(R_2\) are the amplitude reflectances of the inner surfaces of mirrors 1 and 2, where \(T_1, T_2\) are the amplitude transmittances of the mirrors, it is possible to define the intensity transmittance \(T(\nu)\) of a Fabry-Perot resonator as a function of the frequency of the wave \(\nu\), namely as [1], [6]:

\[
T(\nu) = \frac{T_{\text{max}}}{1 + \left(2 \frac{F}{\pi} \right)^2 \cdot \sin^2 \left(\frac{\pi \nu}{\nu_F}\right)},
\]

(2-91)

where \(T_{\text{max}} = T^2/(1-R)^2\), \(\nu_F\) the spacing between adjacent resonance frequencies (free spectral range, FSR), \(T = T_1 \cdot T_2\), \(R = R_1 \cdot R_2\):

\[
\nu_F = \text{FSR} = \nu_m/2d
\]

(2-92)

The most important parameters to characterize the spectral response of the Fabry-Perot resonator are: finesse, FWHM, free spectral range and resolving power, which will be discussed in more detail in the next sections.

2.8.2 FWHM and finesse

Narrow peaks are formed in transmission. The sharpness of the peaks increases with increasing reflectivity. Finesse (equation (2-88)) is expressed as the ratio of the separation of successive peaks to the peak...
Optical theory of a spectrometer

width at half-intensity (Full-Width Half-Maximum). FWHM is the width of the peaks at half-height. Non-parallelism and plate imperfections affect these characteristics of the spectrometer (see Fig. 2-24).

2.8.3 Free spectral range

The spacing between adjacent resonance frequencies is called the free spectral range (equation (2-92) and Fig. 2-24). Generally, the FSR is the wavelength range over which the spectrometer can measure unambiguously. For Fabry-Perot interferometers with a mirror spacing \( d \), the transmission peaks are separated by a wavelength \( \Delta \lambda \) (FSR), defined as:

\[
FSR = \frac{\lambda^2}{2d}
\]  

(2-93)

In a tunable Fabry-Perot the free spectral range and the FWHM bandwidth can be controlled independently. The cavity gap sets the free spectral range, and the mirrors’ reflectivity controls the bandwidth (finesse). The transmission peaks can be made very sharp by increasing the reflectivity of the mirror surfaces. FSR is inversely proportional to the distance between the mirrors, see equation (2-93) [6] [15].

2.8.4 Resolving power

The resolving power or resolution, \( R \) (Rayleigh criterion), was presented in the first chapter and in the previous section. This criterion cannot be applied directly for the Fabry-Perot resonator (interferometer fringes are of rather different form). The resolving power of the Fabry-Perot interferometer is obtained by:

\[
R = \frac{\lambda}{FWHM} = qF = \frac{q \pi \sqrt{R}}{(1 - R)}
\]  

(2-94)

\( F \) is the finesse, \( q \) is the order mode, \( R \) is the reflectivity of the mirrors, \( \lambda \) is the wavelength. Since resolution is the product of finesse and order number, a low finesse does not necessarily mean low resolution, but it implies that to achieve high resolution the interferometer must be used in high order. Therefore, the resolving power depends principally on the
spacing between the mirrors and their reflectivity. Larger spacings and/or higher reflectivities result in a higher resolving powers.

![Diagram](image)

**Fig. 2-24** Transmission peaks are equally spaced as a function of frequency (FSR).

### 2.8.5 Resonator medium losses

Absorption and scattering in the medium between the mirrors are the losses due to the medium of the cavity. If we define $\alpha_m$, the absorption coefficient of the medium, the round trip attenuation factor of the wave is $\exp(-2\alpha_md)$ [1].

### 2.8.6 Lossy mirrors: reflection, finite size, flatness

Sources of loss in the mirrors are: imperfect reflection (the surface is rough), do not reflect all the light (the mirrors are semi-transparent so that light may escape), and they have a finite size (light leaks away) [1]. Also, the spatial distribution of the wave is matched with the size of the mirror. Macroscopic devices typically have available a 5 cm working diameter area and a flatness of the mirrors better than $\lambda/150$ in the visible region of the spectrum. These values are not realistic for micromachining technology.
Optical theory of a spectrometer

The wave intensity of mirror reflectance \( R_1 \) and \( R_2 \) decreases by the factor \( R_1R_2 \) in the course of the two reflections associated with a single round trip. The overall intensity attenuation factor is [1]:

\[
R^2 = R_1R_2 \exp(-2\alpha_md)
\]  \hspace{1cm} (2.95)

If we define an effective overall distributed loss coefficient, \( \alpha_r \), becomes [1]:

\[
\alpha_r = \alpha_m + \frac{1}{2d} \ln \frac{1}{R_1R_2}
\]  \hspace{1cm} (2.96)

The absorption in thin films is treated with more detail in section 2.9.4.

2.8.7 Angle of incidence and non-parallelism

Light rays that are slightly inclined escape from the resonator more quickly than perpendicular rays. This means low resolution. Multiple reflections occur between the two parallel plane mirrors. Spectral information may be derived by varying one of the three terms in the right-hand side of the equation:

\[
q\lambda = 2nd\cos\alpha,
\]  \hspace{1cm} (2.97)

where \( q \) is an integer, \( \alpha \) is the angle of light incidence, \( n \) the refractive index of the cavity medium and \( d \) is the distance between mirrors (for an air gap between the mirrors the optical thickness is equal to \( d \) and \( n = 1 \)). If \( \cos\alpha = [1/2, 1/3, 1/4, \ldots] \), \( q=1, n=1 \) the working mode is altered, but the resolution and transmittance are deteriorated.

If the mirrors are not parallel the light rays escape from the resonator as well. In principle spectral measurements are possible, however, at the cost of a strongly reduced finesse and resolution. The macroscopic devices available with high stability (\( \lambda/5000 \)) have included subsystems to maintain the parallelism between the Fabry-Perot plates better than 0.1 nm (at \( \lambda = 500 \) nm) [16].
2.8.8 Matrix method vs. Airy function for transmission

In studies on the multilayer Fabry-Perot resonator, frequent use has been made of a simplified model (equation (2-86), also called the Airy function for transmission [6]) to calculate the optical response of a multilayer cavity. This simplified model uses an equivalent lumped reflectivity of a multilayer mirror and enables the representation of a multilayer cavity by a resonator with two very thin mirrors, and ignores phase interactions at each reflection [13]. For practical reasons the matrix method (using a characteristic matrix of each layer in the cavity and calculating the optical response of the multilayer stack) is used and is the basis of the algorithms of the CAD tools used.

2.9 Mirrors

The most important part of the Fabry-Perot device are the mirrors. Two types of highly reflective coatings are used: dielectric and metallic. The dielectric mirrors, when properly designed and fabricated, feature high performance (high reflectivity, low absorption losses). However, the deposition of a sequence of two (or more) different dielectric films of well-controlled thickness must to form a stack of many layers. The performance of the filter is greatly influenced by random thickness variations in the deposition of the films. To be effective in a wide optical band, a dielectric mirror must consist of more than fifteen deposited or grown layers. This complicates fabrication (a complex and costly process). As an example we analyze a multilayer structure of two dielectric materials with different refractive indexes (one high the other low) in the next section.

2.9.1 Dielectric HL layers succession

The multilayer structure is arranged in succession and characterized by the sequence HLHL...HLH. The characteristic matrix of this multilayer now is 2N+1 layers, where N is the number of times HL is used. With \( n_1, n_H, n_L \).
as the refractive indexes of air (incident medium), the layer with high $n$ and the layer with lower $n$, respectively. The matrix becomes [11]:

\[
M_{2N+1} = \begin{bmatrix}
0 & \left(\frac{n_H}{n_L}\right)^N \\
-i n_H \left(\frac{n_L}{n_H}\right)^N & 0
\end{bmatrix}
\] (2-98)

This leads to the reflectivity:

\[
R_{2N+1} = \frac{1 - \left(\frac{n_H}{n_L}\right)^{(2N)^2}}{1 + \left(\frac{n_H}{n_L}\right)^{(2N)^2}}
\] (2-99)

This last expression indicates that the reflectivity increases rapidly with both the ratio $n_H/n_L$ and $N$. Clearly this is a disadvantage for silicon-based structures. Use of silicon-compatible materials limits $n_H/n_L$ to $n_{SiN}/n_{SiO2} < 2$, whereas $N$ to 4-6 for practical reasons.

### 2.9.2 Metallic mirrors

Although the metallic-based coatings have much higher losses than dielectric-based coatings, these can be attractive in certain applications due to the simplicity of their fabrication (only one layer must be deposited). Another advantage is that metallic mirrors generally perform well over a wide spectral range. Aluminum, gold and silver are the most commonly used metals for reflective coatings. Fig. 2-25 shows their reflectance in the near-UV, visible, and near-IR spectral region [7].
Fig. 2-25 Reflectance of silver, gold, and aluminum as a function of the radiation wavelength [7].

Fabry-Perot filters using metallic mirrors cannot provide both high finesse and high transmittance simultaneously, due to optical absorption in the metal. However, in many applications the light intensity is controlled within the measurement system.

2.9.3 Operation at visible spectral range

Aluminum would be the most suitable material in terms of fabrication compatibility, but unfortunately it has higher absorption losses than silver or gold in the visible and near-IR regions, Fig. 2-25. For these spectral regions, silver is the best choice, but silver exhibits poor long-term stability (tendency to tarnish) [17]. Gold is more corrosion resistant than silver and would be the best choice if it were not for its poor performance in the visible and ultraviolet range. Unlike macroscopic applications of silver-based reflective coatings [17], the poor environmental resistance of silver is not critical in a microsystem application, because the sealing of the complete system would avoid any mirror degradation caused by environment. Another advantage is that silver is a natural low-pass filter, cutting off the UV range (see Fig. 2-25). Use of silver-based mirrors has previously been reported in non-tunable distributed filters composed of a wedge-shaped dielectric film sandwiched between two reflective (Ag) thin films [18].
2.9.4 Absorption in metallic thin-films

The mobility of the electrons is high in a thick metallic film. Therefore, the electric field strength parallel to the surface vanishes. This boundary condition causes metallic light reflection. A very thin film is almost transparent to light since there are not enough electrons to interact with the incoming light. In this application the thickness of the metallic film is typically selected between these two extremes. Within this median range of metal thickness, the electrons interact with the wave, but they cannot move freely. In this way the movement of the electrons takes energy from the wave and absorption occurs [19].

2.10 Fabry-Perot based spectrometer

As mentioned already in chapter 1, Fabry-Perot resonance cavities can be scanning-type and array-type wavelength-selecting elements.

2.10.1 Scanning type

A tunable Fabry-Perot interferometer can be used as wavelength-selecting element in a scanning-type spectrometer. We used the basic Fabry-Perot resonator from the last section and added some lenses and a photodetector placed in the focal plane to build a spectrometer (see Fig.2-26). By physically adjusting cavity spacing between the two mirrors, one can tune different resonance wavelengths. This tunable bandpass filter can be designed to scan over a small wavelength region to allow precise spectral tuning. The design of a Fabry-Perot based spectrometer requires: high selectivity, low absorption in the mirrors and cavity and a high-resolution photodetector, according the spectral range in measurement. A small change in mirror spacing causes a substantial shift in the resonance wavelength, \( \lambda_0 \), and the FSR changes a little bit as well.
2.10 Fabry-Perot based spectrometer

Fig. 2-26 Scanning-type spectrometer.

2.10.2 Array type

In an array type of Fabry-Perot etalons, the medium between the two mirrors is a solid thin film. Each of the etalons is formed with a film width tuned for a different resonance wavelength. Underneath each etalon a photodetector is implemented in order to measure the current associated with the resonance wavelength. Fig.2-27 shows the array-type spectrometer. The incident light, I, is conditioned in such a way that all etalons are irradiated. Only a fraction of all light, I/n, reaches all etalons for an array of n etalons. The focusing lens is not necessary in this case, because the photodetectors are positioned directly underneath the etalons.

Fig. 2-27 Array-type spectrometer with four Fabry-Perot etalons, each with its own photodetector.
2.10.3 Spectrum reconstruction

Any spectral structure may be considered to be the sum of an infinite number of simple monochromatic lines at different wavelengths. A perfect spectrometer produces an electrical output signal that represents a spectrum identical to the spectrum of incident light. The major limitation is the spectral broadening of the input.

A light source emits a spectrum which consists of a single monochromatic wavelength. After it has passed the spectrometer the line profile has a finite width [3] (Fig. 2-28).

![Diagram showing a) Real spectrum of a monochromatic light source, b) Recorded spectrum with a perfect instrument, c) Recorded spectrum with a real instrument.](image)

Fig. 2-28  

a) Real spectrum of a monochromatic light source.  
b) Recorded spectrum with a perfect instrument.  
c) Recorded spectrum with a real instrument.

Therefore, a relationship between the spectrometer line profile (impulse response), the input spectrum and the recorded spectrum is defined by:

\[ R(\lambda) = L(\lambda) * P(\lambda), \]  

(2-100)

where \( R(\lambda) \) is the spectrum recorded by the spectrometer, \( L(\lambda) \) the input spectrum for to be analyzed and \( P(\lambda) \) the spectrometer line profile. The recorded function \( R(\lambda) \) is the convolution of the input spectrum and the spectrometer line profile [21], [22]. The spectrometer line profile is a function of various parameters: the width of the beam input, the diffraction phenomena, the quality of the coatings, the alignment and aberrations [3].

Scanning type spectrometer
The scanning-type spectrometer measurements yields a time-continuous function, I(t), in which the intensities measured for each spectral component tuned are sequentially represented. The spectral resolution depends on the resolution achieved for the movement of the mirror.

Array-type spectrometer

The array-type spectrometer provides a discrete number of spectral responses, determined by the number of channels of the array (sets the spectral range).

If we define an array of n etalons covering the visible part of the spectrum (390 nm-760 nm) and the response as a function of the wavelength for each etalon is known, the device can reconstruct the emission spectrum.

Consider an array of Fabry-Perot etalons (each with a photodetector) and designed for a specified wavelength. At the entrance of the spectrometer a light of a certain spectral distribution, \( L(\lambda) \), enters the device (Fig.2-29).

The individual responses of all etalons for the complete spectral range gives us the complete response of the spectrometer for all wavelengths. When the input spectrum \( L(\lambda) \) passes through the spectrometer its output is given by n currents corresponding to n etalons (\( I_1, I_2, \ldots, I_n \)).

![Diagram](image)

**Fig. 2-29**  
*a) A spectral structure coming from a light source. b) The individual response of each etalon for the complete spectral range (called also calibration of the spectrometer).*

Using the matrix theory, we can define the matrix \( I_{n\times n} \), which presents the output currents from a spectrometer as a response to the input spectrum to
be analyzed. The calibration matrix $C_{nxn}$ contains the individual response of each etalon over the entire spectral range with a wavelength increment determined by the resolution. The recorded spectrum, given by matrix $R_{1xn}$, is the response of the spectrometer for an input spectrum in the spectral range.

$$
\begin{bmatrix}
    I_1 & I_2 & \ldots & I_n \\
    C_{11} & C_{12} & \ldots & C_{1n} \\
    C_{21} & C_{22} & \ldots & C_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    C_{n1} & C_{n2} & \ldots & C_{nn}
\end{bmatrix} = 
\begin{bmatrix}
    R_1 & R_2 & \ldots & R_n
\end{bmatrix}
$$

(2-101)

The recorded spectrum defined at $R_{1xn}$ is obtained with an accuracy that depends on the spectrometer line profile and on stray light (diffusion from the optical components and incorrect illumination of the spectrometer).

To optimize the coefficients of matrix $C$ for a specific wavelength range or an input light source, some signal processing techniques are used: simple value decomposition [23], and neural networks [24]. Especially a backpropagation neural network is suitable for this application. During the training phase, some known input spectrums are analyzed by the spectrometer. The control algorithm takes into account non-linearities and adjusts the values of matrix $C$ in order to improve the response of the spectrometer. Increasing the number of training cycles (for a large number of spectrum inputs) improves the resolution of the spectrometer.

The general theory presented in this chapter will be used in the next chapters to design practical Fabry-Perot based microspectrometers.
References


Optical theory of a spectrometer


3.1 Introduction

This chapter presents the design, fabrication, and measured characteristics of a bulk-micromachined tunable Fabry-Perot MicroInterferometer (FPMI) for operation in the visible spectral range. The chapter begins with the structure and material selection, followed by the optical design of the Fabry-Perot resonator. Optimization was performed using CAD tools for optical simulations of thin-film coatings. The resonator top plate is a bossed diaphragm. Therefore, an extensive static and dynamic analysis of diaphragm behavior is presented. Subsequently, the fabrication process is described. Standard IC processing and bulk micromachining were applied. Finally, the theoretical predictions are compared with the measured results.

Fig. 3-1 shows schematically the FPMI structure that is the subject of this chapter. Also, the materials used are indicated. The FPMI is formed by two parallel silver mirrors supported by a low-tensile stress silicon nitride membrane with a square aperture (side length between 0.1 to 2 mm) and initial cavity gap from 1.2 μm to 0.5 μm. One of the mirrors is fixed, the other is under tension on a movable Si frame, which is electrostatically
Scanning-type microspectrometer

deflected and uses several distributed electrodes (Fig. 3-2) to control cavity spacing and mirror parallelism.

![Diagram of a scanning microspectrometer with labeled components: Si frame, movable mirror, Al electrodes, fixed mirror, pads, spacers, silicon, silicon nitride, silver, poly, metal.]

**Fig. 3-1** *Fabry-Perot MicroInterferometer (FPMI).*

**Fig. 3-2** *Bottom mirror layout with the control electrodes that allow control of the cavity spacing and parallelism between the two mirrors.*
Electrostatic actuation is used because it is easy to implement in micromachining technology. Spacers were used to set the minimum gap width in order to avoid sticking and electrostatic pull-in. The following sections motivate the selection of materials and the diaphragm structure.

3.2 Diaphragm design

High-performance diaphragm structures are not only very important in MEMS applications (e.g. pressure sensors, microphones and accelerometers), but also in micro-opto-electro-mechanical devices. Thin electrostatically driven diaphragm mirrors have been used for imaging applications because of their high stability at small deflections. Also, adaptive mirrors based on large-area silicon nitride diaphragms were applied for correction of optical aberrations [1].

Thin metallic mirrors (thickness ≤50 nm) cannot be self-supporting. A supporting membrane is required. The diaphragm must feature high transparency in the wavelength region required, flatness over the entire region of operation, elasticity and high mechanical strength to allow fabrication of large and thin structures. Also, the mirrors must have high reflectivity and low absorption.

3.2.1 Selection of the diaphragm geometry

A tunable FPMI involves adjustable spacing between the two mirrors, i.e. one of the mirrors must be movable. Different techniques and structures have been developed. From all the possible types the following ones are considered in this study (see Fig. 3-3):

(a) cantilever [2]
(b) “H”-shaped structure (square suspended membrane with four beams)
(c) simple clamped membrane
(d) perforated membrane
(e) composite SiO₂/SiN diaphragm [3]
(f) silicon nitride membrane/silicon frame.
Scanning-type microspectrometer

FEM simulations show that the cantilever (a) needs relatively low tuning voltage to bend, but flatness and parallelism deteriorates with increased deflection. This is in agreement with literature [2]. Therefore, conditions for optical resonance will be achieved only in a very small region, which strongly limits the application in a Fabry-Perot device. Simple fabrication by bulk or surface micromachining is one of the advantages of this structure.

The square suspended membrane with four beams (b) needs a higher voltage for an equivalent bending as in (a), but the area over which it can be considered flat is not much greater. The simple clamped membrane (c) requires the highest voltage to bend, compared to (a) and (b). The area over which the deflected membrane can be regarded as flat is larger, but far from ideal (see Fig.3-4). In the case of the perforated membrane (d) the etch holes reduce the effect of the membrane internal stress (and also lead to a reduction of the squeezed-film damping). Consequently, structure (d) features a reduced tuning voltage, but the flat area is not better than in (c) [4] [5]. The composite SiO₂/SiN diaphragm (e) takes advantage of the compressive stress in SiO₂ films to compensate for the silicon nitride tensile stress, in order to reduce the internal stress [3]. The fabrication process of the structures (d) and (e) is more complex than that of the others.

The composite (silicon nitride membrane/silicon frame) diaphragm (f), requires the highest voltage for deflection compared to all other structures. However, the silicon frame ensures flatness over a large area (Fig.3-5). The silicon frame is used instead of a “mesa”, because the inner part of the frame is used as active region of the Fabry-Perot device. The fabrication process is simple and comparable to structure (c).

Corrugated silicon membranes exhibit large and linear deflection [6], however, compressive stresses generated in such regions cause membrane buckling. Therefore, corrugated structures were not considered.

From all the diaphragm structures studied, FEM simulations show that the mechanical and optical behavior of the structure (f) is closest to the requirements of the intended application (see Table3-1). Also, the fabrication of this type of diaphragm promises a relatively simple process.
3.2 Diaphragm design

Fig. 3-3 *Diaphragm structures*: (a) cantilever, (b) H-shaped structure held on four sides, (c) simple membrane, (d) perforated membrane, (e) composite SiO₂/SiN membrane, (f) SiN membrane/Si frame.

Fig. 3-4 *Deflection of a membrane without frame* (structure c).
Scanning-type microspectrometer

Fig. 3-5  *Deflection of a membrane with frame (structure f).*

<table>
<thead>
<tr>
<th>Structure type simulated</th>
<th>Fabric.</th>
<th>Low Voltage</th>
<th>Internal stress</th>
<th>Flatness</th>
</tr>
</thead>
<tbody>
<tr>
<td>cantilever (a)</td>
<td>-</td>
<td>++</td>
<td>+</td>
<td>--</td>
</tr>
<tr>
<td>&quot;H&quot;-shaped (b)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>--</td>
</tr>
<tr>
<td>membrane(^1) (c)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>perf. memb. (d)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>SiO(_2)/SiN (e)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Si-frame/SiN (f)</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>++</td>
</tr>
</tbody>
</table>

\(^1\)Simple membrane is used as a reference
3.3 Optical design

3.2.2 Diaphragm materials

Several materials have been considered for diaphragm fabrication: silicon, silicon nitride, oxynitride. Silicon nitride was selected in combination with bulk micromachining, as its mechanical strength and the high SiN/Si etch selectivity in KOH allows the fabrication of large and thin membranes [1]. Also, silicon nitride is transparent in the visible optical range and has been applied to micro-optical devices, such as microinterferometers, despite the fact that control is required to avoid excessive stress levels. The silicon nitride layer is used on the front as a deformable membrane. The mirror was formed by evaporation of a thin silver layer on the SiN-membrane.

3.3 Optical design

Accurate data about the optical properties of coating materials is important for designing and manufacturing coatings. Because optical properties may depend on the process used to deposit the coating, it is best to measure the refractive index and the extinction coefficient of layers produced by the process. When we are designing a coating as a proof-of-concept, the accuracy of refractive index data is less important and we can use material data from published sources and then, after the actual indices have been substituted, the coating design can be optimized again.

3.3.1 Optical properties of the materials

From literature, it is possible to find general material data that can be used for an approximate calculation [8], [9], [10], [11]. The simulations in this thesis use the material data provided by the database of Software Spectra company [9] for silver. Recently, the Sopra company [11], which manufactures ellipsometers, posted in the web a database containing optical data of more than 200 materials.

The refractive index and extinction coefficient for low-stress LPCVD silicon nitride are shown in Fig.3-6 and Fig.3-7 (determined by spectroscopic ellipsometry [8]). The low-stress silicon nitride, when compared to the stoichometric Si₃N₄, is characterized by a stronger variation with wavelength and by higher values of the refractive index. This happens because, in order to reduce the tensile stress, the LPCVD-deposited nitride must be silicon-rich.
Table 3-2 shows the variation as a function of the wavelength of the refractive index and extinction coefficient for silver according the two databases before mentioned [9], [11]. Air has a refractive index very close to 1.

![Graph showing refractive index vs. wavelength for low-stress silicon nitride and Si$_3$N$_4$.]

**Fig. 3-6** Measured dispersion data for the refractive index of LPCVD low-stress nitride together with the data for stoichometric silicon nitride (Si$_3$N$_4$) [8].

![Graph showing extinction coefficient vs. wavelength for different LPCVD deposited films.]

**Fig. 3-7** The measured dispersion of the extinction coefficient $k$ for different LPCVD deposited films [12].
3.3 Optical design

Table 3-2  Refractive index and extinction coefficient of silver as a function of wavelength [9], [11].

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>Database SSpectra, 1996</th>
<th>Database Sopra S.A., 1999</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Refr. index n</td>
<td>Ext. coeff. k</td>
</tr>
<tr>
<td>350</td>
<td>0.1</td>
<td>5.85</td>
</tr>
<tr>
<td>400</td>
<td>0.075</td>
<td>1.93</td>
</tr>
<tr>
<td>450</td>
<td>0.055</td>
<td>2.42</td>
</tr>
<tr>
<td>500</td>
<td>0.05</td>
<td>2.87</td>
</tr>
<tr>
<td>550</td>
<td>0.055</td>
<td>3.32</td>
</tr>
<tr>
<td>600</td>
<td>0.060</td>
<td>3.75</td>
</tr>
<tr>
<td>650</td>
<td>0.070</td>
<td>4.2</td>
</tr>
<tr>
<td>700</td>
<td>0.075</td>
<td>4.6</td>
</tr>
<tr>
<td>750</td>
<td>0.085</td>
<td>5.05</td>
</tr>
<tr>
<td>800</td>
<td>0.10</td>
<td>5.3</td>
</tr>
</tbody>
</table>

3.3.2 Optical simulations

A commercially available thin-film optics software package TFCalc 3.2.5 (supplied by Software Spectra, Inc., USA), was used to perform optimization of mirror layer thickness and composition. All these computer simulations took into account the wavelength dependence of the material properties.

A Fabry-Perot resonator composed of two silicon nitride layers on both sides, silver mirrors with variable thickness, and an air gap of 1 μm (see Fig.3-8) was used as a simulation vehicle to build the graph presented in Fig.3-9. This information enabled us to investigate a trade-off between finesse and transmittance by varying the thickness of the silver mirrors.
Fig. 3-8  *Simulated Fabry-Perot resonator.*

Fig. 3-9  *Transmittance and finesse as a function of silver thickness for a Fabry-Perot cavity (SiN-300 nm, air gap 1000 nm).*

Fig. 3-10 shows simulated transmittance for 40nm-Ag/300nm-SiN mirrors with mirror spacing as a parameter. For cavity gaps 500 nm and 300 nm one single peak is obtained with a FWHM of 10 nm. A finesse of 20 (FSR=100 nm and FWHM=5 nm) is predicted for 1000 nm.
Fig. 3-10  Simulated transmittance for different cavity gaps.

Fig. 3-11 shows a simulated cavity for mirrors with 20 nm thick silver mirrors, which implies a transmittance >80%, but the finesse and FWHM are very poor. Increasing the thickness of the silver mirrors to 50 nm improves the overall optical performance of the Fabry-Perot cavity (see Fig. 3-12).

Fig. 3-11  Simulation of the transmittance of a cavity composed of: SiN membrane 300 nm, 20 nm Ag mirrors, air gap spacing 500 nm.
Scanning-type microspectrometer

Fig. 3-12 Simulation results for a cavity composed of: SiN membrane 300 nm, 50 nm Ag mirrors, air gap spacing 500 nm.

Fig. 3-13 shows the response predicted for a large cavity of 4000 nm. The finesse achieved is higher than 20.

Fig. 3-13 Simulation results for a cavity composed of: SiN membrane 300 nm, 40 nm Ag mirrors, air gap spacing 4000 nm.
3.4 Load deflection: analysis of static response

According to the simulations, a mirror thickness of 40 nm to 50 nm offers the best performance when a trade-off between finesse and transmittance is considered. Cavity gaps ranging from 1 μm to 0.2 μm are required to tune only one resonance peak in the visible spectral range (see chapter two, section 2.7.3).

3.4 Load deflection: analysis of static response

Electrostatic actuation was selected for the cavity tuning. To predict the mechanical behavior of the diaphragm and to optimize its design according to application requirements, we used simulations to avoid expensive “trial and error” process runs. The load-deflection function of silicon diaphragms can be predicted by means of simple calculations based on approximate analytical methods. However, these lack the ability to deal with corrugation profiles, complex shapes, effects arising from internal stress, and stress discontinuities. Therefore, use of finite element methods becomes necessary in modeling and simulating more complex diaphragms.

The mechanical behavior of diaphragms is governed mainly by their geometry (thickness, shape) and material properties (residual built-in stress, Young’s modulus, Poisson ratio). The material data depend strongly on the fabrication process conditions. Therefore, for accurate results, we require reliable data extracted for a particular process.

The mechanical behavior of the diaphragm under a static pressure is discussed. This analysis joins two different forces: elastic and electrostatic forces. The movement of the frame mass causes the membrane to bend, which results in a vertical counterforce: \( F_{\text{spring}} = -Ky \), where \( K \) denotes the spring constant of the diaphragm suspension and \( y \) is the diaphragm displacement.

Tabata et al. [13], [14] presented an analytical solution for load deflection of rectangular flat membranes. But the most detailed analysis is based on finite element modeling [15], [16], [17]. The load-deflection behavior of
Scanning-type microspectrometer

square membranes with initial tensile stress ($\sigma > 0$) was shown to be well described by [15]:

$$P(y) = \frac{c_1 \sigma dy}{(2a)^2} + \frac{f(v)c_2 \epsilon y^3}{(2a)^4 (1 - \nu)},$$  \hspace{1cm} (3-1)

where $P$ denotes the applied pressure, $y$ the center deflection, $a$ one half of the membrane's edge length, $t$ the membrane thickness, $\sigma$ the residual stress, $E$ the Young’s modulus, $c_1 = 13.64$ and $c_2 = 21.92$, the function $f(v)$ depends on Poisson’s ratio through $f(v) = 1.446 - 0.427v$. For $\nu = 0.25$ (SiN Poisson’s ratio) becomes:

$$P(y) = \frac{c_3 \epsilon y}{a^2} + \frac{c_4 \epsilon y^3}{a^4},$$  \hspace{1cm} (3-2)

where $c_3 = 3.41$ and $c_4 = 2.45$ [16].

The first term indicates the part of the stiffness that results from the residual stress of the membrane, and the second term presents the stiffness of the material. For small deflections, the second term in equation (3-1) and equation (3-2) is negligible.

The FPMI diaphragm has a silicon frame (with irregular shape) on the membrane. The load to deflect the diaphragm is applied only at the area under the frame (localized pressure due to voltage applied at control electrodes), defined by its bottom width and length dimensions. Therefore, in order to understand the mechanical behavior of this device (analytical treatment is only valid for a first approximation), numerical calculations based on FEM must be used. An approximate relationship between the elastic force and the deflection of a membrane (with a frame in silicon) was obtained by means of FEM for small deflections ($< 3 \mu m$) [14], [18], [19]:

$$F_{\text{elastic}} = \frac{4c_3 \epsilon y}{(1 - \frac{b}{2a})^2} \text{ with } 0 < b < 2a,$$  \hspace{1cm} (3-3)

where $b$ is the frame size (square shape) [20]. The resulting elastic force within the membrane appears when the mass structure is vertically deflected by a certain amount $y$. This deflection can be achieved either
3.4 Load deflection: analysis of static response

with a distributed load (as in the case of a pressure) or with a non-distributed load (as in the case of an acceleration).

A composite diaphragm (SiN-membrane/Si-frame), with four patterned metal electrodes clamped to a silicon frame and suspended above a ground plate with a gap \( d_{el} \), is used for the top plate in the Fabry-Perot device. The electrostatic force between the two plates is:

\[
F_{elect} = \frac{V^2 \varepsilon A_{el}}{2d_{el}^2}, \tag{3-4}
\]

where \( V \) is the voltage applied between the plates, \( A_{el} \) the electrode area, \( \varepsilon \) the permittivity of the medium between the plates (air gap), and \( d_{el} \) the distance between the electrodes.

Combining equation (3-3) and equation (3-4) results for small deflections in:

\[
y = \frac{V^2 \varepsilon A_{el} \left( 1 - \frac{b}{2a} \right)^2}{4c_3 t \sigma d_{el}^2} \tag{3-5}
\]

From the equation (3-4) follows that the bending of the diaphragm obeys:

\[
Bending = \frac{1}{d_{el}^2}, \tag{3-6}
\]

this shows that for a 1 \( \mu \)m air gap the diaphragm will bend 9 times more than a 3 \( \mu \)m diaphragm for the same voltage [21].

3.4.1 Input data

The input data for FEM simulations are the optical constraints and the material properties (Young’s modulus \( E \), internal stress \( \sigma \), Poisson’s ratio \( \nu \)). We recognize that for a given structure, a higher performance load-deflection characteristic can be obtained using a circular membrane. However, this work is restricted to square diaphragms which can be fabricated using simple bulk micromachining. Different diaphragm lateral dimensions (typically between 3x3 mm\(^2\) and 8x8 mm\(^2\)), membrane
Scanning-type microspectrometer

thicknesses (from 150 nm to 1 μm), frame sizes (from 1.8x1.8 mm² to 4x4 mm²), and square apertures dimensions (between 200 μm to 2 mm) were simulated. These values are mostly set by technology constraints.

The etching of silicon in KOH is very anisotropic: the \{100\} planes and \{110\} planes are selectively etched, while the etch rate in the <111> direction is much lower [22]. As a result, when etching a square membrane in a (100)-cut wafer (thickness 525 μm) in KOH, side-walls are formed under an angle of 54.74° with respect to the surface, as shown in Fig.3-14 and Fig.3-15.

![Diagram](Image)

Fig. 3-14 Dimensions set by technology constraints.

![Diagram](Image)

Fig. 3-15 Input data for the FEM simulations of composite diaphragm.
A top frame width of 50 μm thus implies a bottom frame width (and also a backside mask pattern) of 794 μm. Also, the frame convex corners will be etched away, resulting in a loss of the desired structure [23]. To prevent this, corner-compensation structures have been added at the convex corners of the silicon frame. The values used in FEM simulations (Fig. 3-15) for low-stress LPCVD silicon nitride were: $E = 360$ GPa, $\sigma = 0.125$ GPa (material data extracted from the fabrication process used) [7].

### 3.4.2 FEM techniques

To model a device using FEM one must define the type of construction elements, mesh profile, material properties, boundary conditions, and loads [24]. The conditions and requirements relevant in the FEM simulation of the diaphragm are:

1. the top plate is initially flat, parallel and movable with respect to the bottom plate (fixed infinite ground plate).

2. the movable plate operates in the small-deflection regime until pull-in occurs (linear elastic mechanics).

3. the movable plate has perfect boundary conditions (all six degrees of freedom at each boundary are fixed).

Diaphragm thickness, dimensions and the load-deflection behavior were optimized with FEM in order to achieve large deflection at a minimum value of the electrostatic control voltage.

The membrane was modeled by a three-dimensional shell element (very small thickness), and the silicon frame by a three-dimensional solid.

A mesh refinement at the corner regions of both diaphragm and frame improves precision of the FEM simulations. The membrane residual tensile stress was simulated by defining a thermal-expansion coefficient and applying a temperature load.

### 3.4.3 FEM simulations vs. load deflection results

FEM simulations were used to study the influence of the deformable length $Z$ and the frame size $b$ for the deflection of the diaphragm (Fig. 3-16 and Fig. 3-17) [20].
Fig. 3-16 *Deflection for different values of Z when the force applied and the frame size are constant (membrane thickness is 500 nm).*

Fig. 3-17 *Diaphragm deflection for different values of b (frame size) when the force applied and the diaphragm size are constant (membrane thickness is 500 nm).*

Diaphragms with 3 mm and 4.2 mm frame sizes were fabricated. To test the diaphragms' load-deflection behavior, we used a TENCOR Alpha-Step 500 surface profilometer with a needle force of 187 μN and a vertical resolution of 2.5 nm.
3.4 Load deflection: analysis of static response

The surface profilometer scan across the entire membrane (corresponding to the deformable membrane area, frame area and mirror surface) is shown in Fig. 3-18. The scanning length was 4.2 mm, corresponding to the active area of the diaphragm (edge to edge).

The data obtained was compared with the FEM simulations for the same conditions (frame size 3.0 mm, membrane thickness 500 nm, deformable membrane length 0.6 mm and mirror square aperture 1.2 mm). The results show that low-stress non-planar composite diaphragms can be simulated within 10% of error. This is associated with the measurement of the diaphragm thickness and internal residual stress [20].

![Graph showing FEM Simulation and Experimental data comparison](image)

Fig. 3-18 Comparison of a surface profilometer scan (needle force of 187 μN) with FEM simulations.

Also, the surface profilometer was used to analyze the load deflection only in the square aperture size (Fig. 3-19).
Scanning-type microspectrometer

The scanning length was 2.2 mm (larger than previous one) and the results show that low-stress non-planar composite diaphragms can be simulated within 2% of error, allowing design optimization and predicting long mechanical reliability.

3.4.4 Influence of internal stress

A flexible membrane preserves its shape under lateral tension. This is the main reason for choosing silicon nitride for the membrane material. Tensile stress membranes improve the flatness and optical quality. However, increasing tensile stress requires increasing voltage for electrostatic actuation. Therefore, a compromise between voltage and built-in internal stress is required. The deposition conditions of the LPCVD silicon nitride, as well the annealing temperature used allow setting of the internal built-in stress at a suitable value equal to 0.125 GPa.

3.4.5 Influence of corner shape

Bulk micromachining could influence the corner shape, which may affect:

1. the deflection of the center of the diaphragm.

2. the stress concentration in the corner.
3.4 Load deflection: analysis of static response

Typically, a sharp corner geometry increases the stress concentration. For this reason, small etching variations in the shape of the corner can have a dramatic influence on the device behavior. To avoid overetching of the frame convex corners during the anisotropic etching of silicon, corner-compensation structures were used. Fig.3-20 shows a typical corner shape when compensation structures were used.

The corner overetching depends on the etching time, so this parameter can be used to control the resulting corner shape. An example of a corner with different etching ratios was simulated in order to study the deflection of the diaphragm with a constant force applied (see Fig.3-21). The results show that the deflection is four times higher when the same force is applied in case of a 50% corner overetch (overetching defined as indicated in insert Fig.3-21). Therefore, selective overetching can be used to decrease the value of the electrostatic voltage necessary for deflection [20].

The difference in stress concentration with the same deflection (until 2 μm) between an 100% overetched corner and a perfect corner is not significant (less than 1 MPa at 127 MPa). However, for large deflections (high loads), very high stress concentration values can be encountered, especially in thin membranes.

Fig. 3-20 Etching of (100)-oriented corner-compensation structures in aqueous KOH with resulting convex corner [23].
3.5 Load deflection: analysis of dynamic response

With respect to the dynamic response of the diaphragm, four different forces must be identified: inertial, squeezing-film damping, elastic, and electrostatic forces. The resulting motion equation is described by [25]:

\[
M \frac{d^2 y}{dt^2} + D \frac{dy}{dt} + K_y = \frac{V^2 \varepsilon A_{el}}{2d_{el}^2},
\]

(3-7)

where M is the weight of frame mass (membrane mass is neglected-0.099 mg for a=4 mm), and D is the damping factor.

Between the diaphragm and the bottom plate there is a small separation, and a displacement normal to the device tends to squeeze the air in or allows it to flow out of the gap. The behavior of squeeze gaseous films is, for very small geometries, governed by the viscous flow only [25]. Due to the viscous behavior of the gas, this movement is resisted, which results in a pressure gradient in the gas film (see Fig.3-22). This pressure gradient results in a counterforce, acting on the diaphragm. The faster the mass moves, the larger the pressure gradient and therewith the reaction force.
This study implies that the diaphragm is deflected until the square aperture (middle of the frame) stays flat.

Fig. 3-22  *Schematic configuration of the squeeze-film damper.*

The relation between velocity, density, viscosity, and pressure for an isotropic Newtonian fluid with a laminar gas flow between two moving plates of arbitrary shapes can be calculated by means of the Navier-Stokes momentum equations and continuity equation. These complex equations are not presented here, but can be simplified under the following conditions [26], [27], [28], [29], [30]:

-The surfaces of the plates are parallel and most of the motion is perpendicular to the surfaces. The film thickness is uniform.

-The gas film is isothermal and, therefore, the density is proportional to the absolute pressure. This condition is valid because the conducting path within the gas to the bounding wall is very small and the silicon walls have a relatively high heat capacity.

-Inertial forces are small when compared to viscous forces, which means a small Reynolds number with respect to squeeze motion: \( R = \omega d_0^2/\mu \ll 1 \). Where \( \omega \) is the frequency of the movement of the electrode, \( \rho \) the density, \( \mu \) the viscosity of the gas, and \( d_0 \) the film thickness. With air of atmospheric pressure, \( \mu = 1.8\times10^{-5} \) Pa.s, \( \rho = 1.29 \) kg/m\(^3\) and air gap \( d_0 = 1.2 \) \( \mu \)m, and the relation is valid up to frequencies <9.6 Mrad/s.

-The film thickness \( d_0 \) is large (at least a factor hundred) when compared to the mean free-molecular path \( \lambda \). If this condition is satisfied, the distributed gas-film velocity is identical to the film surface velocity and
continuum flow instead of slip flow exists. If the influence of slip flow is significant an effective viscosity can be used to account for this effect.

When these assumptions are valid a single differential equation can be written which is called the Reynolds equation [25]:

$$\frac{\partial}{\partial x}\left(\frac{\rho d_e^3}{\mu} \frac{\partial}{\partial x} P\right) + \frac{\partial}{\partial y}\left(\frac{\rho d_e^3}{\mu} \frac{\partial}{\partial y} P\right) = 12 \frac{\partial}{\partial r}(\rho d) ,$$

(3-8)

where $\rho$ is air density, $\mu$ is the viscosity or the effective viscosity, $d_e$ is the gap separation, and $P$ is the pressure inside of the cavity.

### 3.5.1 Slip flow

One of the conditions which has to be satisfied to allow the use of the Reynolds equations is that the film thickness $d_0$ is large when compared to the mean free-molecular path $\lambda$. The ratio between these values is defined as the Knudsen number $K_n$ [26] [31]:

$$K_n = \frac{\lambda}{d_0} = \frac{p_0 \lambda_0}{p_a d_0} ,$$

(3-9)

where $\lambda_0$ is the mean free path at pressure $p_0$, and $p_a$ is the actual pressure of the gas. This number is the inverse of the average amount of collisions a gas molecule experiences when it crosses the thickness of the gas film. Three situations may happen: $K_n < 0.01$, the flow may be treated as a continuum; $0.01 < K_n < 15$ slip flow becomes significant, and for $K_n > 15$ fully developed molecular flow results. The effective viscosity that is proportional to the pressure $p_a$ is:

$$\mu_{eff} = \frac{\mu}{1 + f(K_n)} ,$$

(3-10)

Several correction functions $f(K_n)$ can be found in literature, but for small Knudsen numbers the results are comparable and can be defined as:

$$f(K_n) = 9.638K_n^{1.159} ,$$

(3-11)
3.5 Load deflection: analysis of dynamic response

In our case the film thickness (air gap) is 1.2 μm. The free-molecular path at the atmospheric pressure is \( \lambda = 63 \text{ nm} \) and thus \( K_{ij} = 0.05 \), which results in an effective viscosity \( \mu_{\text{eff}} = 0.76 \mu = 0.76 \times 1.8 \times 10^{-3} = 1.37 \times 10^{-5} \text{ Pa.s} \).

3.5.2 Diaphragm response time

If we consider the pumping action of the gas in a narrow gap as the dominant damping mechanism in the motional vibrations, the gas flow has very small Reynolds numbers:

\[
R = \frac{\omega d_{el}^2 \rho \mu^{-1}}{1},
\]

(3-12)

where \( \omega \) is the oscillation frequency of the diaphragm, \( \mu \) is the effective viscosity in this case, \( d_{el} \) is the gap separation, and \( R \) is the Reynolds number, which must be less than one in order to keep the Reynolds relation valid, equation (3-8). Typically \( \rho_{\text{air}} = 1.3 \text{ kgm}^{-3} \), \( \omega << 1000 \text{ s}^{-1} \) for mechanical oscillations, and \( \mu_{\text{eff}} = 1.37 \times 10^{-5} \text{ Pa.s} \). Therefore, \( R = 1.36 \times 10^{-4} \), and equation (3-8) is valid.

From equation (3-8), in order to obtain an analytical solution, small pressure variations and small displacements are assumed [25], [32], [33], [34]. Integration of the pressure change over the plate area results in the squeeze-film damping force. An important parameter in squeeze-film calculations is the non-dimensional squeeze number \( \sigma \), defined by:

\[
\sigma = \frac{12 \mu_w^2 \omega}{P d_{el}^2} = \frac{12 \mu_{\text{eff}} (2a)^2 \omega}{P d_{el}^2},
\]

(3-13)

where \( w \) is the characteristic width of the plate. At low squeeze numbers (low frequencies) the air can move out of the gap, resulting in a damped movement due to the viscosity of the gas, whereas at high squeeze numbers, the air is compressed and cannot move out of the gap [33]. For \( w = 2a = 8 \text{ mm} \), \( \omega << 1000 \text{ s}^{-1} \), \( d_{el} = 1.2 \text{ μm} \), the ambient gas pressure \( P > 1 \times 10^5 \text{ Pa} \) and \( \mu_{\text{eff}} = 1.37 \times 10^{-5} \text{ Pa.s} \) yields \( \sigma < 1 \). The compressibility
Scanning-type microspectrometer

effects [25] [32] can be ignored (low non-dimensional squeeze numbers, \( \sigma < 1 \)) and the damping force under these conditions is [32]:

\[
F_{damp} = D \frac{d}{dy} = \frac{\mu w^3 L f(w/L)}{d_e} \frac{d}{dy},
\]  

(3-14)

where \( w \) and \( L \) are the width and length of the plate, and \( f(w/L) \) is a function which depends on the aspect ratio of the rectangular plate. For square plates, \( f(w/L) = 0.42 \) [25].

Another essential parameter in the motion equation (3-7) is the weight of the frame mass (see Fig. 3-23). Its value is obtained by first calculating the pyramidal mass \( (M_1) \), where \( C_1 \) is the side length of the square mass at the base, and \( C_2 \) the side length of the mass at the truncated top. Subtracting from this mass the second mass \( (M_2) \) with the same shape and dimensions \( C_3 \) (side length of the base) and \( C_4 \) (side length of the top), we obtain the frame mass. The height of the frame mass is defined by the wafer thickness. Therefore, results for the weight frame mass (by integration) [34]:

---

Fig. 3-23 *Dimensions of the frame mass (top view).*
3.5 Load deflection: analysis of dynamic response

\[ M_1 = \int_{z=0}^{t_{wafer}} \int_{y=-C_1/2}^{(C_1 - C_2)/2} \int_{x=-C_1/t_{wafer}}^{(C_1 - C_2)/2} \rho \, dx \, dy \, dz \]

\[ = \frac{\rho t_{wafer} (C_1^3 - C_2^3)}{3(C_1 - C_2)} = \frac{\rho t_{wafer}}{3} \left( C_1^2 + C_1 C_2 + C_2^2 \right), \]

\[ M_{frame} = \frac{0.5}{6} \left[ (C_1^3 - (C_1 - t_{wafer}/2)^3) - (C_3^3 - (C_3 - t_{wafer}/2)^3) \right] \]

\[ C_2 = C_1 - t_{wafer} \sqrt{2} \]

\[ C_4 = C_3 - t_{wafer} \sqrt{2}, \]

where \( \rho \) is the density of the material \( (\rho_{Si} = 2.33 \times 10^3 \text{ kg m}^{-3}) \), and \( t_{wafer} \) is the thickness of the silicon wafer. The pyramidal shape obtained is due to the anisotropic etching of silicon in KOH (see section 3.2.5).

With \( D, M \) and \( K \) known, it is possible to solve the second-order differential equation in order to estimate the recovery time of the diaphragm:

\[ \frac{d^2 y}{dt^2} + D \frac{dy}{dt} + Ky = 0, \]

setting the initial conditions: \( y(0) = y_0 \), and \( (dy/dt)_{t=0} = 0 \), \( \zeta \), the damping ratio and \( \omega_n \), the undamped natural frequency are expressed as [32]:

\[ \zeta = \frac{D}{2M_{frame} \omega_n} \]

\[ \omega_n = \sqrt{\frac{K}{M_{frame}}} \]
using $K = 3410$ N/m, a large membrane ($b = 4$ mm, $t = 500$ nm, $\sigma = 125$ MPa, $a = 4$ mm), $M_{\text{frame}} = 2.15 \times 10^{-6}$ kg and $D = 13639$ N.s/m, $\zeta = 7.96 \times 10^4$, $\omega_n = 3.98 \times 10^4$ s$^{-1}$.

When $\zeta$ equals 0.707, the mechanical system is critically damped, meaning that the amplitude response is flat over the largest frequency range. From the values of $K$ and $D$ we conclude that the motion of the diaphragm is highly damped. The equation of motion has the solution [33]:

$$y(t) = T \exp \left[ \zeta \omega_n \left( \sqrt{1 - \left( \frac{1}{\zeta} \right)^2} - 1 \right) t \right] + X \exp \left[ (1-\zeta) \omega_n \left( \sqrt{1 - \left( \frac{1}{\zeta} \right)^2} + 1 \right) t \right]$$

(3.21)

with

$$T = y(0) \left( \frac{\sqrt{1 - \left( \frac{1}{\zeta} \right)^2} + 1}{\sqrt{1 - \left( \frac{1}{\zeta} \right)^2}} \right)$$

$$X = y(0) \left( \frac{\sqrt{1 - \left( \frac{1}{\zeta} \right)^2} - 1}{\sqrt{1 - \left( \frac{1}{\zeta} \right)^2}} \right)$$

(3.22)

The second term is small and decays to zero in a very short time. The characteristic response time can therefore be approximated as:

$$\zeta \omega_n \left( \sqrt{1 - \left( \frac{1}{\zeta} \right)^2} - 1 \right) t = -1,$$

(3.23)

developing in series $[1-(1/\zeta)^2]^{1/2} = 1-0.5(1/\zeta)^2$, gives for the response time:

$$t = \frac{D}{K}$$

(3.24)

For a different frame mass and a different membrane length, the response time ranges from 4 s ($2a = 8$ mm, $b = 4$ mm) to 0.12 s ($2a = 2.8$ mm, $b = 2$ mm). The response time can be decreased by opening circulation...
holes in the silicon nitride membrane to free the air from the squeezing cavity.

3.5.3 Pull-in voltage

When a voltage is applied between the two plates, the diaphragm bends toward the ground plate due to attractive forces. At a critical voltage, called the pull-in voltage or $V_{pi}$, the linear elastic restoring force is exceeded by the nonlinear attractive force, and the diaphragm collapses (this event is abrupt and readily observed experimentally). Deflection increases with voltage until pull-in is reached. The value of $V_{pi}$ depends on the material constants and diaphragm geometry. Fig. 3-24 shows the FEM simulations of the air gap as a function of applied voltage for four different devices (A,B,C,D, see Table 3-3). The cavity gap is 1.2 μm, membrane thickness 500 nm, residual stress 125 MPa. As the voltage is swept from 0 to the pull-in voltage, the air gap decreases from 1.2 μm to 740 nm. A further increase in voltage results in pull-in of the diaphragm. The increase of the electrostatic force, when the diaphragm as it is moved closer to the lower plate, is responsible for this phenomena.

Literature shows a pull-in voltage depending on the structure, residual stress level, and pull-in at 1/3 of the gap spacing between the plates when the plate-spring model and 2/5 of the gap spacing for the Rayleigh approximation is used [26]. The experimental results shows a pull-in voltage that is higher (between 10% to 20%) than that in the FEM simulations (see Fig. 3-24).

Table 3-3  Geometry dimensions of four different devices.

<table>
<thead>
<tr>
<th></th>
<th>Device A</th>
<th>Device B</th>
<th>Device C</th>
<th>Device D</th>
</tr>
</thead>
<tbody>
<tr>
<td>2a</td>
<td>8 mm</td>
<td>4.2 mm</td>
<td>6 mm</td>
<td>2.8 mm</td>
</tr>
<tr>
<td>b</td>
<td>4 mm</td>
<td>3 mm</td>
<td>4 mm</td>
<td>2 mm</td>
</tr>
<tr>
<td>Z</td>
<td>2 mm</td>
<td>0.6 mm</td>
<td>1 mm</td>
<td>0.4 mm</td>
</tr>
<tr>
<td>$A_{electr.}$</td>
<td>10 mm$^2$</td>
<td>8 mm$^2$</td>
<td>8 mm$^2$</td>
<td>3 mm$^2$</td>
</tr>
</tbody>
</table>
3.6 Fabrication

Micromachining technology has developed dramatically over the last years due to its versatile application in areas such as microsensors and MEMS. One of these technologies is bulk micromachining, which uses anisotropic wet etch and wafer bonding to build three-dimensional microstructures, with both lateral and vertical dimensions controlled within extremely small tolerances (less than 1 μm).

The design of the tunable FPMI requires the removal of silicon at well-controlled locations. The vertical dimensions are controlled by stopping the etch-process at the desired depth. At the same time, the lateral dimensions must be controlled, protecting with a photoresist mask the parts of the wafer that are to be maintained. Batch processing of the device facilitates mass production and provides, in principle, the potential for low unit cost.

Fig. 3-25 shows the cross section of the FPMI. The electrodes control the deflection of the movable SiN membrane/Si frame diaphragm and prevent the sticking of the upper mirror on the bottom mirror.
3.6 Fabrication

![Diagram of micromachined Fabry-Perot filter](image)

Fig. 3-25 *Cross section of the micromachined Fabry-Perot filter.*

### 3.6.1 Fabrication sequence

The whole fabrication process requires 5 masks (see Fig.3-26) [34]. On the same wafer, upper and bottom dies are fabricated. A 100 nm double-side polished wafer is used as starting material. First, 400 nm recesses are formed on a wafer front for which the LOCOS technique is used. Then a 300 nm low-stress (<0.15 GPa) LPCVD silicon nitride layer is deposited on both wafer sides and protected against damage during subsequent processing by a 300 nm LPCVD polycrystalline silicon layer.

Then PECVD oxide is deposited on a wafer front side with thickness (0.3-1 μm) corresponding to the required initial resonance cavity gap. The PECVD-oxide/poly-Si stack is patterned by wet etching to form spacers between upper and bottom dies for later die attachment.

The 300 nm Al interconnect and control/sensing electrodes (deposited by sputtering) are ‘buried’ in 400 nm recesses to increase initial spacing of the electrodes and avoid sticking during operation.

The poly-Si/nitride stack on the wafer back is patterned by means of dry plasma etching to prepare windows for anisotropic KOH etching.
Silver mirrors are evaporated and patterned using lift-off technique on the wafer front side. This is the last step before anisotropic KOH etching (33 wt% KOH solution at 85°C) performed in a sealed holder to protect the front side of a wafer with Al electrodes and Ag mirrors. The anisotropic KOH etching was performed using a sealed stainless steel holder. The etch rate depends on the temperature and concentration of the solution and, in this case, was 1.7 μm/min, which results in a total etch time of approximately 5 hours to etch through a 525 μm thick wafer. The etch of silicon is very anisotropic (see section 3.4.1). As a convex corner is
attacked from all sides with similar underetching, it was necessary to implement corner-compensation structures in order to achieve convex corners.

Subsequently, the SiN membranes were completely cleared from Si over the entire area. The wafer, while in the holder, was carefully taken out of the KOH solution and rinsed for 30 minutes in de-ionised water. The sealed holder was then opened and the wafer was separately rinsed for 30 minutes. The protective photoresist layer on the wafer front side was subsequently removed in acetone. The resulting structures (Si frame hanging on a SiN membrane) are very fragile and extreme care must be taken during further handling. Therefore, any 'aggressive' drying method is not allowed. Slow drying in a nitrogen flow resulted in stains on most of the membranes. These stains probably consisted of organic residues from the protective photoresist layer. Imersion in isopropyl alcohol after acetone cleaning gives an improvement. Obviously, another solution is to completely avoid the protective photoresist layer. However, the best solution will be deposition of the Ag layer at the very end of the fabrication sequence [24]. A SEM photograph of the surface morphology of an on-SiN evaporated silver layer is shown in Fig.3-27.

Fig. 3-27  *SEM photograph of the evaporated silver layer surface.*
3.6.2 Assembly of the FPMI

A press-on contact was used to provide connection (electrical ground) between the top die and the bottom die. A hole is formed on the top die using the KOH etch step, and large metal pads were patterned at the same position in both dies. When one applies a force with a needle, the upper contact will touch the lower contact and these stick together (see Fig. 3-28).

![Diagram](image)

**Fig. 3-28** *Top-view: of the bottom plate (left) and top plate (right).*

To facilitate dicing of the finished wafer into the individual dies, deep V-shaped trenches are formed around each die during anisotropic KOH etching. After careful cleaning, the bottom die is mounted on a print board and wire-bonded. The upper die is then attached and fixed by gluing.

Two square-shaped pyramidal openings at the periphery of the device (both upper and bottom die) are used during the assembly of the bottom and upper die to perform the manual alignment. The photograph of a fabricated FPMI is shown in Fig. 3-29.
3.7 Experimental results

The optical response was measured with a 5.1 mm² photodiode (Hamamatsu type S1336-5BQ) and HP 4142B DC source/monitor controlled by a HP 9000/700 computer. A 100 W tungsten lamp (Oriel 6333) and a 1/8 m Oriel 77250 monochromator with a ruled grating (Oriel 77298) were used as light source (see Fig. 3-30). The light from the exit slit of the monochromator was first collimated by a f=25.6 mm lens (Melles Griot 06GLC004) and focused onto the Fabry-Perot interferometer. A focusing lens with an 8 mm focal length (Melles Griot 06GLC002) focused the transmitted light in the photodetector. All lenses were corrected for spherical aberration, coma, and astigmatism.
3.7.1 Spectral response

The control voltages required for tuning the FPMI resonance cavity width and adjustment of the mirror parallelism were set manually. Firstly, the parallelism between the two mirrors is obtained through analysis of the interference pattern projected. Secondly, electrostatic actuation is applied to set the cavity gap width, while maintaining parallelism. Fig. 3-31 shows an example of the measured spectral response in transmittance (cavity gap ~500 nm) with FWHM of 12 nm, which is in reasonable agreement with simulation (Fig. 3-32). The difference in the values of transmittance between the measurements and the simulation can be due to losses of the transmitted light when focusing the beam in the detector, roughness surface, and the difference in the thickness of the SiN layer values from the simulation compared to the real situation (were used SiN membranes 500 nm thick).

In Fabry-Perot microinterferometers with four electrodes, it is possible to achieve parallelism between the two plates and to simultaneously correct for die bonding misalignment. The air gap that formed the cavity must be small to provide the largest air gap change for a given range of applied voltage.

Fig. 3-30 The experimental arrangement used to spectrally characterize the tunable FPMI.
3.7 Experimental results

Fig. 3-31 Optical transmittance measured for an air cavity gap of about 500 nm and silver mirrors with a thickness of 40 nm.

<table>
<thead>
<tr>
<th>Illuminant:</th>
<th>WHITE</th>
<th>Angle:</th>
<th>0.0 (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium:</td>
<td>AIR</td>
<td>Reference:</td>
<td>550.0 (nm)</td>
</tr>
<tr>
<td>Substrate:</td>
<td>AIR</td>
<td>Polarization:</td>
<td>Ave — —</td>
</tr>
<tr>
<td>Exit:</td>
<td>AIR</td>
<td>Remark:</td>
<td>FF(Ag-40,SiN-300,Air-4000) (nm)</td>
</tr>
<tr>
<td>Detector:</td>
<td>HPD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3-32 Simulated transmittance of a FPM1 using 300 nm-SiN/40 nm-Ag mirrors with an air gap of 500 nm.

A reflectance measurement was done without adjusting the parallelism between the two plates (for 300 nm-SiN/40 nm-Ag mirrors and an air gap
of about 1 μm), of which the result is presented in Fig. 3-33. In these measurements a focused beam with a diameter of about 100 μm was used.

![Graph showing reflectance vs wavelength](image)

**Fig. 3-33** Measured reflectance of a FPMI using 300 nm-SiN/40 nm-Ag mirrors with an air gap of about 1 μm (without adjustment of parallelism).

### 3.7.2 Stiction and hysteresis

At pull-in voltage the diaphragm snapped down, coming into contact with the bottom plate indicating the threshold voltage was reached. Hysteresis was observed in the motion of the diaphragm during the reverse voltage sweep. It was necessary to reduce the voltage drastically to release the diaphragm from position near to the displacement at pull-in. Charging effects result from surface charge build-up of the silicon nitride membrane. Using a silicon dioxide bottom plate instead a silicon nitride one eliminates this problem. Stiction was also observed when the diaphragm reached the bottom plate. Stiction is a common problem with movable micro-mechanical structures [21]. Voltages must be maintained below the threshold voltage to avoid contact. Polysilicon spacers were used for the same purpose.
3.8 Conclusions

We fabricated a tunable bulk micromachined Fabry-Perot microinterferometer to be used in an on-chip integrated scanning-type spectrometer (which includes FPMI, integrated photodiode, and read-out electronics). Optical measurements show that for perfectly parallel mirrors, transmittance of 10% approximately and a FWHM of less than 13 nm are achievable. The materials and device properties enable a FPMI with a finesse exceeding 30 and a FWHM smaller than 3 nm (with 50 nm silver mirrors and working at a high resonance mode).

It is difficult to achieve complete parallelism through manual adjustment of the voltages applied to the control electrodes. Therefore, a future version of the device should include a servo-control system (with distributed sensing electrodes for control of the cavity spacing), allowing automatic adjustment of the mirror parallelism and improving the response time.

The general disadvantage of this scanning-type microspectrometer is the mechanical movement. For this reason the array-type microspectrometer is analyzed in more detail in the next chapter.
References


3.8 Conclusions


Array-type microspectrometer

4

4.1 Introduction

In this chapter, an array-type microspectrometer based on Fabry-Perot etalons is described [1]. In contrast to the scanning-type microspectrometer from the previous chapter, this solution contains no movable parts. Using a standard bipolar process an array of photodetectors is fabricated. Subsequently, on top of each photodiode a Fabry-Perot etalon is integrated, using a simple post-processing module. The single-chip spectrometer is intended for operation at visible wavelengths and can easily be tuned during fabrication to cover a different spectral range.

4.2 Motivation

Previously developed microspectrometers [2], [3], [4], [5], [6], fabricated by means of bulk or surface micromachining, contain movable parts to perform wavelength tuning. As a result, these are less reliable and suitable only for operation in a limited spectral band (mostly near-IR) [6], [7]. Moreover, high-voltage electrostatic actuation is necessary for resonance cavity tuning.
Array-type microspectrometer

Operation in the visible spectral range (with transmission in a single peak) requires cavity length below 300 nm (see section 2.7.1.2 for details) [8]. Fabrication and modulation of such a narrow air gap between the two mirrors is severely hindered by capillary forces inside of the cavity. Also, electrostatic pull-in and subsequent sticking of the two mirrors limits the operating range of the device to one third of the initial air gap [9].

An array composed of fixed cavities of different widths has several operational advantages over devices that involve mechanical scanning.

4.3 Design

4.3.1 Selection of materials

According to the simulation results from chapter three, the best option in terms of optical characteristics for a Fabry-Perot resonance cavity for the visible range, is to use silver mirrors with an air gap between them. Because the array type has fixed cavities of different widths, an array of Fabry-Perot etalons (cavity medium different from air) were implemented. Often, a dielectric film such as glass or silica is the material used to serve as cavity medium. Two different dielectric materials were investigated: silicon nitride (SiN) and PECVD silicon dioxide (SiO2). SiN films are characterized by stronger dispersion and larger values of the refractive index (see Chapter 3 for details). The dependence of refractive index on the wavelength for the PECVD SiO2 films is more constant (varies from 1.47 to 1.46 between 400 nm to 800 nm wavelengths). Therefore, PECVD SiO2 thin film was selected (see the etalon structure in Fig.4-1).

Due to fabrication constraints the bottom mirror of the etalon will be a thin aluminum film instead of a silver one. Only the top mirror is a thin film of silver deposited at the last step in the fabrication process.
4.3 Design

Fig. 4-1 *The Fabry-Perot etalon with Al bottom mirror.*

Fig. 4-2 shows the simulated transmittance and finesse as a function of silver thickness for a Fabry-Perot etalon with two silver mirrors and a silicon dioxide layer with thickness 1 μm between them.

Fig. 4-2 *Simulated transmittance and finesse as a function of silver thickness for a Fabry-Perot etalon: Ag mirrors with a SiO₂ layer of 1 μm between them.*

4.3.2 Optical simulations of the etalon structure

A thin-film optics software package (see Chapter 3) was used for structural optimization of the FP filter. The transmittance of a 60 nm-Ag/
1000 nm-SiO$_2$/60 nm-Ag layer stack (Fig. 4-3) shows a FWHM of 1.8 nm and a finesse of 40. The Ag layer thickness is a trade-off between achievable FWHM and absorption loss. When aluminum is used for the bottom mirror (due to fabrication constraints), the performance slightly decreases due to higher absorption of aluminum films compared to that of silver films (see Fig. 4-4 and Fig. 4-5). A FWHM of 4.1 nm and a finesse of 18 are achieved. Fig. 4-6 shows a single peak for a cavity thickness of 300 nm. For a 45 nm silver mirror, an 20 nm aluminum mirror and cavity gaps between 225 to 300 nm in 5 nm increments, it is possible to cover the range 400 nm to 500 nm with an inter-channel shift of about 6 nm (see Fig. 4-7).

An oxide layer is present between the Fabry-Perot etalon and the photodiode; it introduces a wavelength-dependent transmission of the incident radiation. Its thickness was optimized for a flat transmittance over the visible spectra range (approximately 50 nm) [10] [11].

![Graph showing transmittance vs. wavelength for 60 nm Ag/1000 nm SiO$_2$/60 nm Ag layer stack.](image-url)
Fig. 4-4  Absorptance vs. wavelength for a 45 nm thick Ag and 20 nm thick Al thin-films. Materials data from Sopra S.A [12].

Fig. 4-5  Simulated transmittance vs. wavelength 60 nm Ag/1000 nm SiO$_2$/20 nm Al layer stack.
Array-type microspectrometer

Fig. 4-6  *Simulated transmittance vs. wavelength* 45 nm Ag/300 nm SiO₂ / 20 nm Al layer stack.

Fig. 4-7  *Simulated transmittance vs. wavelength* for 45 nm/SiO₂/20 nm Al layer stack. The SiO₂ layer thickness is used as a parameter and changes from 225 nm to 300 nm in 5 nm increments.
4.3.3 Phototransducers

Each channel in the microspectrometer is composed of a Fabry-Perot etalon with an opto-electric transducer underneath: a photon detector. This device converts light into electrical current through the photoelectric effect (to convert photons in electron-hole pairs). In a standard bipolar process the p-n junction photodiode is conveniently available (see Fig.4-8). The absorption of light in silicon is wavelength dependent [10]. Long wavelength light (low-energy photons) penetrates much deeper before being absorbed.

Isolation of individual photodiodes is used (Deep Boron diffusion, DP layer, see Fig.4-8). The top surface between different channels is covered with metal in order to shield stray light.

Limiting factors in the opto-electrical conversion in a photodiode are the size and uniformity of the effective sensing area, and dark current. Dark current is the current that flows in a photodiode when there is no optical radiation incident on the photodiode. It is usually measured and then subtracted from the flux. As the dark current is temperature dependent, one measurement at the beginning of the experiment is usually not sufficient. A dark-current-compensation channel can be implemented using a photodiode completely covered with metal.

The dominant noise source in a photodiode is shot noise. Since individual photoelectrons are created by absorbed photons at random intervals, the resulting signal has some variation with time. The variation of the detector current with time creates noise (due to the desired signal photons or by background flux).

Another form of noise is the fixed pattern noise. It is defined as the photoreponse variation between adjacent photodiodes. In this application it is significant, because we are dealing with an array of photodiodes. The fixed pattern noise is controlled by the size of the photodiode relative to the minimum lithographic feature size. As the dimensions of the photodiodes (1 mm²) are relatively large, this noise is negligible [13].
Array-type microspectrometer

Fig. 4-8 The schematic structure of the photodiode in cross section.

The cross-talk resulting from photon-produced minority-carrier diffusion is a problem in optical detection with more than one channel. The carriers generated outside the junction depletion layer may diffuse to a neighboring photodiode and thus introduce an undesired current component. In certain silicon processes the diffusion length of such carriers can be up to 100 µm [14]. So, this minority-carrier diffusion causes cross-talk in a closely-spaced photodiode array. Moreover, the problem cannot be solved by means of a light shield, because the carriers are primarily produced under a photodiode. The solution is to form a deep p-diffusion that contacts the p-substrate and is connected to ground potential. In this way the photocurrent in a diode is contributed to only by the photocarriers generated in or near the depletion layer between the p-diffusion and the n-diffusion regions. Minority carriers generated above or below the n-diffusion/p-substrate junction or diffusing from neighboring photodiodes will be swept away from the n-diffusion region by the junction field, or are bypassed to the power supply.

4.3.4 The complete structure

The entire structure of the array-type microspectrometer is schematically shown in Fig. 4-9 (an individual channel in cross section) and in Fig. 4-10 (the complete 4x4 array microspectrometer). The device fabrication is fully compatible with a standard IC bipolar processing, as the upper mirror of silver is the only non-compatible material used. This step is performed outside the standard processing line.
Fig. 4-9  *Schematic structure of the microspectrometer: individual channel in cross section.*

Fig. 4-10  *A 4x4 array microspectrometer in perspective. Each of the Fabry-Perot cavities is tuned to transmit in different spectral band.*

The impinging spectrum is filtered in the Fabry-Perot resonators and the intensity of the selected spectral component is measured in transmission using an underlying integrated photodiode array. On top of each photodiode an Al/SiO$_2$/Ag layer stack is deposited functioning as the Fabry-Perot optical filter.

### 4.3.5 Optical simulation of the complete structure

The complete structure (Fabry-Perot etalon plus photodiode underneath) was modeled by means of TFCalc. Simulated reflectance vs. wavelength of the complete structure was analyzed to investigate the influence of the oxide thickness between the etalon and the silicon substrate. Table 4-1 shows the values used for the refractive index and extinction coefficient of monocristalline silicon [15]. The thickness range of the oxide passivation
Array-type microspectrometer

layer was investigated between 50 and 300 nm. Fig. 4-11 shows, for a 50 nm oxide layer, higher transmittance in the region 450-500 nm than a 300 nm oxide layer. Also, the performance in the ultraviolet region is improved and an oxide thickness of 50 nm presents an almost wavelength-independent transmission over the entire visible part of the spectrum [11].

![Graph showing reflectance vs. wavelength for SiO₂ layers at 50 and 300 nm](image)

**Fig. 4-11** Simulated reflectance vs. wavelength of an etalon 45 nm Ag/300 nm SiO₂/20 nm Al layer stack for an oxide layer of 50 nm and 300 nm thicker underneath.

**Table 4-1** Measured dispersion of the refractive index and extinction coefficient of monocrystalline silicon [10].

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>Refractive index</th>
<th>Extinction coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>5.65</td>
<td>0.40</td>
</tr>
<tr>
<td>450</td>
<td>4.65</td>
<td>0.090</td>
</tr>
<tr>
<td>500</td>
<td>4.30</td>
<td>0.060</td>
</tr>
<tr>
<td>550</td>
<td>4.10</td>
<td>0.040</td>
</tr>
<tr>
<td>600</td>
<td>3.90</td>
<td>0.035</td>
</tr>
<tr>
<td>650</td>
<td>3.80</td>
<td>0.020</td>
</tr>
<tr>
<td>700</td>
<td>3.75</td>
<td>0.010</td>
</tr>
<tr>
<td>750</td>
<td>3.60</td>
<td>0.0095</td>
</tr>
<tr>
<td>800</td>
<td>3.55</td>
<td>0.0075</td>
</tr>
</tbody>
</table>
4.4 Fabrication

The photodiode array was fabricated using standard microelectronic bipolar processing. The Fabry-Perot etalons were added in a post-process module. The fabrication sequence is straightforward. Only the post-process module, which was used to fabricate the 16 etalons, each with a different thickness, requires further explanation. The formation of the Fabry-Perot etalon starts with the deposition of a 20 nm Al layer after completion of the bipolar process (including metallization and thinning of the oxide above the photodiodes). The oxide passivation layer on top of the photodetector (between the Fabry-Perot etalon and Si substrate surface) is thinned to reduce its influence on the spectral response. The thin Al layer is evaporated and patterned using lift-off. Subsequently, a PECVD oxide layer is deposited with a thickness equal to the maximum cavity length. The thickness of the PECVD silicon dioxide layer, which is enclosed in between two semi-transparent metallic mirrors, determines the wavelength for tuning. In N subsequent plasma etching steps (for which different photoresist masks are used), the initially deposited PECVD oxide layer is thinned that $2^N$ channels are formed, each with a different resonance cavity length. An example with four masks for a 4x4 array-type microspectrometer is shown in Fig.4-12. After the deposition of the silicon dioxide, each mask used has a different etching time (T, T/2, T/4, T/8). A relative thickness uniformity better than 2 nm was achieved between cavity lengths.

<table>
<thead>
<tr>
<th>Etch time</th>
<th>Mask1</th>
<th>Mask2</th>
<th>Mask3</th>
<th>Mask4</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4-12 Principle used for fabricating a 4x4 array microspectrometer. After the deposition of the silicon dioxide, each mask is used with a different etch time. In this example the first mask will etch one half of whole thickness, the second one fourth, etc.

A silver layer is deposited at the very end of the fabrication sequence and patterned using lift-off. Fig.4-13 presents a complete description of the bipolar postprocessing fabrication sequence.
Array-type microspectrometer

(a) Completed bipolar process.

(b) Thinning of the oxide above photodiode active area.

(c) Deposition of lower aluminum mirror by means of lift-off.

(d) PECVD oxide deposition and sequential thinning.

(e) Bond pad opening.

(f) Deposition of upper Ag mirror by means of lift-off (not included in the standard processing line).

Fig. 4-13  Post bipolar processing fabrication sequence.
Fig. 4-14 shows a SEM photograph presenting the cross section of one of the channels.

![SEM photograph showing the cross-section of one of the channels.](image)

Fig. 4-14  *SEM photograph showing the cross-section of one of the channels.*

Fig. 4-15 shows a photograph of the fabricated device with overall dimensions of 4.7 x 4.7 mm².

![Photograph of the fabricated micro-spectrometer with 4x4 channels. The die area is 4.7x4.7 mm².](image)

Fig. 4-15  *Photograph of the fabricated micro-spectrometer with 4x4 channels. The die area is 4.7x4.7 mm².*
4.5 Experimental results

The electrical characteristics and spectral responsivity were measured using an HP4142B DC Source/Monitor (full-scale range from $10^{-15}$ A to 1 A and a resolution of $10^{-13}$ A). A 100 W quartz tungsten halogen lamp with a monochromator TRIAX-180 (1200 g/mm grating with a spectral dispersion of 3.6 nm/mm and a spectral resolution of 0.3 nm at 546 nm), was used as light source. A collimator lens was used to image the light on the entrance slit. At the exit slit, a pinhole and a focusing lens were used to achieve a beam with a diameter of about 400 μm [16]. The measurements were calibrated with a commercially available photodiode, Hamamatsu S1336-5BQ.

System operation is demonstrated using a 16-channel micro-spectrometer designed for operation in the spectral range between 400 to 500 nm with inter-channel shift of about 6 nm. Each of the channels consists of a 20 nm-Al / SiO₂ / 45 nm-Ag layer stack, where the oxide layer thickness changes between 220 nm and 300 nm with 5 nm steps.

Optical spectral measurements in reflectance (using an external detector, Leica Microscope Photometer System) show that each of the channels is sensitive to only one narrow spectral band, with FWHM of 16 nm over the visible spectral range (see Fig.4-16). Independent measurements were performed on a cavity of 1000 nm thickness, which indicate a FWHM of 6 nm and a finesse of 12 (see Fig.4-17).

Subsequently, measurements were performed on the device with integrated photodiodes. A dark current of 0.2 pA at 0 V and 1 pA at -4 V (see Table4-1, conclusions) was measured for a photodiode active area of 1 mm². The dark current directly limits the signal noise ratio (SNR). The photocurrent, at a given source spectral power, is the result of a trade-off between sensitivity, selectivity, and the number of channels used. An increased Ag-layer thickness improves selectivity but causes higher absorption losses and low sensitivity, and thus decreases SNR.
4.5 Experimental results

![Graph](image)

**Fig. 4-16** Measured reflectance vs. wavelength for 45 nm Ag/\(\text{SiO}_2\)/20 nm Al layer stack. The \(\text{SiO}_2\) layer thickness is used as a parameter and changes from 225 nm to 300 nm in 5 nm increments.

![Graph](image)

**Fig. 4-17** Measured reflectance vs. wavelengths for a 45 nm Ag/1000 nm \(\text{SiO}_2\)/20 nm Al layer stack.

Fig. 4-18 presents the spectral responsivity (A/W) between 400 nm to 800 nm for all 16 channels using on-chip photodiodes. The ratio between
the base line and the peak maximum ranges from 4 to 7. The relatively high stray light, beam divergence and the roughness surface are responsible for the background signal. Stray-light compensation methods must be used in order to compensate for the non-idealities of both the incident light beam and the Fabry-Perot etalon.

Fig.4-19 shows the reflectance measured for 32 channels. In this case, the initial PECVD oxide thickness was 350 nm, and on every second die an additional 5\textsuperscript{th} etching step (time=T/16) was performed. Two adjacent dies (with and without the 5\textsuperscript{th} etching step) together form the 32-channel device. This technique can be extended further to yield a device with 64 or 128 channels.

![Graph showing spectral responsivity](image)

**Fig. 4-18** Spectral responsivity of the 16-channel microspectrometer for a 45 nm Ag / SiO\textsubscript{2} / 20 nm Al layer stack. The SiO\textsubscript{2} layer thickness is used as a parameter and changes from 225 nm to 300 nm in 5 nm increments.
Fig. 4-19 *Measured reflectance vs. wavelength for 32-channel microspectrometer for a 45 nm Ag / SiO₂ / 20 nm Al layer stack. The SiO₂ layer thickness is used as a parameter and changes from 225 nm to 350 nm in 4 nm increments.*

### 4.6 Conclusions

A single-chip spectrometer, using Fabry-Perot resonance cavities on a distributed photodetectors array, was fabricated and characterized (see Table 4-2). The advantage of the device presented is that it can easily be tuned during fabrication to cover different spectral bands, by adjusting the etching time only, without affecting the device layout. Such a device is extremely suitable for applications in microsystems (e.g. μTAS), because of its small size, high spectral selectivity, and low cost. Also, microspectrometers for the UV and IR are feasible with this technique.

The next chapter deals with an integrated optical microsystem based on this array type of Fabry-Perot etalons in CMOS that includes on-chip power management, signal-conditioning circuits, A/D conversion, and a digital bus interface for external/internal interface.
### Array-type microspectrometer

#### Table 4-2  Electrical and optical characteristics

<table>
<thead>
<tr>
<th>Feature</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>Bipolar</td>
</tr>
<tr>
<td>Device area</td>
<td>4.7 x 4.7 mm(^2)</td>
</tr>
<tr>
<td>FWHM @(t_{\text{SiO}_2}=300) nm</td>
<td>16 nm</td>
</tr>
<tr>
<td>Finesse @(t_{\text{SiO}_2}=300) nm</td>
<td>Single peak</td>
</tr>
<tr>
<td>FWHM (simulated) @60 nm Ag mirrors @(t_{\text{SiO}_2}=1) (\mu)m</td>
<td>1.8 nm</td>
</tr>
<tr>
<td>Finesse (simulated) @60 nm Ag mirrors @(t_{\text{SiO}_2}=1) (\mu)m</td>
<td>40</td>
</tr>
<tr>
<td>Dark current @ 0 V</td>
<td>0.2 pA</td>
</tr>
<tr>
<td>Dark current @ -4 V</td>
<td>1 pA</td>
</tr>
<tr>
<td>Spectral response maximum @ -4V</td>
<td>0.013 A/W@388 nm</td>
</tr>
<tr>
<td>Incident angle range</td>
<td>±10°</td>
</tr>
</tbody>
</table>
References


5.1 Introduction

The two previous chapters presented two concepts of spectrometer: scanning type and array type. The array-type device is directly suitable for monolithic co-integration with many different standard processes. CMOS offers the best solution for integration of complex electronic functions. Therefore, an on-chip integrated CMOS optical microspectrometer with a light-frequency converter and a bus interface is presented. Monolithic integration of a spectrometer with circuits features small size, low cost, maximum functional packaging density, and superior system reliability. This optical microsystem may be used for stand-alone operation. However, a more promising option is using Multi Chip Module (MCM) techniques to place the optical microsystem on an active silicon platform in order to build a microinstrumentation system. In this option the optical microsystem or the microspectrometer are merely components in the microinstrumentation system, which is, in turn, a part of a practical measurement-control system, as will be shown in this chapter.
5.2 Microinstrumentation system

The development of integrated smart microsystems which include sensors, microactuators, low-power signal-processing circuits, microcontrollers, and an interface for the external world using a wireless or hardwired link (e.g. RS-232, RS-485) promises to have a significant impact in many innovative products, as these can be applied in such diverse areas as industrial process automation, automotive systems and optical systems [1].

5.2.1 Concept

A microinstrumentation system, shown schematically in Fig. 5-1, includes all components of a complex measurement system on the smallest possible material carrier: a silicon chip.

![Block diagram of a microinstrumentation system for high-resolution spectrometer system.](image-url)
The microinstrumentation system is a general concept, but within the scope of this thesis it is designated as an option to transform a relatively simple single-chip 16-channel microspectrometer into a high-resolution spectrometer.

The microinstrumentation system is composed of a universal platform which is to be populated with the required sensors, microactuators, and a microcontroller using MCM (Multi Chip Module) techniques. The microinstrumentation system is built around an embedded MCU (Micro Controller Unit) (see Fig. 5-2) having on-chip memory, timers, serial communications facilities, low-power modes, and, when required, incorporates an A/D converter.

The sensors are scanned periodically by the MCU (based on the variations in the parameters measured), or preferably by the use of interrupt and service requests in order to minimize power consumption. Also, the digital compensation of the data is strongly sensor dependent and is chosen such that processing time, and MCU memory code are minimized.

![Diagram of microinstrumentation system](image)

**Fig. 5-2** A microinstrumentation system based in an active platform with attached transducers and MCU.

To control power consumption between sensor scans, the MCU stays in sleep mode. The power is removed from all idle devices and only the
sensors with permission to generate service and interrupt request are able to wake the system from its sleep mode. The temperature sensor has high priority to interrupt the system when the platform temperature reaches a threshold.

The internal bus of the microinstrumentation system is a serial sensor bus intended for use in small data-acquisition and control systems, aiming at an efficient and low-cost interconnection of sensors, actuators, and bus masters.

5.2.2 Microinstrumentation bus interface

This bus interface chip was designed for use in a multi-chip-composed microinstrumentation system (internal/external bus interface). Power consumption at 5 V@100 kHz is less than 500 µW. For 5 V@4 MHz power consumption amounts to less than 2 mW, due to smart power management of all functional blocks. The bus interface is able to transmit a digital code, bitstream, analog voltage, frequency, duty cycle and is also suitable to support calibration procedures, and service and interrupt request for the smart sensors or microactuators.

Present microsystems require a sophisticated local bus to address sensors, microactuators, and controllers. Available bus protocols lack the flexibility that is needed to deal with a multi-sensor system on the die level. Suitable buses like I²C (having a too complex protocol) and the basic Integrated Smart Sensor bus (ISS-bus) [2] do not provide calibration, test facilities, nor interrupt mechanisms. A microinstrumentation system contains an active silicon platform, including all infrastructural functions, such as power/thermal management of the system, test facilities, and an on-system sensor bus protocol that allows flexible and low-power data handling using interrupt [1].

The bus interface should be versatile enough to ensure efficient communication between all sensors and systems on the platform, however, it should be simple enough to be integrated on-chip within the platform [3]. As an additional feature, the bus interface should be able to handle both digital and semi-digital signals, such as pulse width and frequency-modulated pulse series. Moreover, the bus allows self-test by using analog excitation signals and simultaneous semi-digital or digital readout.
5.2 Microinstrumentation system

The upgraded version of the ISS bus interface presented here is based on a single controller to coordinate the activity on the bus and is characterized by: a maskable interrupt mechanism, calibration facilities, small size, and low power consumption, which makes it very suitable for implementation on microsystems.

Implementation of an interface bus in a complex structure like a microsystem requires techniques to reduce the power consumption. As the circuit density on chips and systems continues to increase, the difficulty of providing adequate cooling either adds significant cost to the system or limits its functionality [4]. Assuming that short-circuit current, glitching, and leakage can be kept in bounds through a good design, the dominant power source in CMOS is the dynamic power consumption.

Functional circuit design

Apart from simplicity, the improved ISS bus interface has two convenient features, which makes it very suitable for a microsystem. First, analog data can be transferred over the bus. Data generated by a sensor with limited signal-processing capability are analog usually, so this requirement is necessary, but not present in the usual standard interfaces, which are designated to interconnect only digital subsystems. Secondly, the use of the Manchester encoding scheme for transmission of the data at the logical level adds further flexibility. In such a scheme, the clock is embedded into the data allowing four logical levels instead of two (see Fig. 5-3).

The physical structure consists of two wires, a data line and a clock line, both open drain driven. The data line allows half-duplex communication between the modules connected to the bus. In order to increase flexibility, we added a second data line to be used only in case of duplex transmission (e.g. in case of on-line sensor calibration or a testing procedure).

In case of the embedded systems, particularly for instrumentation systems, sensor modules should also be able to signal announce to the controller when data is available, or more generally, when some particular event has happened. In the realized interface this flexibility is obtained by adding an interrupt request and a service request protocol.

An interrupt request message, if it is not disabled by the configuration used for that particular module, can be sent over the bus at any moment, even if the controller is in the middle of another conversation. A service request
on the other hand, is allowed only if the bus is in the idle state. All messages excluding the request messages are initiated by the controller.

```
Clock

Data

"1"  "0"  "Error"  "Idle"  "Bus Free"
```

Fig. 5-3  Manchester encoding scheme.

The length of the frame is variable, and depends on its significance. The structure of the bus interface is shown schematically in Fig. 5-4.

Fig. 5-4  The block diagram of the bus interface.

Smart power management

The interface bus is composed of some synchronous blocks. Especially those that are pipelined have a significant portion of the total power dissipated by the clock (responsible for 50% of the total power dissipated) [5]. Waste of power due to the clocking of blocks which are idle for a
significant period of time in normal or standby modes must be avoided. To manage the power consumption, we implemented two modes:

- **Selective shut-down of different blocks based on the level of activity required to run a particular application.** Different blocks of the chip may be idle for a certain period of time when different applications are running (this happens with the service and interrupt request blocks).

- **An idle mode for reducing power dissipation in the standby mode.** Wake up is initiated only at the start of frame to enable verification of the address.

**Features**

The features of the microinstrumentation bus interface are:

- simplicity of structure;
- only two communication wires are used in the minimum configuration;
- reliable data transfer by using the Manchester encoding with error-detection schemes;
- flexibility of signal type, as synchronous and asynchronous transmission of digital data is possible in combination with semi-digital signals, such as bitstreams, or even analog signals;
- flexibility of signal handling based on a maskable interrupt mechanism;
- sensor self-test capability over the bus using separate directional data lines, to be used as a separate die in microsystems and also suitable for on-chip integration with sensors;
- smart power management for reducing power consumption [6].

These functions have been realized in a 1.6 μm n-well CMOS process. A microphotograph of the bus interface (1.5x0.7 mm²) is shown in Fig. 5-5 [7], [8].
5.2.3 Possible configurations for optical microsystems

An optical microsystem (described in next section) including a bus interface module can be plugged with a MCU on a passive/active silicon platform (see Fig. 5.6). Also, the bipolar array type described in Chapter 4 can be glued on an active silicon platform and connected to a bus interface module in order to send the analog signal to the MCU for A/D conversion. These two configurations allow more than one module glued on the platform (until 16, according to the four bits to address).

Fig. 5-6  a) Several bipolar array-type spectrometers on a silicon platform. b) Several optical microsystems joined together with a MCU on a silicon platform.
5.3 A CMOS on-chip integrated optical microsystem

5.3.1 Monolithic vs. hybrid integration

In the previous section, a trend towards assembly on a single substrate of different modules (sensors, actuators, and electronics) was shown; such MCM techniques are widely used for hybrid integration and are an important alternative to on-chip co-integration of circuits and sensors in a single chip [9]. The subtleties involved in the choice of one of these alternatives are not widely recognized and deserve some discussion.

The pro-integration arguments are: small size and low weight, reliability, cost in high volume production could be lower than for a multi-component sensor, temperature compensation could be better because circuit and sensor both have the same temperature, reduced parasitics (very important for capacitive sensors), and increased immunity to Radio Frequency Interference (RFI) due to small dimensions and short length of the signal paths. Especially in the case of distributed systems, which require an array of devices to be sensed or actuated, the number of bonding wires can become too large and impractical; integrating both parts into one chip eliminates this problem.

The counter-integration arguments are strongly dependent on a particular application. To merge analog and digital circuits and a sensor increases the cost of the design and of the required modeling of the physical performance of the sensor. These costs can be supported only for very high-volume applications. Process compatibility is by far the most serious technical problem for monolithic integration. However, in some cases it is possible to divide the process into two parts, while the extra processes inserted in the sequence of the standard IC process steps have no effect on the performance and characteristics of the already integrated circuit.

In the case of the hybrid solution, there is much more freedom in maximizing the performance of the sensors or actuators, as the micromachining is not restricted by compatible temperatures, processing steps and materials. The electronic part can be fabricated in the optimized technology (CMOS, bipolar or BiCMOS).

When the microspectrometer is used in applications that are moderately demanding on resolution, such as (bio)medical devices (sometimes implantable), chemical analysis based on optical absorption, automated
color measurements on-line in an industrial process line, a CMOS on-chip integrated optical microsystem [10] has huge potential. However, when high resolution is required, a hybrid system seems more appropriate. Section 5.2.3 showed that, conveniently, a set of 16-channel devices can be included in a microinstrumentation system.

5.3.2 A CMOS single-chip microspectrometer

The optical microsystem is composed of an array-type spectrometer (presented in chapter four) [11], readout circuits, a light-frequency converter, and a bus sensor interface (see Fig. 5-7). The optical quality and long-term stability of the scanning-type spectrometer (tunable device presented in Chapter 3) is lower than that of (fixed-cavity) Fabry-Perot etalons (Chapter 4). Therefore, an array-type spectrometer was selected for the integrated optical microsystem.

Compatible CMOS post-processing was used to fabricate the array of Fabry-Perot etalons, tuned for a different resonance in the visible range. Five external connections (including $V_{dd}$ and $V_{ss}$) are sufficient for operation of the microsystem. Frequency output and a serial bus interface allow easy multi-sensor, multi-chip interfacing using a microcontroller or an external port of a personal computer (e.g. parallel port).

![Array of non-tunable Fabry-Perot cavities](image)

Fig. 5-7 A single-chip optical microsystem based on an array-type spectrometer of Fabry-Perot etalons integrated with photodiodes underneath the 16 etalons, readout electronics, and the bus interface.
5.3.3 Fabry-Perot etalon

CMOS postprocessing consists of depositing an Al/SiO$_2$/Ag layer stack on top of each photodiode after the CMOS process has been completed (integrated-circuit fabrication). This stack functions as a tuned Fabry-Perot resonance cavity. This type of etalon was described in Chapter 4 in detail. The two differences between this etalon and that of Chapter 4 are that compatibility with a CMOS process is required, rather than with a bipolar process, also the photodiode should be fabricated in a CMOS process. Fig.5-8 shows the Fabry-Perot etalon structure plus the photodiode.

![Diagram of Fabry-Perot etalon with photodiode](image)

Fig. 5-8 *CMOS Fabry-Perot etalon with the photodiode underneath: a cross section.*

5.3.4 The photodiodes

Each Fabry-Perot etalon has an associated CMOS photodetector. The photodetectors have been realized in a conventional 1.6 μm n-well CMOS process. In an n-well CMOS process, has three different types of photodiodes and one vertical phototransistor available [12]. The three different photodiodes result from the n-well/p-substrate junction, the n-well/p-diffusion, and the p-substrate/n-diffusion junction. The device used for photodetection is the vertical pnp phototransistor (Fig.5-8), with the deep junction formed by the p-epilayer and the n-well, and the shallow junction formed by the n-well and a p+ implanted layer that is normally used for the drain/source contacts (sp). Both junctions are connected in
A CMOS optical microsystem

parallel and operate at reverse bias. The upper junction contributes little to the generated photocurrent, especially at long wavelengths. Because of its large capacitance per area, this junction is mainly used for photocurrent integration. The sensors are arranged in a 4x4 array of square photodiodes with an active area of 500x500 μm² each. A typical dark current of 30 fA (12 pA/cm²) with both junctions reverse biased at 5 V in parallel was measured.

Cross-talk reduction techniques described in section 4.3.3 of Chapter 4 are implemented so that the operating current comes from the photocarriers generated in or near the depletion layer between the p-diffusion/n-well junction. The utilization of the three-layer photodiode is an effective approach to improve photodiode performance.

5.3.5 Readout electronics

Low-noise read-out of photodiodes [13] has been obtained to cover a 10^5 dynamic range of light intensity level. The photocurrent-to-frequency circuits developed feature a dynamic range of sensitivity that is comparable to that of more complex analog circuits commercially available [14].

Fig. 5-9 shows the block diagram of the read-out circuit. Only one photodiode can be connected to the comparator at a time using multiplexer \( S_1 - S_N \). At a voltage \( V_j \) lower than \( V_{ref} \), the comparator output remains at a high logic level and, after synchronization with the clock, closes switch \( S_{ch} \), which forces the photodiodes to be quickly charged during one clock period to \( V_{dd} \). When this switch opens again, the photocurrent discharges the junction capacitance \( C_{j1} \) and \( C_{j2} \), until the comparator detects \( V_j < V_{ref} \), which causes the cycle to repeat. The flip-flop chain divides the signal by two, resulting in a symmetrical output signal that can be transferred to the bus via a buffer.

Basically, the photodiode readout circuit can be considered a first-order (or relaxation) oscillator circuit and, since the circuit has only one pole (or frequency controlling element), it can be tuned over a very wide range [10]. The charge generated by the photoelectric effect directly modulates the charge in the integrating junction capacitance, thereby modulating the output frequency.

The sensitivity of the current-to-frequency conversion is linear and can be calculated as \( S_{if} = f_0 / I_{ph} = 1/(2\Delta Q) \) since the clock frequency is much
higher than the oscillator frequency. With ΔQ, the variation of the charge across the two junctions, depending on the junction capacitances of the photodiode. The total light-to-frequency conversion factor is given by the product of S_{if}, the spectral responsivity of the photodiode and the transmittance of the Fabry-Perot filter.

The charging switch S_{ch} is based on a complementary pair with equal-size PMOS and NMOS transistors to minimize capacitive charge injection at high oscillation frequencies. The currents in the multiplexer switches S_{1}-S_{N} are sufficiently low to allow a relatively high channel “on” resistance. The leakage current in the “off” state should be neglected compared to the dark current of the diode. The channel of these MOS transistors has a W/L of 2 μm/16 μm.

![Fig. 5-9 Photodiode-readout circuits.](image)

### 5.3.6 Optical microsystem bus interface

Fig. 5-10 shows the block diagram of the optical microsystem bus interface. The structure is basically the same as the one presented for the general microinstrumentation system shown in Fig. 5-4. It differs in the sense that the third bus line is used to put the frequency output after the light-frequency conversion on the bus. Also, the use of a different clock...
A CMOS optical microsystem

signal for the light-frequency conversion instead of the clock signal of the bus (SCL line) is allowed.

![Diagram of CMOS optical microsystem bus interface]

Fig. 5-10  Block diagram of the optical microsystem bus interface.

5.4 Fabrication

The electronic circuits and photodetectors were realized in a conventional double-metal, single-polysilicon 1.6 μm n-well CMOS process. After completion of the standard CMOS process, except for the last silicon nitride deposition step for scratch protection, a postprocessing module was used to build the Fabry-Perot etalons on top of the photodetectors. This etalon consists of a thin-film stack (silver/PECVD oxide/aluminum).

Fig. 5-11 presents a complete description of all CMOS postprocessing sequence steps.
Fig. 5-11  CMOS postprocessing fabrication sequence.
A CMOS optical microsystem

The photodetector was formed in the n-well with the p-epilayer by means of shallow boron implantation. Since the oxide layer on top of the photodiode is rather thick just after the CMOS process, this oxide layer was thinned to 50 nm to minimize the effect of the oxide layer on the transmittance. Subsequently, the stack is built similar as in the bipolar process described in Chapter 4.

A microphotograph of the completed chip is shown in Fig. 5-12. The die measures 4.2 mm by 3.9 mm. The analog circuits can be seen in the upper part: a sensor array, analog switches, a test diode, a metal-covered diode (for dark-current compensation), a reference circuit, a reference capacitor, and a comparator. The lower part holds the bus interface, the multiplexer, and some other digital circuits. Only four external connections to the chip are strictly needed: $V_{dd}(+5V)$, ground, the clock input SCL, and the bidirectional dataline (SDL), since this line can also be used for transmission of the frequency output. The other pads are the chip address pins and pins for testing purposes.

![Fig. 5-12 A photograph of the optical microsystem.](image-url)
5.5 Experimental results

A device is addressed, via an ISS bus interface by a conventional 8-bit microcontroller. A standard internal data-acquisition PC card can be also used for this purpose. A typical addressing sequence transmitted serially over the data bus has been recorded and is shown in Fig. 5-13. The frame transmitted was composed of a start bit, four bits for the device address and four more bits related to the internal sensor configuration. A network with three bus interfaces was implemented and the request block was successfully tested.

A typical service-request sequence transmitted serially over the data bus has been recorded and is shown in Fig. 5-14. When clock and data lines are idle, the sensor demands a service request by lowering the data line, and in reply the master issues a request for the address sensor.

Fig. 5-13  A plot of clock and data lines from the oscilloscope.
Fig. 5-14  A plot of service request proceeding.

The optical microsystem bus interface frames use eight bits for addressing. The four most significant bits are used for addressing the chip, so up to 16 chips can be addressed. The four least significant bits are used to select one of the 16 photodiodes. After selection, the corresponding sensor places its output frequency over the data asynchronous bus line. This sequence can be seen from the oscilloscope plots of Fig.5-15 and Fig.5-16.

Fig. 5-15  Oscilloscope plot of the bus signals for a high light level.
Fig. 5.16 Oscilloscope plot of the bus signals for a low light level.

Finally, the specifications of the single-chip spectrometer are listed in Table 1.

Table 5-1 *Optical microsystem test results.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Test result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating voltage</td>
<td>--</td>
</tr>
<tr>
<td>Power dissipation @ 1MHz</td>
<td>Clk= 1 MHz</td>
</tr>
<tr>
<td>Power dissipation @ 100 kHz</td>
<td>Clk= 100 KHz</td>
</tr>
<tr>
<td>Max clock frequency</td>
<td>--</td>
</tr>
<tr>
<td>Dark frequency</td>
<td>25 °C</td>
</tr>
<tr>
<td>Free spectral range</td>
<td>--</td>
</tr>
<tr>
<td>FWHM</td>
<td>--</td>
</tr>
<tr>
<td>Responsivity (no FP etalon)</td>
<td>$\lambda=480$ nm</td>
</tr>
<tr>
<td>Sensitivity (no FP etalon) [14] ref. TLS230 from Texas Instruments</td>
<td>$\lambda=670$ nm</td>
</tr>
</tbody>
</table>
5.6 Conclusions

A single-chip CMOS optical microspectrometer containing an array of 16 addressable Fabry-Perot etalons (each with a different resonance cavity length), photodetectors and circuits for readout, multiplexing, and driving a serial bus interface was fabricated. The result is a chip that can operate using only five external connections (including $V_{dd}$ and $V_{ss}$) covering the optical visible range. The frequency output and the serial bus interface allow easy implementation on a microinstrumentation system. Power consumption is 1250 $\mu$W for a clock frequency of 1 MHz.
5.6 Conclusions

References


A CMOS optical microsystem


Discussion and conclusions

6

6.1 Introduction

This chapter summarizes the results achieved with the two different types of microspectrometers presented in this thesis and discusses the advantages and drawbacks involved. Moreover, performance is compared with other microspectrometers described in literature. Finally, the conclusions along with recommendations for future work are given.

6.2 Scanning-type vs. array-type microspectrometer

The scanning-type microspectrometer has the advantage of being tunable in a certain range using e.g. electrostatic actuation [1]. But achieving parallelism between the two mirrors requires complex position servo control, using multi-spot capacitive sensing and multi-spot electrostatic actuation, giving a response time >1 s. Also, electrostatic actuation requires an operating voltage exceeding 20 V. A possible solution is to reduce the cavity gap (lower voltage for the same displacement). However, applications demanding cavity thicknesses of less than 400 nm (first order for the visible part of the spectrum) are complicated with this type of
device (stiction between the two mirrors). Table 6-1 presents a qualitative comparison between the two spectrometers.

Table 6-1  *Scanning-type vs. array-type spectrometer.*

<table>
<thead>
<tr>
<th></th>
<th>Scanning</th>
<th>Array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tunable</td>
<td>++</td>
<td>--</td>
</tr>
<tr>
<td>Voltage for electrostatic actuation</td>
<td>--</td>
<td>n.a.</td>
</tr>
<tr>
<td>Parallelism and stability between the two mirrors</td>
<td>-</td>
<td>++</td>
</tr>
<tr>
<td>Assembly of the microspectrometer</td>
<td>-</td>
<td>n.a.</td>
</tr>
<tr>
<td>Fabrication process</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Cavities below 400 nm (first mode operation)</td>
<td>-</td>
<td>++</td>
</tr>
<tr>
<td>Response time</td>
<td>--</td>
<td>++</td>
</tr>
<tr>
<td>Mirror flatness</td>
<td>-</td>
<td>++</td>
</tr>
<tr>
<td>Device area*</td>
<td>+</td>
<td>--</td>
</tr>
<tr>
<td>Resolution</td>
<td>+</td>
<td>++</td>
</tr>
</tbody>
</table>

*The necessary area for an array-type device to cover the same spectral range of a scanning-type device.

n.a. - not applicable

Both devices are fabrication compatible with standard VLSI technology. This allows manufacturing of the device in great quantities, thus offers the potential for low unity cost of the sensor.

6.3 *Performance comparison of scanning-type devices*

Table 6-2 provides a comparison between the performance of the device presented in this thesis and the performance of two previously published devices: a cantilever Fabry-Perot resonator [2] and a Distributed Bragg Reflectors (DBR) Fabry-Perot resonator [3] for the infrared region. These devices are both tunable and realized on a silicon substrate. In the first chapter, these devices were presented, as examples of microspectrometers.
6.4 Performance comparison of array-type devices

Table 6-2  Micromechanical Fabry-Perot performance comparison

<table>
<thead>
<tr>
<th></th>
<th>Tuning sensitivity</th>
<th>Finesse</th>
<th>FSR</th>
<th>Air gap</th>
<th>Number of layers required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning</td>
<td>140 nm</td>
<td>7</td>
<td>84 nm</td>
<td>1.2 µm</td>
<td>2</td>
</tr>
<tr>
<td>(2.8 mm x 2.8 mm)</td>
<td>23 V</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cantilever</td>
<td>120 nm</td>
<td>2.4</td>
<td>60 nm</td>
<td>6.27 µm</td>
<td>3</td>
</tr>
<tr>
<td>(101 µm)</td>
<td>53 V</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DBR-FP</td>
<td>60 nm</td>
<td>35</td>
<td>375 nm</td>
<td>3 µm</td>
<td>19</td>
</tr>
<tr>
<td>(infrared)</td>
<td>65 V</td>
<td></td>
<td>(calculated)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The tuning sensitivity is defined as the voltage required to obtain a certain transmission wavelength shift. Our device exhibits a larger tuning sensitivity: 140 nm @ 23 V, but at a smaller air gap. The low sensitivity of the DBR FP device is probably due to the many stiff DBR layers in the movable mirror structure. Finesse (FSR/FWHM) is used as the figure of merit for the Fabry-Perot devices, and the DBR FP presents the best performance. However, our scanning-type microspectrometer has a simple fabrication process (only two layers, as compared to the nineteen layers for the DBR FP) and can easily be integrated with circuits or other devices. Increasing the thickness of the silver mirrors to 50 nm and the air gap to 3 µm would lead to a finesse equal to 35, but measurements would suffer from a decreased light transmission.

6.4 Performance comparison of array-type devices

The performance of the array-type microspectrometer presented in this thesis [4] is compared with that of the low-cost spectrometer [5] (presented in the first chapter) and with a commercially available spectrometer, (Spectronic 20 [6]). Both spectrometers are based on a grating element and use a CCD detector.
Discussion and conclusions

Table 6-3  Array type performance comparison.

<table>
<thead>
<tr>
<th></th>
<th>FWHM (nm)</th>
<th>FSR (nm)</th>
<th>Accuracy (nm)</th>
<th>Resolution</th>
<th>Dispersive element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Array type</td>
<td>16 1.8</td>
<td>400-520</td>
<td>±3</td>
<td>25*</td>
<td>FP etalon</td>
</tr>
<tr>
<td>(this thesis)</td>
<td></td>
<td>400-472</td>
<td></td>
<td>222**</td>
<td></td>
</tr>
<tr>
<td>Low-cost</td>
<td>9.1 @632.8</td>
<td>450-750</td>
<td>2.55</td>
<td>69.8</td>
<td>Grating</td>
</tr>
<tr>
<td>Spectronic</td>
<td>20</td>
<td>340-950</td>
<td>±2.5</td>
<td>47.5</td>
<td></td>
</tr>
<tr>
<td>20 [6]</td>
<td></td>
<td></td>
<td>@950 nm</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*First mode operation. **Fourth mode operation and R=λ/FWHM.

The free spectral range in the array-type device is limited by the low quantum efficiency of the photodiodes at low wavelengths. The high resolution in Fabry-Perot based spectrometers is obtained for high interference orders. In our case a resolution of 222 is achieved in mode q=4. But for a single peak, mode q=1, a reasonable resolution of 25 was measured.

The technique used for fabrication of the Fabry-Perot etalons allows, for simple expansion of the number of channels to 512 or 1024, thus increasing the operation range.

6.5  Surface roughness

Atomic force microscopy (AFM) was employed to examine the surface morphology of the realized Fabry-Perot structures. T. Kwa et al. [7] measured a typical root mean square (RMS) roughness of 2.68 nm for the SiN membrane (~500 nm thick) after KOH etching. The mirror surface in the scanning-type microspectrometer has an approximate RMS roughness of 6.4 nm (see Fig.6-1). The RMS roughness of the surface mirror in the array-type microspectrometer is typically 15 nm (see Fig.6-2). A possible explanation for the high value compared to the scanning type can be the back etching applied to the PECVD oxide in order to define the cavity gap.
An etalon with silver mirrors (45 nm thick) and 300 nm PECVD oxide between them presents a typical RMS roughness of 8.6 nm (see Fig. 6-3).

Fig. 6-1  AFM surface scan of a 40 nm Ag layer evaporated on 500 nm SiN membrane (scanning-type device mirror surface).

Fig. 6-2  AFM surface scan of a 50 nm/350 nm PECVD oxide/20 nm Al/250 thermal-oxide layer stack (array-type device mirror surface).
Discussion and conclusions

Fig. 6-3  *AFM surface scan of an Fabry-Perot etalon with two silver mirrors (45 nm thick) and resonance cavity of PECVD oxide (300 nm thick).*

A further improvement of the mirror flatness can be done by optimizing the deposition processes [7], [8]. As evaporation is a directional deposition technique, lateral thickness variations are caused by lateral variations in impinging material dose. To ensure greater degree of uniformity, the wafers are often moved in various ways above the source. A high degree of uniformity can be achieved by a circular motion with two separated axes of rotation, one fast and the other slow [9]. Due to the small dimensions of the mirrors, the thickness non-uniformity is generally not as critical as the layer roughness.

6.6 Conclusions

6.6.1 Scanning-type and array-type microspectrometer

Two different types of lensless microspectrometers for use in the visible spectral range have been realized in silicon by means of conventional IC-process and bulk micromachining. The wavelength-selecting element used was a Fabry-Perot optical resonance cavity. A good mirror-quality micromachined surface was achieved using silver.
6.6 Conclusions

The scanning-type device is based on a two-wafer design and was done using bulk-micromachining technology: 1) a top wafer, which is a Si-frame/SiN-membrane composite diaphragm actuated electrostatically, and 2) a passive bottom plate with a highly reflective coating (no microelectronic processing).

The array-type device is fabricated in two steps and a special modular approach was used to make fabrication suitable for complete integration of electronics and sensor in a single chip: 1) a standard IC fabrication process (bipolar or CMOS) to fabricate microelectronics devices, and 2) a low-temperature postprocessing module for the formation of Fabry-Perot etalons.

Both microspectrometers can be applied in the visible part of the spectrum with reasonable resolution (6-10 for the scanning type and 25-222 for the array type) within the constraints of the available optical path length.

6.6.2 Single-chip CMOS optical microsystem

The ultimate result of the research presented in this thesis is the design and fabrication of the first complete CMOS single-chip optical microsystem with readout electronics, light-frequency converter, and bus interface. This optical microsystem features small size, low cost, maximum functional packaging density, and superior system reliability. The frequency output and serial bus interface allow easy multisensor, multichip interfacing using a microcontroller or the parallel port of a personal computer. Only four external connections, including $V_{dd}$ (+5 V), $V_{ss}$ (ground), and clock line (SCL), are necessary, since the bidirectional data line (SDL) can also be used for transmission of the frequency output.

6.6.3 Applications

Equipment for optical spectral analysis already finds widespread application in science and industry. Miniaturized spectrometers reveal potential for application in: industrial automation, automotive systems, biomedicine, environmental monitoring, emission line characterization, and chemical analysis by optical absorption. Fig.6-4 illustrates the chemical analysis of human blood for the detection of the quantity of hemoglobin present [10], [11]. Also, the characterization of powder mixtures in fluidized beds or in stirred tanks [12] (to determine the time necessary to obtain a homogeneous mixture) and color determination (of,
Discussion and conclusions

e.g., the ripeness of fruit) are applications where a microspectrometer would be useful. Miniaturized spectrometers are now being used in paint stores to enable provision of custom-mixed paint [13]. Therefore, the microspectrometer could be applied in any application in which the optical spectrum is a characteristic feature.

![Graph](image url)

Fig. 6-4 Spectral reflectance of human blood [10].

6.7 Future work

Background signal compensation

The Fabry-Perot etalons suffer from background signal (see chapter 4 Fig. 4-18) due to stray light, beam divergence and roughness surface. When an ideal plane wave of normal incidence interacts with an ideal Fabry-Perot resonance cavity, only a narrow spectral band around the resonance wavelength is transmitted (see Fig.6-5). But any imperfection of the incident light wave (components with non-normal incidence, stray light) and the Fabry-Perot etalon itself (pinholes, scattering at the rough surface of the mirror) causes increased transmittance outside the narrow resonance band to which the Fabry-Perot etalon is tuned (see Fig.6-6). This results in an increased parasitic background signal. Compensation or a significant reduction of the average background signal level must be implemented to improve the spectral selectivity of the Fabry-Perot etalons. A novel compensation method [14] is under investigation.
6.7 Future work

Fig. 6-5 Interaction of an ideal light beam with an ideal Fabry-Perot optical filter.

Fig. 6-6 Non-idealities of the incident light beam and the Fabry-Perot structure itself causes a large background light level in the entire spectrum band.
Discussion and conclusions

CMOS optical microsystem

The CMOS optical microsystem is a prototype and optimization of the layout (including the elimination of all pads for testing) will reduce the chip area by one third [15]. Also, the response of the photodiodes in the short-wavelength region could be improved.

Scanning-type microspectrometer

The scanning-type microspectrometer also has the potential to be used for other applications outside the optical domain, e.g., as optical pressure sensor, accelerometer, displacement and temperature sensors.
References


Discussion and conclusions


Summary

Equipment for optical spectral analysis is widely applied in science and industry. Optical microsystems, e.g. microspectrometers, merged with microelectronics, promise new applications. This thesis describes the design, fabrication and characterization of two different types of microspectrometers - the scanning type and the array type - for application in the visible part of the spectrum (390-760 nm). A fully integrated optical microsystem (chip size of 4.2 mm by 3.9 mm), based on a sixteen-channels array-type configuration, with electronic circuitry on-chip (readout electronics for the photodetectors, light-frequency converter and a bus interface) is presented.

The wavelength-selecting element used is a Fabry-Perot resonance cavity, acting as an interferometer. The Fabry-Perot interferometer consists of two semi-transparent parallel mirrors separated by a gap. The optical transmission characteristic of such an element consists of a series of sharp resonant transmission peaks when the gap equals multiples of a half wavelength of the incident light. The filtered spectrum is projected onto the integrated photodiode, where the presence and intensity of the spectral components is detected.

Both microspectrometers are fabricated using conventional IC fabrication processing. Bulk-micromachining techniques are applied in the fabrication of the scanning-type microspectrometer. After the standard microelectronic process (bipolar or CMOS), a post-process module adds the Fabry-Perot etalons in the array-type microspectrometer. Metallic mirrors were used due the high reflectance exhibited in the entire visible part of the spectrum (especially silver mirrors). Moreover, the deposition of a single layer is preferable to the multilayer dielectric mirrors. Optically smooth and flat mirror surfaces were obtained. The fabrication of the Fabry-Perot microinterferometer in silicon enables the integration of electronic circuitry with the photodetectors.

The realized microspectrometers in silicon can be applied in the visible part of the spectrum with a reasonable resolution (25 for first-mode operation).
Samenvatting

Apparatuur voor optische spectraalanalyse vindt reeds wijdverspreid toepassing in wetenschap en industrie. Optische microsystemen, zoals microspectrometers geïntegreerd met microelectronica, maken nieuwe toepassingen mogelijk.

Dit proefschrift beschrijft het ontwerp, de fabricage en de karakterisatie van twee verschillende typen microspectrometers (scanning type en array type) in silicium voor gebruik in het zichtbare deel van het spectrum (390-760 nm). Er wordt een volledig geïntegreerd optisch microsysteem gepresenteerd (chip-oppervlak 4.2 mm bij 3.9 mm), gebaseerd op zestien kanalen, geconfigureerd in een matrix, samen met electronische circuits op de chip (uitleseselectronica voor de fotodetectoren, een licht-frequentie omzetter en een bus interface). Het golflengte-selecterend element dat wordt gebruikt, is een Fabry-Perot resonantieruimte, welke dienst doet als interferometer. De Fabry-Perot interferometer bestaat uit twee gedeeltelijk reflecterende parallelle spiegels met lage verliezen, gescheiden door een ruimte. De optische transmissiekarakteristiek van een dergelijk element wordt gekarakteriseerd door een serie scherpe resonante transmissiepieken telkens als de afstand tussen de spiegels gelijk is aan een veelvoud van de halve golflengte van het invallende licht. Het gefilterde spectrum wordt geprojecteerd op de geïntegreerde fotodiode, alwaar de aanwezigheid en intensiteit van de spectrale componenten wordt gedetecteerd.

Beide microspectrometers zijn vervaardigd met gebruikmaking van conventionele IC fabricagetechnieken. Voor de microspectrometer van het scanning type worden bulk-micromachining technieken toegepast. Na de standaard microelectronica-fabricage (bipolair of CMOS) worden in het geval van de array-type spectrometer de Fabry-Perot etalons aangebracht, middels post-processing. De gebruikt spiegels zijn van metaal, vanwege de hoge mate van reflectie over het gehele zichtbare gebied van het spectrum (vooral zilveren spiegels). Bovendien heeft depositie van een enkele laag de voorkeur boven meerlagige dielelectrische spiegels, doordat optisch gladde en vlakke spiegels worden verkregen. De fabricage van de Fabry-Perot microinterferometer is silicium maakt on-chip integratie van zowel electronische circuits als fotodetectoren mogelijk.
De gerealiseerde silicium microspectrometers kunnen worden toegepast in het zichtbare gedeelte van het spectrum, met een redelijke resolutie (25 voor een enkele piek in het zichtbare gedeelte van het spectrum).
Resumo

Equipamento para análise espectral de luz é utilizado tanto na indústria como em laboratórios de análise e investigação. Micro-sistemas ópticos como os micro-espectrômetros integrados com microeletrônica prometem novas aplicações.

Esta tese descreve a concepção, fabrico e caracterização de dois diferentes tipos de micro-espectrômetros - sintonizável e de matriz - em silício para a luz visível do espectro electromagnético (entre aproximadamente 390 nm até 760 nm de comprimento de onda). No final, um micro-sistema óptico (dimensões 4.2 mm por 3.9 mm), baseado num espectrômetro de matriz de dezasseis canais, com electrônica incluída (elettronica de leitura dos fotodiódos, conversão luz-frequência e um Bus para externo interface) é apresentado.

A cavidade de ressonância Fabry-Perot é usada como filtro óptico para selecionar diferentes comprimentos de ondas, actuando como um interferómetro. Um interferómetro Fabry-Perot consiste em dois espelhos paralelos e semi-reflectivos separados de uma certa distância. A transmissão óptica através da cavidade resulta numa série de picos de largura estreita quando a distância entre os espelhos é um múltiplo de metade do comprimento de onda da luz incidente. A intensidade da luz filtrada é medida num fotodiodo, também integrado juntamente com a cavidade.

Os dois espectrômetros foram fabricados usando os processos convencionais para fabrico de circuitos integrados e técnicas de micromausinação para o espectrômetro sintonizável. O espectrômetro de matriz depois do standard fabrico da microeletrônica (bipolar ou CMOS) precisa de um módulo para fabrico da cavidade Fabry-Perot (neste caso a cavidade é um filme fino). Espelhos metálicos foram usados pois exibem alta reflectividade na parte visível do espectro electromagnético. Também a deposição de uma única camada metálica para espelho é preferível á deposição de uma multi-camada de materiais dielétricos. Espelhos de prata com superfícies anti-rugasos e planares foram realizados.

O fabrico do micro-interferómetro Fabry-Perot em silício permite a integração conjunta com a microeletrônica bem como dos fotodiódos.
Resumo

Os espectrômetros realizados em silício podem ser aplicados na análise espectral da luz visível com razoável resolução (25 para o primeiro modo de operação).
List of Publications

JOURNAL PUBLICATIONS


PRESENTATIONS


List of Publications
Acknowledgements

This work could not have been completed without the help and support of many people. To them I would like to express my deepest gratitude. In particular I would like to thank the following people:

Prof. J. H. Huijsing, my promotor and Prof. S. Middelhoek for giving me the opportunity to be a Ph.D. student at the Laboratory for Electronic Instrumentation. The laboratory has a research environment so stimulating and of such high quality that I greatly enjoyed the last four years I spent in Delft.

Dr. Reinoud Wolffensbuttel, my project supervisor, for his continuous support and guidance, and for enabling me to gain experience abroad.

Dr. Marian Bartek, for his technology support, for processing the devices, for his helpful suggestions and critical remarks on this thesis and for his pleasant company around California and Nevada.

The members of the Delft Institute for Micro-Electronics and Submicron technology (DIMES) for technical assistance in fabrication of the devices. In particular J. Groeneweg, who has deposited all the metallic mirrors used in the devices.

Ger de Graaf for his technical support and helpful cooperation during the design of the device in CMOS with the light-frequency converter.

Piet Trimp, Jeff Mollinger, Rob Stolk, Maureen Meekel and Jeroen Bastemeijer for their technical support and Evelyn Sharabi, Sabine van den Boer, Ingeborg Egmond and Willem van der Sluys for their administrative support.

My fellow Ph.D. students at the Laboratory, for providing a research environment that was so stimulating and friendly.

Dept. of Industrial Electronics at the University of Minho, Portugal and my colleagues for making my stay in Delft possible. In particular, Prof. C. Couto for his continuous support.
Acknowledgements

Room-mates J. Foerster, E. Cretu, S. H. Kong and Davey Wijngaards, for sharing their experiences with me. As Dutch is not my mother language, I wish to thank Davey for his translation of the summary and the “stellingen”.

R. Kazinczi, for producing and analysing the AFM images.

Drs. Mirjam Nieman and Wendy Sturrock for carefully correcting the English of this thesis.

Prof. K. Najafi and the staff members of the University of Michigan, Ann Arbor, USA for enabling me to gain experience in MOS processing, and E. Cretu for his pleasant company around Mackinac island and Michigan upper peninsula.

The Portuguese foundation, Fundação Ciência Tecnologia/Program Praxis XXI for the financial support during my stay in Delft.

The Dutch Technology Foundation (STW) that financially supported this project.

My family and especially my mother have provided the essential encouragement and support all these years.

Finally, I would like to thank my wife Fátinha for her love, support, and patience during these past four years.
About the Author

José Higino G. Correia was born in Carmona, Angola on November 19, 1965. In 1990 he completed his Physical Engineering degree at the University of Coimbra, Portugal. In February, 1991, he started working as an Assistant Lecturer at the Dept. of Industrial Electronics, University of Minho, Portugal. In July, 1994, he became a Lecturer on the basis of the research work "Intelligent Signal Processing for Load Cells". In September, 1995, he joined the Laboratory for Electronic Instrumentation, Dept. of Electrical Engineering, Delft University of Technology, as Ph.D. student where he performed research on the application of silicon micromachining techniques for optical microsystems. During his Ph.D. research, he spent three months at the University of Michigan, Ann Arbor, USA as a visiting scholar, to work in their IC laboratory.