Stellingen

behorende bij het proefschrift

INFRARED MICRO-SPECTROMETER
BASED ON A MULTI-SLIT GRATING

Seong Ho KONG

Delft, 16 oktober 2002
1. Gebruik van infrarood voor spectrale analyse voegt veel toe aan zichtbare analyse, aangezien het infrarode deel van het elektromagnetisch spectrum het gedeelte is waar zich de emissie- en absorptielijnen bevinden van vrijwel alle moleculen alsook van vele atomen en ionen (B. Stuart, Modern infrared spectroscopy, John Wiley & Sons Ltd. 1996).

Using infrared for spectral analysis adds much to visible analysis, since the infrared part of the electromagnetic spectrum is where the emission and absorption lines of virtually all molecules as well as those of numerous atoms and ions are (B. Stuart, Modern infrared spectroscopy, John Wiley & Sons Ltd. 1996).

2. Hoewel de micro-spectrometer in verhouding tot grote macroscopische instrumenten een mindere spectrale resolutie heeft, wordt deze beperking in veel toepassingen raakzaak gecompenseerd door de kleine omvang en lage kosten.

Compared to bulky macroscopic devices, the micro-spectrometer has an inferior spectral resolution but its small size and low cost more than compensate for this limitation in many applications.

3. Naast de IC compatibiliteit en de superieure mechanische eigenschappen maakt de hoge mate van transparantie van silicon voor golflengtes voorbij 1 µm deze zeer bruikbaar voor een IR spectrometer.

In addition to its IC compatibility and superior mechanical properties, the highly transparent property of silicon in the wavelength range exceeding 1 µm makes it very useful for an IR spectrometer.

4. In termen van fabricage compatibiliteit, eenvoud en ondoorzichtigheid bij infrarode golflengtes is aluminium het meest geschikte materiaal voor de realisatie van een tralies met goede prestaties in de IR micro-spectrometer.

In terms of fabrication compatibility, simplicity and opaqueness at infrared wavelengths, aluminum is the most suitable material to realise a multi-slit grating with a good performance in the IR micro-spectrometer.


Even objects that we think of as being very cold, such as ice cube, emit infrared. Similarly, even cold-hearted thugs have their standard of good and evil.


The Dutch weather provides the perfect motivation for the development of a sensitive wind or humidity sensor.


Dutch people have a micro-scaled recognition towards Korea. Some of my neighbours even say that you are a Chinese if you are from Korea.

8. De westere wetenschap probeert alle details te analyseren terwijl de oosterse filosofie het geheel wil beschouwen zonder ontleding. Dat is waarom oosterse geleerden een grotere kans hebben een identiteitsconflict te krijgen alvorens tot een kritisch besluit te komen.

Western science tries to analyse all the details while oriental philosophy tries to regard the whole without decomposition. That is why oriental scientists are more likely to have an identity conflict before reaching a critical decision.

9. De natuur is een open boek, helas geschreven in haar eigen taal, zodat de meeste mensen haar niet kunnen lezen.

Nature is an open book, unfortunately written in its own language, so that most people can not read it.


The home town of Guus Hiddink, Varsseveld, became almost like the Holy Land for enthusiastic Korean soccer fans. In that sense, missionary work could be done much better through soccer.
Infrared Micro-Spectrometer

Based on a Multi-Slit Grating

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op woensdag 16 oktober 2002 om 10:30 uur

door

Seong Ho KONG

Master of Science, Tohoku University, Japan
geboren te Young-cheon, Kyungpook provincie, Zuid-Korea
Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. J.H. Huijsing

Toegevoegd promotor:
Dr. ir. R.F. Wolffenbuttel

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof. dr. ir. J.H. Huijsing, Technische Universiteit Delft, promotor
Dr. ir. R.F. Wolffenbuttel, Technische Universiteit Delft
Prof. dr. ir. G.C.M. Meijer, Technische Universiteit Delft
Prof. dr. P.M. Sarro, Technische Universiteit Delft
Prof. dr. J.H. Correia, University of Minho (Portugal)
Prof. dr. ir. P.P.L. Regtien, Universiteit Twente
Prof. dr. R. Puers, Katholieke Universiteit Leuven

ISBN 90 5166 925 9

Published by:

Eburon Publishers
P.O. Box 2867
2601 CW Delft
The Netherlands
Phone: +31-(0)15 2131484
Fax: +31-(0)15 2146888
info@eburon.nl / www.eburon.nl

Copyright © 2002 by S.H. Kong

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the publisher or the author.

PRINTED IN THE NETHERLANDS
# Table of Contents

## Chapter 1. Introduction

1.1 Motivation and objectives ........................................... 1  
1.2 Technical considerations ........................................... 3  
  1.2.1 The advantages of integrated spectrometers ............. 3  
  1.2.2 Optical properties of silicon and aluminium .......... 4  
  1.2.3 IC compatibility ........................................... 4  
1.3 Organisation of the thesis ......................................... 4  

## Chapter 2. Infrared spectrometers

2.1 Overview of infrared light and its applications ................. 7  
2.2 Spectral analysis using infrared light ............................ 11  
  2.2.1 Introduction ............................................ 11  
  2.2.2 Spectral quantities ....................................... 13  
    2.2.2.1 Wavelength, frequency and wave number .......... 13  
    2.2.2.2 Resolving power ................................... 13  
    2.2.2.3 Luminosity ......................................... 14  
  2.2.3 Wavelength-selecting elements ................................ 15  
    2.2.3.1 Prisms ............................................. 15  
    2.2.3.2 Gratings ........................................... 17  
    2.2.3.3 Interferometers ................................... 29  
  2.2.4 Optical path .............................................. 31  
  2.2.5 Infrared detectors .......................................... 33  
    2.2.5.1 Photon types ....................................... 33  
    2.2.5.2 Thermal types ...................................... 36  
    2.2.5.3 Performances of different types of IR detector .... 42  
  2.2.6 Classification of Infrared spectrometers .................... 43  
    2.2.6.1 Prism-based spectrometer ............................ 43  
    2.2.6.2 Grating-based Spectrometer .......................... 45  
    2.2.6.3 Fourier-transform Infrared (FT-IR) Spectrometer .. 46  
2.3 Conventional macro-spectrometer vs. integrated micro-  
  spectrometer .................................................. 50
Chapter 3. Design of the multi-slit grating-based infrared micro-spectrometer

3.1 Introduction .................................................................................. 61
3.2 Basic design considerations .......................................................... 62
    3.2.1 Device structure and principle .............................................. 62
    3.2.2 Optical path ......................................................................... 63
    3.2.3 Alignment ............................................................................. 67
    3.2.4 Lensless spectrometer ......................................................... 68
3.3 The multi-slit grating ................................................................. 69
    3.3.1 Diffraction by a multi-slit grating ......................................... 69
    3.3.2 Optical property of grating material ..................................... 71
    3.3.3 Resolving power of multi-slit grating ................................... 73
3.4 The silicon optical path ............................................................... 75
3.5 The thermopile-based infrared detector array ............................. 79
    3.5.1 Construction ........................................................................ 79
    3.5.2 Performance - Optimisations of IR detector ....................... 80
      3.5.2.1 Thermal conductance .................................................... 80
      3.5.2.2 AT as a function of beam length .................................... 86
      3.5.2.3 Optimization of beam length ....................................... 89
      3.5.2.4 Integrated Peltier cooler ............................................... 94
    3.5.3 Choice of thermoelectric material ...................................... 103

Chapter 4. Fabrication of the multi-slit grating based infrared micro-spectrometer

4.1 Technologies used for fabrication .............................................. 111
    4.1.1 Introduction ........................................................................ 111
    4.1.2 Silicon IC technology ......................................................... 113
      4.1.2.1 Oxidation .................................................................... 113
      4.1.2.2 Photolithography .......................................................... 118
      4.1.2.3 Patterning .................................................................... 120
      4.1.2.4 Chemical Vapour Deposition (CVD) .............................. 124
      4.1.2.5 Ion implantation ............................................................ 125
      4.1.2.6 Metallisation ................................................................. 126
    4.1.3 Bulk micromachining technology ....................................... 127
### Table of Contents

4.1.4 Wafer level low-temperature silicon fusion bonding ........131
4.2 Fabrication of the IR micro-spectrometer ..........................132
    4.2.1 Multi-slit grating .............................................134
    4.2.2 IR detector array ............................................137
    4.2.3 Low-temperature wafer-wafer fusion bonding ............140

**Chapter 5. Measurements**

5.1 Introduction ..........................................................153
5.2 Electrical and thermoelectrical measurements ..................154
5.3 Optical quality of fabricated gratings ............................157
5.4 Thermoelectrical response of the thermopile-based IR detectors ..159
5.5 Spectral response of the IR micro-spectrometer .................160

**Chapter 6. Conclusions**

6.1 The integrated infrared micro-spectrometer ......................163
6.2 Applications ..........................................................164
6.3 Future work ............................................................165

**Summary**

169

**Samenvatting**

171

**List of publications**

173

**Acknowledgements**

177

**About the author**

181
1.1 Motivation and objectives

The great success of silicon integrated-circuit (IC) technology motivated researchers to develop a large variety of silicon-based sensors [1.1][1.2]. Recently, efforts have been focused on Micro Electro Mechanical System (MEMS)-based micro-fluidic devices [1.3][1.4], which open up new possibilities for the miniaturization of conventional chemical and biochemical analysis systems. Those micro-fluidic devices are combined with chemical or optical sensors controlled by integrated circuits, resulting in the micro total analysis systems (μTAS).

Spectrometers can not only be applied for quality inspection in industry and agriculture; they could also be a part of μTAS since they can analyse gaseous or liquid substances. Spectrometers read the chromatography results by measuring the spectral absorption of a chemical constituent placed between a light source and a dispersing element. Alternatively, the spectrometer can be applied to determine the composition of a compound
**Introduction**

![Image of spectroscope](image)

**Fig. 1-1.** Grating spectroscopy with a photocell read-out [1.5].

by using a white light source and observing distinct absorption lines or bands in a specified wavelength range. Such a spectroscopic sensor system requires a wide-band light source, an open path for interaction with the substance, and a dispersing element, which disperses the light on to a detector array. In an earlier period a spectrum was directly examined using a photocell [1.5], as can be seen in Figure 1-1.

Based on a grating spectrometer, a very high resolution ($2 \times 10^6$) had been achieved by Lindblom et al. [1.6]. As illustrated in Figure 1-2, the light

![Diagram of spectrometer](image)

**Fig. 1-2.** Complex and bulky optical arrangement of a very high-resolution grating spectrometer mounted in a pressure chamber PR [1.6].

---

*Infrared Micro-spectrometer based on a Multi-slit Grating*
enters through slit $S$, passes through window $W$, is collimated by mirror $C$, diffracted successively by gratings from $G1$ to $G4$ and finally focused by mirror $F$ onto detector $IMD$. However, as the device has dimensions in the order of a metre and a weight of more than a half ton, it is not suitable to be applied for $\mu$TAS. An integrated spectroscopic device, which has a compact size, is preferred for such an application. Whereas the external illumination of the substance is required, most of the analysis components are integrated on a single chip resulting in a micro-spectrometer. Recent progress in micromachining technology makes implementation of the dispersion and detection elements in a silicon micro-spectrometer possible, so that it can perform on-chip analysis of the spectrum.

Using infrared for spectral analysis is very useful, since the infrared part of the electromagnetic spectrum is where the emission and absorption lines of almost any type of molecules are located [1.7]. Moreover, the optical properties in the infrared spectral range of standard materials used in the device, e.g., silicon and aluminium, are promising, as is shown in the next section.

The micro-spectrometer may operate either by using an array of detectors, each with a uniform spectral response, or by scanning the dispersed light using a single calibrated detector. This thesis will introduce an infrared micro-spectrometer based on a multi-slit grating and a thermopile-based infrared detector array.

### 1.2 Technical Considerations

#### 1.2.1 The advantages of integrated spectrometers

Compared to bulky macroscopic devices, the micro-spectrometer has an inferior spectral resolution but its small size and low cost more than compensate for this limitation in many applications. Moreover, the limited performance of the miniaturized spectrometer has a potential for further improvement, because it allows integration of microelectronic circuits directly into the system. The integrated compact micro-spectrometer would also reduce the assembly problem, while allowing an easier alignment.
1.2.2 Optical properties of silicon and aluminium

The versatility of silicon, for instance its IC compatibility combined with the superior mechanical properties, has prompted applications not only for a variety of sensors but also for actuators. A very useful property of silicon in an IR spectrometer is the fact that it is a highly transparent material for the wavelength range exceeding 1 μm [1.8], as will be discussed in more detail in Chapter 3.

The high reflectivity of aluminium, which is the conventional metallisation material in microelectronics technology, at both the visible and infrared wavelengths enables the realisation of optical components in a compatible way with standard IC processing. In terms of fabrication compatibility, simplicity and opaque optical property, aluminium is the most suitable material to realise a multi-slit grating with a good performance in the infrared micro-spectrometer.

1.2.3 IC compatibility

When silicon and aluminium are used for the optical path and grating material, respectively, the whole fabrication process except KOH etching remains simple and fully compatible with standard IC processing. Therefore, if the KOH etching is the final step prior to dicing, wire bonding and packaging, the wafer does not need to return to the cleanroom [1.9].

1.3 Organisation of the thesis

This chapter introduces the motivation and objectives of this thesis. In Chapter 2 the overview of infrared light, its applications and a brief history of research on infrared radiation are described. The basic diffraction theory associated with multi-slit grating and the operating principle of the optical components of the infrared spectrometer are discussed. Moreover, the actual potential and limitations of integrated micro-spectrometer are presented by introducing recently developed miniaturized infrared spectrometers. Chapter 3 presents the design considerations of miniaturized infrared micro-spectrometer based on a multi-slit grating. In the design presented, a fully-
integrated, miniaturized spectrometer has been pursued, where much effort is put in the compatibility of its fabrication process with standard microelectronic technology. Moreover, it discusses the approach used to optimise the device performance. Chapter 4 is concerned with the fabrication techniques used to realise the designed micro-spectrometer. The measurement results are reported in Chapter 5. Finally, the conclusions of this thesis are drawn in Chapter 6 with a brief description of some envisioned applications and guide lines for future work.
References


[1.7] B. Stuart, Modern infrared spectroscopy, John Wiley & Sons Ltd. 1996.


2.1 Overview of infrared light and its applications

The entire range of possible wavelength values of radiant energy is referred to as the electromagnetic spectrum. It is usually subdivided into a few distinct regions depending on the radiant energy as can be seen in Figure 2-1. Infrared light lies between the visible and microwave regions of the electromagnetic spectrum. The Latin prefix “infra” means “below” or “beneath”. Thus “infrared” refers to the region where photons carry an energy below that of the red end of the visible colour spectrum.

The existence of radiant energy beyond the long-wavelength end of the visible spectrum was discovered in 1800 by the English astronomer Sir William Herschel. For a study of the heating effects of sunlight, he passed sunlight through a glass prism to form a spectrum. As he moved a simple mercury-in-glass thermometer through the spectrum, he was able to measure the heating effect of the various colours, starting from the blue region and progressing toward the red region. Herschel found that when he
moved the thermometer beyond the red end of the spectrum, where there was no light visible, there still was an energy source causing heating. This experiment led to the discovery of a new region of the electromagnetic spectrum: the infrared one.

**Fig. 2-1. The electromagnetic spectrum.**
The heat that we feel from sunlight, a fire, a radiator or an electric lamp is infrared. The higher the temperature of a radiation source, the higher the spectral radiant energy at all wavelengths. Moreover, with increase in temperature, the wavelength of which the spectrum peaks shifts to shorter wavelengths. This is known as Wien's law [2.1]. Since the primary source of infrared radiation is heat or thermal radiation, any object which has a temperature radiates in the infrared. Even objects that we think of as being very cold, such as an ice cube, emit infrared. When an object is not quite hot enough to radiate visible light, it will emit most of its energy in infrared. For example, hot charcoal may not give off light but it does emit infrared radiation, which we feel as heat.

Near-infrared (NIR) light is closest in wavelength to visible light and far infrared (FIR) is closer to the microwave region of the electromagnetic spectrum. However, the nomenclature of the infrared regions is confusing and there is no general agreement on the names the various regions. The short-wavelength limit of the NIR is usually taken as 750 nm. The end of the NIR is taken by most people to be 3 μm - but sometimes as 2 or 2.5 μm. The region of mid infrared (MIR) starts at 3 μm and ends anywhere between 20 and 50 μm. This region is also called the 'fingerprint' region because it contains the characteristic vibrational spectra of many molecules. The fingerprint region is most popularly defined as 2.5 ≤ λ ≤ 25 μm or 4000 ≤ k ≤ 400 cm⁻¹, where λ and k represent the wavelength and wave number, respectively [2.2]. Beyond the MIR, there is a region of FIR. Where the FIR ends is a matter of taste but usually it would not extend beyond 1 mm. Sometimes, FIR stops at 100 μm and the region between 100 μm and 1 mm is known as the sub-millimetre region.

Infrared is applied in many different fields such as remote control, industrial inspection, process control, night vision, damage detection, fire control, environmental monitoring, astronomy, spectral analysis systems and so on. Night-vision capability based on infrared sensor technology can be classified into two main categories: uncooled thermal detectors and cryogenically operated photon detectors [2.3]. The advent of uncooled infrared sensors has created the opportunity to achieve a night vision sensor with a wide operating spectral range at low cost. However, its sensitivity is generally lower than that of a photon detector.

Apart from the night-vision capabilities, infrared imaging is often used in
Infrared spectrometers

industry for non-destructive testing. It is an effective method that provides a quick thermal evaluation. It can help to solve many heat-related problems in design, production, or use.

After Sir William Herschel's discovery of infrared showed that the sun emits infrared radiation, astronomers tried to see if other objects in the universe radiate infrared light. In 1856, thermocouples were used to detect infrared radiation from the moon [2.4]. In the early 1900s, infrared radiation was successfully detected from the planets Jupiter and Saturn and from some bright stars such as Vega and Arcturus [2.4]. However, the insensitivity of the early infrared instruments limited the detection of other near-infrared sources.

A major breakthrough in the development of sensitive infrared detectors was made by Frank Low in 1961 [2.4]. He built the germanium-based bolometer, which is hundreds of times more sensitive than other previously developed detectors. A thin strip of germanium is used as a thermally sensitive resistor. It is placed in a container, which has a small opening and is cooled down to 4 degrees Kelvin using liquid helium for the best performance. When infrared radiation comes through the opening and hits the germanium strip, its conductivity is changed. The change in conductivity can be measured and is directly related to the amount of infrared radiation entering the container.

Another giant leap in the infrared technology is the development of infrared array detectors in the 1980s motivated by the IR image sensor proposed by Shepherd et al. [2.5]. Basically, a detector array is a set of single detectors that results in enhancement of observational capability without the need to move optical components. The major advantages of such a staring array system are the elimination of the mechanical scanner, increased sensitivity due to the longer integration times available and the absence of scan inefficiencies. More advanced recent technology allows the system to produce images containing a large number of pixels [2.6][2.7][2.8].

Recently, analytical applications of infrared have rapidly been gaining importance in the area of molecule detection and biotechnology. Due to these breakthroughs in infrared technology, infrared applications have become more and more important in many fields and continue to bring new visions.
2.2 Spectral analysis using infrared light

2.2.1 Introduction

Spectroscopy involves the detailed study of the electromagnetic radiation from an object, mostly in the area of chemistry and physics. Spectrometers are instruments that separate light into its different wavelength components (=disperse), creating a spectrum. Within this spectrum, we can study emission and absorption of light by atoms and molecules. An emission line occurs when an electron drops down to a lower orbit around the nucleus of an atom and loses energy. An absorption line occurs when electrons move to a higher orbit by absorbing a photon or energy in another form. Spectroscopic measurements of emission and absorption of light can tell us a great deal about the energies associated with a particular atom or a particular chemical bond in a molecule, since these wavelengths are directly related to the energy changes that take place within the molecule or in its interactions with other molecules.

Figure 2-2 shows a plot of intensity versus wavelength exhibiting how much light is present or absent at each wavelength. A peak in the plot shows the position of an emission line and valley shows an absorption line. The spacing and location of these lines are unique to each atom and molecule.

The energy levels and the internal structure of the atom can be determined from the positions of the spectral lines in a spectrum, and the probabilities

![Fig. 2-2. Emission and absorption lines in a spectrum.](image-url)
of the transitions between individual levels can be derived from the line intensity. In addition, the spectral line intensity is proportional to the number of emitting atoms. The line intensity is therefore used to determine the composition rate of the individual elements in a substance being investigated, by comparing specimens containing known amounts of those elements.

There are three types of spectra which an object can emit:
(1) continuous,
(2) emission, and
(3) absorption spectra.

Continuous spectra (also called a black-body spectra) are emitted by any object at a certain temperature which radiates heat. The spectrum comprises a continuous band with every wavelength having some amount of radiation. As mentioned, the atoms and molecules in a gas will absorb only certain wavelengths of light. The pattern of these lines is unique to each element and tells us, for example, what elements make up the atmosphere of the sun. We usually see absorption spectra from regions in space between us and a hotter source where a cooler gas lies.

An emission spectrum occurs when the atoms and molecules in a hot gas emit light at certain wavelengths, causing bright lines to appear in a spectrum. As with absorption spectra, the pattern of these lines are unique for each element.

Infrared spectroscopy can also be used to measure the wavelength and intensity of light due to absorption of infrared light. An infrared spectrum is obtained by passing radiation of a certain wavelength range through a sample and determining what fraction incident radiation is absorbed at a particular wavelength. The energy at which any peak in an absorption spectrum appears corresponds to the frequency of a vibration of a part of the sample molecule. A spectrometer can determine the composition of samples based on this absorption.

Infrared spectral analysers are amongst the most powerful analytical techniques in many fields since infrared analysis can be used for almost any type of sample. The infrared part of the electromagnetic spectrum is where the emission and absorption lines of virtually all molecules as well as those
of numerous atoms and ions are [2.9].

2.2.2 Spectral quantities

2.2.2.1 Wavelength, frequency and wave number

It is common to describe a monochromatic radiation by its wavelength rather than by frequency. In the infrared region of the electromagnetic spectrum, the usual measurement unit of wavelength is micrometer (μm) for convenience (1 μm = 10⁻³ nm) instead of nanometre (nm).

Wavelength \( \lambda_v \) in a vacuum is related to frequency \( v \) by the relation \( \lambda_v = c / v \), where \( c \) is the speed of light in a vacuum. In a medium of refractive index \( n \), the wavelength becomes \( \lambda = c / n v = \lambda_v / n \). Although it is convenient in experimental work to characterize a monochromatic radiation by its wavelength, in theoretical work frequency is nevertheless used as the fundamental parameter. Indeed, in the quantum theory of the emission of luminous radiation by atoms frequency \( v \) is used, which is given by \( E_2 - E_1 = h v \), in which \( E_2 \) and \( E_1 \) represent the energy of the emitting atom in the initial and the final states, and \( h \) is Planck's constant. It is known that measurements in which frequency is used instead of wavelength are less accurate and more difficult. Frequency and wavelength have the relationship of \( v = c / \lambda_v \). Therefore in spectroscopy, wave number is commonly used instead of frequency since it is proportional to frequency \( (k = 1 / \lambda_v = v / c) \) and is known to have the same accuracy as wavelength. The unit commonly used as a measure of wave number is the inverse centimetre (cm⁻¹).

2.2.2.2 Resolving power

When an instrument can resolve two adjacent narrow spectral components of equal intensity, \( \lambda - (\Delta \lambda / 2) \) and \( \lambda + (\Delta \lambda / 2) \), centered on wavelength \( \lambda \), the resolving power is defined as:
\[ R = \frac{\lambda}{\Delta \lambda} \]  

(2.1)

It expresses the ratio of the wavelength to the narrowest spectral width that can be distinguished. The smaller the value of \( \Delta \lambda \), the more highly the spectrum is resolved, so \( \Delta \lambda \) is the resolution of the instrument.

When there are two monochromatic spectral components, \( \lambda \) and \( \lambda + \Delta \lambda \), each of these produces a distribution of intensity. The resulting intensity is the sum of those spectral components (see Figure 2-3). It is generally accepted that two spectral components are resolved if the central maximum of one spectral component is superimposed with the first minimum of the other.

![Figure 2-3](image)

**Fig. 2-3.** Distribution of intensity resulting from superposition of the diffraction patterns that have neighbouring wavelengths (Rayleigh criterion).

### 2.2.2.3 Luminosity

The performance of a dispersing instrument is not simply dependent on the resolving power. It is not sufficient to have only a spectral wave band that is as narrow as possible. Each of those elemental wave bands should have sufficient radiant energy to achieve a reasonable measurement. This factor is defined as luminosity, which describes the amount of light that passes from the source through the instrument and on to the detector. The
luminosity can be therefore defined as the ratio of the intensity at the centre of a spectral line to the radiance at the entrance slit.

2.2.3 Wavelength-selecting elements

A wavelength-selecting or dispersing element is indispensable to perform spectral analysis. The dispersing element spreads out the spectral continuum of radiation energy from a source, so that only narrow regions of the spectrum pass through optical path to the detector. The dispersing element determines both the resolution and the spectral range of the instrument with respect to a known optical-path length.

Prisms, gratings and interferometers are commonly used for the purpose of dispersing. In earlier years, prisms were the usual dispersing elements of infrared instruments. Now gratings or interferometer instruments have become more common.

2.2.3.1 Prisms

A prism refracts a light beam at an angle with different amounts for different wavelengths, based on the fact that the refractive index changes with wavelength. A prism shows a non-linear relationship between wavelength and refraction angle and its theoretical or intrinsic resolving power is not high. It is therefore not often used in state-of-the-art equipment. When the full aperture is illuminated, the theoretical resolving power of prisms is [2.10]

\[ R_0 = e \frac{dn}{d\lambda}, \]  

(2.2)

in which \( e \) is the width of the prism base (see Figure 2-4 (a)). If the full aperture is not illuminated (see Figure 2-4 (b)), the Eq (2.2) becomes

\[ R_0 = (e_1 - e_2) \frac{dn}{d\lambda}. \]  

(2.3)

If the beam traverses a few prisms in succession, the base widths \( e \) will be added. In fact, two or three prisms are often connected together in series in
Infrared spectrometers

Fig. 2-4. Influence of method of illumination on intrinsic resolving power of a prism.

order to obtain sufficiently high values of the base width $e$ without making the system too big.

Material for prisms need to be highly dispersive for a high resolving power, as can be imagined from Eq (2.2), and need to be very transparent in the relevant spectral region. However, these two requisites are not independent of each other. As the intrinsic resolving power of a prism is proportional to the width $e$ of the base, the apex angle of the prism should have the highest possible value. However, the angles of incidence and emergence increase with the apex angle, increasing the loss of light by reflection at the surfaces of the prism. This loss is dominant especially when the prism is made of a high-refractive index material. In practice, spectrograph prisms always have apex angles between about 50 and 70 degrees to reduce the loss of light by reflection [2.10].

All the available materials for prisms working in infrared spectral regions have limited bands of transmission, and the larger prisms needed for better throughput are very costly. In spite of their distinctly poorer performance in terms of resolving power and luminosity, prisms have by no means been ousted by gratings in spectrometers. Prisms have the advantages of having no overlapping orders and good efficiency. The spectral regions that prisms can cover are part of the ultraviolet, the visible and the infrared as far as about 40 $\mu$m. Quartz and fused silica cover the visible and ultraviolet region, down to 200 nm. These materials extend their transparency into the near-infrared spectral range, to about 3.5 $\mu$m. Calcium aluminate and arsenic trisulphide prisms can be transparent to 5 and 12 $\mu$m, respectively. Table 2-1 shows the characteristics of compound materials for prisms that are transparent in the infrared spectral ranges.
TABLE 2-1. Characteristics of infrared optical materials [2.10][2.11].

<table>
<thead>
<tr>
<th>Compound material</th>
<th>Long-wavelength limit [μm]</th>
<th>Refractive index</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO₂</td>
<td>4.5</td>
<td>1.45 at 1.0 μm</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>6.5</td>
<td>1.76 at 1.0 μm</td>
</tr>
<tr>
<td>MgF₂</td>
<td>7.5-9</td>
<td>1.38 at 3.3 μm</td>
</tr>
<tr>
<td>LiF</td>
<td>8</td>
<td>1.38 at 2.0 μm</td>
</tr>
<tr>
<td>MgO</td>
<td>9.5</td>
<td>1.66 at 3.3 μm</td>
</tr>
<tr>
<td>CaF₂</td>
<td>11</td>
<td>1.42 at 2.0 μm</td>
</tr>
<tr>
<td>As₂S₃</td>
<td>13</td>
<td>2.40 at 3.3 μm</td>
</tr>
<tr>
<td>ZnSe</td>
<td>21.8</td>
<td>2.40 at 3.3 μm</td>
</tr>
<tr>
<td>AgCl</td>
<td>25</td>
<td>1.45 at 1.0 μm</td>
</tr>
<tr>
<td>NaCl</td>
<td>15-22</td>
<td>1.98 at 10 μm</td>
</tr>
<tr>
<td>KBr</td>
<td>25</td>
<td>1.45 at 1.0 μm</td>
</tr>
<tr>
<td>CsBr</td>
<td>40</td>
<td>1.45 at 1.0 μm</td>
</tr>
<tr>
<td>KBr</td>
<td>33</td>
<td>1.53 at 10 μm</td>
</tr>
<tr>
<td>KCl</td>
<td>26</td>
<td>1.46 at 1.0 μm</td>
</tr>
<tr>
<td>ZnS</td>
<td>14.5</td>
<td>2.20 at 3.3 μm</td>
</tr>
<tr>
<td>CdTe</td>
<td>28</td>
<td>2.69 at 3.3 μm</td>
</tr>
<tr>
<td>BaF₂/CaF₂</td>
<td>12</td>
<td>1.42 at 3.3 μm</td>
</tr>
<tr>
<td>CsI</td>
<td>55</td>
<td>1.74 at 10 μm</td>
</tr>
</tbody>
</table>

2.2.3.2 Gratings

Diffraction gratings are defined as devices which impose a periodic deformation on the wavefront of any incident optical radiation and are divided according to several criteria such as their geometry, material, efficiency, manufacturing method, usage and so on. In this thesis, mainly multi-slit diffraction gratings are discussed even though almost no multi-slit gratings are used in macro-scale research instruments.

A multi-slit grating practically consists of a series of identical parallel and equally-spaced slits formed on a plane or spherical surface. In fact, each slit
of the multi-slit grating acts as an optical point source. The usefulness of gratings is derived from the fact that there exists a unique point where the light diffracted is in phase. Only when the path-length difference between the light diffracted by successive slits equals the wavelength of incident light, there will be constructive interference.

**Diffraction pattern of a single slit**

When the slit dimensions approach the wavelength of the radiation, the wave theory of light should be taken into account instead of geometric optics. It is assumed that the slit width is constant, plane waves arrive at the slit and the slit is illuminated uniformly. In the case that the distance between the slit and the screen is large enough, the projected pattern spreads out considerably, bearing little or no resemblance to the actual shape of the slit. This is Fraunhofer or far-field diffraction, which will be discussed in Chapter 3.

According to the Huygens principle, every unobstructed point of a wavefront arriving at the slit is a spherical secondary wavelet with the same frequency as that of the primary wave, as shown in Figure 2-5 [2.13]. The superposition of these Huygens wavelets in the forward direction results in

---

**Fig. 2-5. Optical diagram for the diffraction of light at a single rectangular slit, showing Huygens wavelets.**
a diffraction pattern. To calculate the intensity distribution of the diffraction pattern on the screen, one must take into account the contribution of every secondary wave superimposed at each position on the screen. Depending on whether the light waves are in phase or out of phase, based on the difference in the path length travelled by the different secondary waves arriving at the same point on the screen, constructive or destructive interference occurs.

The intensity distribution of a single-slit diffraction pattern is derived using the geometry and coordinates for the superposition of the waves, shown in Figure 2-6. The incident plane wave on the slit is divided into $M$ sections. The interval of each of the sections is $\Delta y$. The $d^{th}$ interval is placed at $+y_\alpha$ with respect to the central interval at $y=0$. The optical-path difference between the $d^{th}$ wave and the reference wave ($\alpha=1$) is $y_\alpha \sin \theta$ as can be seen in Figure 2-6. Thus, when the Fraunhofer condition ($X \gg D$) is implied, the travelled distance (optical-path length) of the $d^{th}$ source to the screen among secondary waves, $\rho_\alpha$ is

$$\rho_\alpha = \rho_1 - y_\alpha \sin \theta,$$

(2.4)

**Fig. 2-6.** A single slit for calculating the diffraction pattern in the far field ($X \gg D$).
Infrared spectrometers

and for the phase factor

$$\cos 2\pi \left( \frac{p_a}{\lambda} \cdot \frac{t}{T} \right)$$  \hspace{1cm} (2.5)

or

$$e^{i2\pi \left( \frac{p_a}{\lambda} \cdot \frac{t}{T} \right)} = e^{i(kp_a - \omega t)}$$  \hspace{1cm} (2.6)

$$= e^{i(kp_1 - kp_1 \sin \theta - \omega t)} = e^{i(kp_1 - \omega t)} e^{-ikp_1 \sin \theta}$$

where $t$ is the time and $\lambda$ and $T$ are the wavelength and period of the oscillation, respectively. The wave number $k = 2\pi / \lambda$ and the angular frequency $\omega = 2\pi / T$ are used.

Then the newly generated wave is

$$u_\alpha = u_0 e^{i(kp_1 - \omega t)} e^{-ikp_1 \sin \theta}, \hspace{1cm} (2.7)$$

where $u_0$ is the amplitude of the secondary source uniformly distributed across the slit. If the intensity reaching the slit is $C^2$, the $\alpha^{th}$ amplitude of the secondary source, which has a length $\Delta y$, is $C\Delta y/D$, that is,

$$u_\alpha = \left( \frac{C\Delta y}{D} \right) e^{i(kp_1 - \omega t)} e^{-ikp_1 \sin \theta}, \hspace{1cm} (2.8)$$

The superpositioned amplitude on the screen can be given by summing all contributions from $M$ sources, thus

$$u = \frac{C}{D} e^{i(kp_1 - \omega t)} \sum_{\alpha=1}^{M} \Delta y e^{-ikp_1 \sin \theta} \hspace{1cm} (2.9)$$

When $M$ becomes very large, so that the slit is divided into an infinite number, the Eq (2.9) can be expressed by
Spectral analysis using infrared light

\[ u = \frac{C}{D} e^{i(kp_1 - \omega t)} \int_{-D/2}^{D/2} e^{-i\beta y} dy. \] (2.10)

The integral in Eq (2.10) can be solved directly if \( \beta = k\sin \theta \) is used as follows:

\[ \int_{-D/2}^{D/2} e^{-i\beta y} dy = 2i e^{i\beta (D/2)} \frac{e^{-i\beta (D/2)}}{2i(i\beta)} = \frac{2i}{i\beta} \sin \beta \frac{D}{2} = D \frac{\sin \left( \frac{kD}{2} \sin \theta \right)}{k \frac{D}{2} \sin \theta}. \] (2.11)

As a result, the amplitude (Eq (2.10)) can be expressed as

\[ u = Ce^{i(kp_1 - \omega t)} \frac{\sin \left( \frac{kD}{2} \sin \theta \right)}{k \frac{D}{2} \sin \theta}, \] (2.12)

and the intensity at a position \( Y \) on the screen is

\[ u u^* = |u|^2 = C^2 \left( \frac{\sin \left( \frac{kD}{2} \sin \theta \right)}{k \frac{D}{2} \sin \theta} \right)^2. \] (2.13)

By setting \( C^2 = 1 \), the central maximum can be normalized to unit intensity, that is, \( |u|^2 = 1 \) at \( Y = 0 \) (\( \theta = 0 \)). Thus, the ratio of the intensity at a position \( Y \) on the screen, \( I \), to the intensity on the screen at \( Y = 0 \) (\( \theta = 0 \)), \( I_0 \), is

\[ \frac{I}{I_0} = \left( \frac{\sin \left( \frac{kD}{2} \sin \theta \right)}{k \frac{D}{2} \sin \theta} \right)^2 = \left( \frac{\sin \left( \frac{\pi D}{\lambda} \sin \theta \right)}{\frac{\pi D}{\lambda} \sin \theta} \right)^2. \] (2.14)

Figure 2-7 shows the diffraction pattern of a slit with a width \( D \), where the plot \( I/I_0 \) is a function of \( \sin \theta \). The minimum intensity occurs at \( \sin \theta = \pm m \frac{\lambda}{D} \), \( m = 1, 2, 3 \cdots \).
**Diffraction pattern of a double slit**

In the case of the diffraction pattern made by a double slit, the contribution of both slits must be included. Figure 2-8 shows the two-slit system and its parameters defined for the calculation of the diffraction pattern. The total contribution is the sum of integrals, taken over each slit (see Eq (2.10)), that is,

\[
\frac{D}{D} e^{i(kp_1-\omega t)} \left( \int_{-D/2}^{D/2} e^{-ik\sin \theta} dy + \int_{-a-D/2}^{-a+D/2} e^{-ik\sin \theta} dy \right). \tag{2.15}
\]

By introducing the assumptions \( y^* = y + a \) and \( \beta = k\sin \theta \), we can write Eq (2.15) as

\[
\frac{D}{D} e^{i(kp_1-\omega t)} \left( \int_{-D/2}^{D/2} e^{-i\beta y} dy + \int_{-D/2}^{D/2} e^{-i\beta(y^*-a)} dy^* \right). \tag{2.16}
\]

\[
= \frac{D}{D} e^{i(kp_1-\omega t)} \left( \int_{-D/2}^{D/2} e^{-i\beta y} dy (1 + e^{i\beta a}) \right)
\]
Fig. 2-8. Two-slit system and its parameters defined for the calculation of the diffraction pattern.

By factoring $e^{i\beta(a/2)}$, this becomes again based on Eq (2.11)

$$u = C \frac{\sin \left[ \frac{\beta D}{2} \right]}{\beta D/2} (e^{i(kp_1 - \omega t)}(1 + e^{i\beta a})$$

$$= C \frac{\sin \left[ \frac{\beta D}{2} \right]}{\beta D/2} e^{i(kp_1 - \omega t + \beta a^2/2)} e^{-\beta a^2/2 + e^{i\beta a^2}/2}$$

so that

$$u = 2C \frac{\sin \left[ \frac{\beta D}{2} \right]}{\beta D/2} e^{i(kp_1 - \omega t + \beta a^2/2)} (\cos \beta a^2/2). \quad (2.17)$$
For the intensity, we obtain

\[ uu^* = 4C^2 \left( \cos \beta_2^a \right)^2 \left( \frac{\sin \left( \beta_2^D \right)}{\beta_2^D} \right)^2, \]  \hspace{1cm} (2.18)

where \( \left( \sin \left( \beta_2^D \right) \right)^2 / \left( \beta_2^D \right)^2 \) and \( 4C^2 \left( \cos \beta_2^a \right)^2 \) represent the intensity distributions due to a single slit and the interference of waves from two slits placed apart at a distance \( a \), having a phase difference of \( \beta_2^d = \pi a \sin \theta / \lambda \), respectively. Figure 2-9 shows the resulted diffraction pattern by a double-slit grating.

![Graph showing diffraction pattern](image)

**Fig. 2-9.** Diffraction pattern of a double-slit grating (\( a=2d \)).

**Diffraction pattern of a multi slit**

In this section, the diffraction of waves with an array of \( N \) regularly spaced finite slits is considered. The slit width, \( D \), and the distance with which the slits are equally spaced (grating constant), \( a \), are defined in Figure 2-10 similarly to the two-slit system. The intensity distribution on the screen is obtained by summing all contributions from \( N \) slits. The contribution of the \( j^{th} \) slit to the amplitude on the screen is expressed using \( \beta = k \sin \theta \) as in the
two-slit system

\[ u_j = \frac{C}{D} e^{i(kp_1 - \omega t)} \int_{j} e^{-i\beta_{y_j}} dy_j \]  \hspace{1cm} (2.19)

The sum of the integrals for \( N \) slits represents the total amplitude, thus

\[ u = \frac{C}{D} e^{i(kp_1 - \omega t)} \left( \int_{-D/2}^{D/2} e^{-i\beta y} dy + \int_{0}^{a + D/2} e^{-i\beta y} dy + \int_{a - D/2}^{0} e^{-i\beta y} dy + \ldots + \int_{(N-1)a - D/2}^{(N-1)a + D/2} e^{-i\beta y} dy \right) \]  \hspace{1cm} (2.20)

The integral parts in the parentheses can be solved as

\[ \frac{2\sin \left( \frac{\beta D}{2} \right)}{\beta}, \ 2e^{-i\beta a} \frac{\sin \left( \frac{\beta D}{2} \right)}{\beta} \] and so on.

**Fig. 2-10.** Multi-slit array and its parameters defined for calculating the diffraction pattern.
As a result, Eq (2.20) can again be written as

\[ u = C e^{i(kp_1 - \omega t)} \frac{2 \sin \left( \frac{BD}{2} \beta \right)}{\beta} \left( 1 + e^{-i\beta a} + e^{-2i\beta a} + \ldots + e^{-(N-1)i\beta a} \right). \]  

(2.21)

\[ = C e^{i(kp_1 - \omega t)} \frac{\sin \left( \frac{BD}{2} \beta \right)}{\beta} \frac{1 - e^{-IN\beta a}}{1 - e^{-i\beta a}}. \]

Thus, the intensity is

\[ uu^* = C^2 \left[ \frac{\sin \beta \frac{D}{2}}{\beta} \right]^2 \left( \frac{1 - e^{-i\beta a}}{1 - e^{-i\beta a}} \right) \left( \frac{1 - e^{iN\beta a}}{1 - e^{i\beta a}} \right). \]  

(2.22)

\[ = C^2 \left[ \frac{\sin \beta \frac{D}{2}}{\beta} \right]^2 \left[ \frac{\sin N\beta a}{\sin \frac{\beta a}{2}} \right]^2. \]

When \( N=2 \), Eq (2.22) results in Eq (2.18) for the double slit. In order to normalize the central maximum to unit intensity, we can set \( C^2 = 1 \) as in Eq (2.14). Thus, the ratio of the intensity at a certain position on the screen, \( I \), to the intensity at the middle of the screen, \( I_0 \), is

\[ \frac{I}{I_0} = \left[ \frac{\sin \beta \frac{D}{2}}{\beta} \right]^2 \left[ \frac{\sin N\beta a}{\sin \frac{\beta a}{2}} \right]^2 \left[ \frac{\sin \left( \frac{\pi D}{\lambda} \sin \theta \right)}{\frac{\pi}{\lambda} \sin \theta} \right]^2 \left[ \frac{\sin \left( \frac{N\pi a}{\lambda} \sin \theta \right)}{\frac{\pi}{\lambda} a \sin \theta} \right]^2. \]  

(2.23)

where \( \theta \), \( \lambda \), \( d \), \( a \) and \( I_0 \) denote the angle of diffraction, wavelength, slit width, grating constant and intensity for \( \theta = 0 \), respectively. The first factor of Eq (2.23) represents the diffraction due to a single slit. The second one describes the interference factor, which is the superposition of \( N \) waves equally spaced with a distance \( a \).
In the case of $N=4$ and $a=2d$, a plot of the relative intensity, $I/I_0$ as a function of $\sin \theta$ is shown in Figure 2-11. The resulting diffraction pattern is the product of diffraction and interference factors so that it occurs within the envelope of the first factor of Eq (2.23). The principal maxima are at $(\pi a/\lambda)\sin \theta = m\pi$, where $m = 0, 1, 2, \ldots$. The positions of the diffraction intensity maxima for a given wavelength and grating constant can be derived from this. Thus,

$$\text{asin} \theta = m\lambda.$$  \hspace{1cm} (2.24)

This is called the grating equation. The integer $m$ is called the order of diffraction. By selecting $a=2d$, all even-order maxima ($m=2, 4, \ldots$) are eliminated in the diffraction pattern, as can be seen in Figure 2-11. This is often conveniently used to avoid overlapping of diffraction-order maxima which have different values in their wavelength. Consequently, the diffraction pattern from a multi-slit grating is characterized by

- the wavelength
- the grating constant
- the slit width of grating
- the number of slits.

![Graph showing diffraction pattern](image)

**Fig. 2-11.** Diffraction pattern of a multi-slit grating with $N=4$ and $a=2d$. 

---

*Infrared Micro-spectrometer based on a Multi-slit Grating*
A larger number of slits results in narrower peaks, a smaller grating constant yields a larger separation between the spectral lines and a smaller slit width produces a wider envelope. Because infrared has a longer wavelength, it overcomes the limitations of the lithography to realise a multi-slit grating that yields sufficient separation between the spectral lines for analysing them, and it results in wider dispersion angle with the same optical-path length and grating constant. Figure 2-12 shows a photo of the diffraction pattern formed by a single slit.

Fig. 2-12. Diffraction patterns for slit systems shown at the left [2.1].
2.2.3.3 Interferometers

There are a number of practical devices for optical spectral analysis known as interferometers. In an interferometer, the wavelength is selected by changing the difference of the optical path length between two or more beams which interfere with each other. An interferometer consists of two or more separate fully reflecting and semi-transparent mirrors that reflect and split or combine their optical signals coming from separate optical paths to produce interference. Figure 2-13 shows a simplified illustration of the most important interferometers.

The Michelson interferometer (see Figure 2-13 (a)) is one of the most important and commonly used devices [2.13][2.14]. The incident wave passes through the semi-transparent mirror or beam splitter of the

![Diagram](image)

**Fig. 2-13.** Schematic illustration of the most important interferometers: (a) Michelson, (b) Mach-Zehnder and (c) Sagnac (d) Fabry-Perot.
interferometer, which divides the beam into multiple optical paths. The beam splitter ideally transmits and reflects 50 percent of the incident wave without introducing any additional phase shifts. The divided beams subsequently travel to the mirrors, $M_1$ and $M_2$. After reflection from these mirrors, the beams recombine at the beam splitter into a single beam and travel to the detector. Only a specific single wavelength is transmitted to the detector and any other light travels back to the source.

The Mach-Zehnder interferometer consists of two beam splitters and two mirrors (see Figure 2-13 (b)). A difference between two optical paths after the first beam splitter can be introduced by slightly tilting one of the beam splitters. This interferometer is relatively difficult to align since the two paths are separated. A common application of this interferometer is observing the density variation in gas-flow patterns within research chambers [2.1]. The principle of the Mach-Zehnder interferometer is also used to demonstrate micro-scaled biosensors using silicon integrated-circuit (IC) technology [2.15].

The other instrument known for its easy alignment and stable operation is the Sagnac interferometer (see Figure 2-13 (c)). The main feature of this device is that there are two identical but oppositely directed closed-loop paths. Slightly shifting one of those mirrors will result in a path-length difference, and hence in a fringe pattern.

Among all different types of interferometers the Fabry-Perot interference filter is most frequently used, especially for miniaturized thin-film devices due to the simplicity of its structure and its preparation process. For example, miniaturized Fabry-Perot interferometer-based spectrometers were designed for the application of optical-fibre communications at 1.55 µm by Jerman et al. [2.16] and for optical filter for the visible part of the spectrum by Raley et al. [2.17]. A tunable bulk-micromachined and array-type Fabry-Perot micro-interferometer for the visible spectral range was demonstrated by Correia et al. [2.18][2.19]. These will be discussed later.

The Fabry-Perot interferometer usually consists of two parallel mirrors, so the gap between the mirrors is constant over the whole area. Light is repeatedly reflected and transmitted with little losses within the system; this is known as the Fabry-Perot optical resonator (see Figure 2-13 (d)). When an ideal Fabry-Perot resonator is assumed, the two mirrors must be perfectly
parallel and lossless. Also, no absorption occurs in the resonator medium. Non-idealities are accounted for in the finesse, defined as the ratio of the separation of successive peaks to the peak width at half maximum.

A standard micro-fabrication process limits both the lateral and vertical dimensions small and constrains shapes of the component rather flat. The plane plates Fabry-Perot configuration is therefore very suitable for the fabrication of wavelength-selecting elements for an integrated microspectrometer [2.20].

2.2.4 Optical path

In the design of an infrared system, it is necessary to consider the optical properties and parameters of the media which comprise the path of the radiation, for example, window, lens, filters and atmosphere. Mostly, the transmission or absorption and scattering characteristics of the media are the main parameters to be taken into account. For the application of infrared sensor systems especially in long atmospheric paths, a study of both absorption and scattering in the atmospheric media is very important. Some key absorbers include water vapour, carbon dioxide, ozone, nitrous oxide and carbon monoxide [2.21]. Figure 2-14 shows a typical atmospheric transmission as a function of the wavelength. The detailed theory about atmospheric transmission is quite complex and not relevant to the micro-

![Graph showing atmospheric transmission](image)

**Fig. 2-14.** Typical atmospheric transmission (1 km path length) [2.21].
Infrared spectrometers

![Graph showing intrinsic, extrinsic, and free-carrier absorption in a typical semiconductor.]

Fig. 2-15. Spectral absorption in a typical semiconductor [2.22].

devices, which are usually packaged in well-defined cleanroom air, and thus this discussion in omitted in this thesis.

Whereas most dielectric materials are transparent in the visible spectral range and are opaque in the infrared range, most semiconductors are opaque in the visible but transparent in a large part of the infrared range [2.22]. Figure 2-15 shows the absorptance of a typical semiconductor as a function of wavelength. In the wavelength region that is smaller than $\lambda_0$, the incident photons have enough energy to excite electrons across the forbidden energy band gap. At wavelengths greater than $\lambda_0$, which corresponds to the band gap energy, the absorption decreases dramatically. In extrinsic materials that have an energy state in the forbidden band, an incident photon causes the excitation of a carrier from the extrinsic energy state to a free conduction state, resulting in absorption. This is possible only when the material has been cooled down to prevent thermal excitation [2.22]. The free-carrier absorption at longer wavelengths occurs by the interaction between the radiation and free electrons or holes, as can be seen in Figure 2-15.

Silicon is highly transparent for wavelengths exceeding 1 $\mu$m, beyond which free-carrier absorption can be disregarded [2.23]. Therefore the bulk silicon can be positively used to define a optical path in integrated electro-optical micro-devices without etching of the silicon bulk.
2.2.5 Infrared detectors

There is a considerable variety in devices available for the detection and measurement of infrared radiation. These infrared detectors convert the electromagnetic radiant energy directly or indirectly into electric energy. Infrared radiation detectors can be classified according to the physical mechanisms or principles involved in the conversion of incident radiant flux into a measurable signal.

2.2.5.1 Photon types

The principle of operation for all photon detectors is that the incident radiation causes the excitation of bound carriers into mobile states, and this transition can be monitored as an electrical output signal. Photon detectors respond faster than thermal type detectors that are discussed later since they interact directly with the incident radiation to generate electrical carriers instead of changing temperature in thermal detectors. However, photon detectors are strongly wavelength dependent while thermal-type detectors show a flat spectral response over a wide range.

There are several methods to monitor the change within a device caused by IR radiation, depending on the device structure. The most common methods are to measure the change in resistivity of a photoconductive element (photoconductors) and the voltage generated across a junction (photovoltaic detectors). Photon detectors also include photoemissive detectors, which are different from photoconductors and photovoltaic detectors in that electrons physically leave the detection material when a photon is absorbed.

Photoconductive detectors

Photoconductive detectors measure a change either in voltage or in current across the sensitive element, usually a load resistor. The incident radiation increases the number of free-charge carriers in the sensitive element so that the effective resistance of the detector decreases. Photoconductive detectors require a read-out amplifier which biases the detectors with a constant current; this causes higher on-chip power dissipation. The incident photon must have an energy greater than the energy gap of the semiconductor material used for the sensing element, that is, photoconductors have
wavelength cut-offs, which is always a long-wavelength threshold. With intrinsic photoconductors, the operating wavelength range is set by the energy gap of the semiconducting material. One method of extending the wavelength range is the use of extrinsic photoconductors. In these extrinsic devices the carrier level is either close to the conduction band or to the valence band. The long-wavelength limit is determined by the impurity ionization energy [2.24]. Thus, an extrinsic photoconductor can work for a longer wavelength. The major disadvantage of extrinsic devices is that they require a much lower operating temperature than intrinsic devices. And the quantum efficiency is lower for extrinsic devices because the dopant is less abundant than the host material. For detecting even longer wavelengths, free-carrier devices are employed. These devices use semiconductors with very high carrier mobilities. The incident radiation causes intraband transitions of the electrons in the conduction band, changing the electron's mobility and hence the resistance of the material. However, this device must be cooled down to liquid-helium temperature to reduce the rate of transfer of irradiative energy to the crystal lattice [2.24]. The limits of the performance of intrinsic and extrinsic infrared photoconductors are discussed by M. M. Blouke et al. [2.25].

Photovoltaic detectors

Junction or photovoltaic detectors have a junction that consists of two different types of material. In general, it is fabricated with p-type and n-type regions formed in the same material. The incident photons are absorbed in the material and create electron-hole pairs, and the minority carriers diffuse to the junction where they are swept through the depletion region. Consequently, a forward biasing is formed at the junction, producing either an open-circuit voltage or a short-circuit current. Figure 2-16 shows the current-voltage characteristic of the photovoltaic devices. When there is no incident radiation on the device a reverse current, called the dark current, is observed. If the detector is illuminated with a proper wavelength, the I-V characteristic is displaced and the signal can be measured, as can be seen in Figure 2-16. Photovoltaic detectors operate in the diode's reverse bias region; this minimizes the current flow through the device which in turn minimizes the power consumption. In addition, photovoltaic detectors have low noise because the reverse bias diode junction is depleted of minority carriers.
In order to achieve a higher efficiency, one should use a front surface with a high absorption coefficient and low recombination velocity, and a junction depth that is small compared to a diffusion length. Most early junction detectors have homo-junction structures, that is, p- and n-type regions are prepared in the same material. To achieve a larger gain with higher efficiency, devices with a hetero-junction or multi-layer structure have been developed. Hetero-junction detectors consist of two different semiconductors that have similar lattice structures, whereas a multi-layer structure has a multi-graded band gap. Photovoltaic detectors that have the highest performance are fabricated with mercury cadmium telluride (HgCdTe) and indium antimonide (InSb) semiconductors [2.27][2.28][2.29]. Especially the HgCdTe, introduced in the late 1950s in Great Britain, is today most widely used of all tunable band-gap materials. Using the relation $E = \frac{hc}{\lambda}$, the band-gap of silicon (1.11 eV) corresponds to a wavelength of 1.13 μm. Materials with a smaller band-gap, such as InSb and HgCdTe, are thus needed. The spectral range of InSb detectors extends from 1 to 5 μm, while the range of HgCdTe detector is dependent on the doping concentration. As spectral responsivity extends further into longer wavelengths, more cooling is necessary to prevent the contribution of thermally excited electrons.
**Infrared spectrometers**

**Photoemissive detectors**

Photoemissive detectors usually have a photocathode placed in a vacuum chamber with high voltage. The surface of the photocathode emits free electrons when it absorbs the photon from exposure of light with sufficient energy. The electrons are accelerated to the anode placed on the opposite side in the vacuum chamber by a positive electrostatic field. Once the electrons reach the anode, they are measured as corresponding current representing the amount of light on the photocathode. For rapid detection of light or larger gain to sense very low light levels, a photomultiplier tube (PMT) can be applied. In a PMT, the accelerated electrons fall upon a metal surface where they emit secondary electrons that are again accelerated to generate more electrons at the next metal surface and so on. This multiplication occurs through a number of metal electrodes, resulting in a large number of electrons collected on the anode. The whole arrangement thus acts as a combination of a simple photo cell with a high-gain amplifier in a self-contained unit.

Photon detectors have a higher sensitivity than thermal-type detectors, which will be discussed in the next section, although photon-type ones require bulky and expensive cooling systems and they have very sharp cut-off operating wavelengths. Nevertheless, for very-high-performance applications and exact wavelength operations the solid-state photon detectors are preferable.

**2.2.5.2 Thermal types**

Thermal-type IR detectors show a change in some measurable property that results from a temperature change of the sensitive element, caused by absorption of IR radiation. Thermal detectors therefore utilize a material with a strong temperature-dependent property, such as electrical conductivity or thermal expansion. The most common detection mechanisms of thermal-type IR detectors are the resistive bolometer, the pyroelectric detector and the thermoelectric detector. Figure 2-17 shows the mechanism of thermal-type infrared sensors. A thermal-type or uncooled IR detector consists of three parts: a thermal isolation structure, an IR absorber (sensitive element) and thermometer. Infrared radiation falls on the sensitive element connected to the substrate by thin supporting legs for thermal isolation. The sensitive element is heated up by the incoming IR radiation.
Finally, the temperature change of the sensitive element is usually measured by:

- resistance change (bolometer),
- pyroelectric effect,
- thermoelectric effect (thermopile-based detector), or
- gas pressure change (Golay cell).

Among these, micro-bolometers, pyroelectric and thermoelectric devices are the most popular uncooled infrared detectors. Flannery et al. [2.26] and Marasco et al. [2.30] discussed resistive bolometer and pyroelectric device in more detail. Many reports also investigate the thermoelectric device and attempt to improve its performance [2.31][2.32][2.33][2.34].

The most important factor in the design of thermal detectors is thermal isolation between the sensitive area and the substrate for a high performance [2.35]. In general, the supporting legs are very long and thin to reduce the value of conductance. However, the thermal isolation acts negatively for the response time since the thermal response time $\tau$ can be defined as

$$\tau = \frac{C}{G}. \quad (2.25)$$

For a good thermal response time, the heat capacity of the sensitive element must be low enough. This requirement always needs a trade-off between low conductance and low heat capacity. Because the pixel area is defined by the system requirement, the only adjustable parameter is the thickness. The use of a monolithic thin-film technology is therefore desirable to realize a device with a high performance.
Resistive bolometers

A bolometer contains a resistance element that changes in resistance when heated by the radiant energy. The device is operated by passing an accurately controlled bias current through the resistance element, and monitoring the output voltage. The temperature change of the bolometer due to the absorption of IR radiation is usually so small that the resistance change can be assumed to be linear with the temperature change. Assuming the output signal is the voltage, it can then be written as [2.35]

\[ V_{out} = i_b \alpha r \Delta T, \]  

(2.26)

where \( i_b \) is the bias current, and \( \alpha, r \) and \( \Delta T \) the temperature coefficient, resistance of sensitive element and temperature increase due to IR radiation, respectively. When the radiation falling on the sensitive area is sinusoidally modulated with angular frequency \( \omega \), the responsivity is given by

\[ R = \frac{i_b \epsilon \alpha r}{G \sqrt{1 + \omega^2 \tau^2}}, \]  

(2.27)

where \( G \) and \( \epsilon \) are total thermal conductance and optical absorptance, respectively.

The temperature coefficient of resistance can be either positive or negative. At room temperature, for example, it is positive for metals and normally negative for semiconductors.

Pyroelectric detectors and ferroelectric bolometers

Below a temperature \( T_c \) known as the Curie point, pyroelectric materials exhibit a large spontaneous electrical polarization. If the temperature of such a material is altered, for example by incident radiation, the polarization changes, as can be seen in Figure 2-18. This change in polarization may be observed as an electrical signal if electrodes are placed on opposite faces of a thin slice of the material to form a capacitor. When the polarization changes, the charges induced in the electrodes can be made to produce a voltage across the slice if the external impedance is comparatively high. The sensor will only produce an electrical output signal when the temperature
changes; that is, when the level of incident radiation changes.

There is a variant of the pyroelectric effect; the field-enhanced pyroelectric effect also known as the ferroelectric bolometer effect. If the magnitude and direction of the spontaneous polarization can be reversed by an external field, then such a material is said to show ferroelectric behaviour. All single pyroelectric crystals which show ferroelectric behaviour are pyroelectric, but the opposite is not necessarily true [2.39].

The main difference between pyroelectric and ferroelectric materials is that the direction of the spontaneous polarization in ferroelectric material can be switched by applying electric field. Ferroelectric materials show a remanent polarisation, which can be oriented and switched by an electric field due to an applied voltage [2.39]. That means the polarization does not fall to zero when the external field is removed. Therefore, an electric field in the negative direction should be applied to depolarize the crystal.

**Fig. 2-18. Pyroelectric effect.**

**Thermoelectric detectors**

The thermoelectric effect takes place in a circuit consisting of two different electrically conducting materials that are joined at two points, the so-called thermocouple. The thermocouple detects the temperature gradient between the joined points, hot and cold junction by the Seebeck effect [2.40]. The signal form is an electromotive thermoelectric voltage when the circuit is opened. Several thermocouples can be connected in series to form a
Infrared spectrometers

Fig. 2-19. Three thermocouples connected in series electrically and in parallel thermally.

thermopile with a higher output voltage for the temperature difference between the hot and cold junctions, \( T_1 \) and \( T_2 \) (Figure 2-19). The magnitude of the voltage also depends on the type of material and the temperature difference between the junctions.

In order to employ the thermoelectric effect as an IR detector, hot junctions of thermopile are placed on a thermally isolated structure, usually on a thin film prepared by deposition, whereas the cold junctions are deposited on the surrounding bulk substrate. The thermoelectric voltage \( V_{out} \) is given by

\[
V_{out} = N(\alpha_1 - \alpha_2)\Delta T,
\]

(2.28)

where \( \alpha_1 \) and \( \alpha_2 \) are Seebeck coefficients and \( N \) represents the number of thermocouples connected in series. The responsivity of a thermoelectric detector is [2.34]

\[
R = \frac{\varepsilon N \alpha}{G \sqrt{1 + \frac{\omega^2 \tau^2}{2}}} \left[ \frac{V}{W} \right],
\]

(2.29)

where \( \varepsilon \) is optical absorptance.

Thermoelectric detectors have a low responsivity compared to resistive bolometers and pyroelectric detectors. However, they operate unbiased,
Fig. 2-20. Schematic structure of thermopile-based infrared detector.

show no current induced noise [2.35] and are relatively simple to prepare in a compatible way with an IC process. Figure 2-20 shows the schematic structure of a thermopile-based infrared detector cell formed on a thin membrane for the thermal isolation of a sensitive element.

Golay cell

Golay cell uses the concept of monitoring the pressure change in a gas due to radiation heating. The Golay cell has a gas chamber and mostly a parallel planar capacitor consisting of a thin diaphragm formed by bulk-micromachining. Infrared radiation heats the gas in the chamber and the resulting deflection of the diaphragm induced by the expansion of the gas is detected as a capacitance change of the capacitor [2.36], or as a quantum mechanical tunnelling current between a fixed tip and the moving membrane [2.37]. The gas chamber of the device is completely sealed and can be filled with various kinds of gas that have different thermal expansion coefficients. The Golay cell is one of the most efficient detecting devices offering excellent sensitivity and response time at room temperature in a wide wavelength range, due to the large thermal expansion coefficient and the small heat capacity of the small volume of gas.

Bimaterial element detector

A microcantilever-based thermal detector array employing thermally sensitive bimaterial elements has been developed by combining surface
micromachining and conventional IC process [2.38]. IR radiation on the bimaterial element causes mechanical movement of capacitive cantilever plate due to the differential expansion.

### 2.2.5.3 Performances of different types of IR detector

The detectivity, which is defined as the rms incident radiant power required to produce an output signal equal to the detector noise level, is not an ideal parameter for comparing different detectors because it varies depending on both the sensitive area and the bandwidth. The specific detectivity, expressed as $D^*$, is a very useful parameter for comparing detectors of different sizes (see Chapter 3). Figure 2-21 shows the specific detectivity as a function of wavelength for different types of detectors. The spectral response of the thermal-type IR detector is independent of the wavelength of the incident radiation and hence thermal detectors have a flat response over a very wide spectral range. The limiting feature in the spectral range is the transmittance of the window material used to manufacture the sensor housing. By using different windows materials it is possible to detect radiation at different frequencies. This contrasts with the spectral responsivity of photon detectors, which has sharp cut-off wavelengths.

![Normalized detectivity for different types of detectors as a function of wavelength](image)

**Fig. 2-21.** Normalized detectivity for different types of detectors as a function of wavelength [2.41].
2.2.6 Classification of Infrared spectrometers

Spectrometers are often classified according to the type of dispersive or wavelength-selective elements employed or the physical effect producing dispersion. In this section, infrared spectrometers classified by the type of dispersing elements are briefly introduced.

Out of the enormous efforts of the early researchers for IR spectroscopy came the recognition that each compound had its own unique IR absorption pattern even in different molecules. However, the real use of IR absorption pattern by chemists was delayed until the 1940s, because of the difficulty in measuring the spectra. The rapid advance in the science of electronics made it possible to use electronic amplification and to record very small signals from a tiny thermocouple in an IR spectrometer. The next significant improvement came with the technique for making a thermocouple detector that has a higher response time. This improvement made it possible to eliminate drift in the spectrometer record and opened the door to the development of double-beam spectrometer [2.42]. Moreover, these results attracted more researchers to develop more advanced IR spectroscopy techniques and expanded its potential and popularity rapidly.

During the past several decades the infrared spectroscopy technique has become increasingly important and useful in qualitative and quantitative analysis due to its powerful and versatile performance with many different instrumental techniques.

2.2.6.1 Prism-based spectrometer

Infrared spectrometers with dispersive elements based on prisms have been commercially available since the 1940s and diffraction gratings were introduced as alternative dispersive elements in the mid 1950s. The popularity of prism-based instruments fell away in the late 1960s when the improved technology to manufacture gratings promised a mature quality at low cost.

Figure 2-22 shows a classical mounting of a prism monochromator in which the collimator, prism(s) and focusing lens are simply mounted. Figure 2-23 shows the optical layout of a Littrow mounting [2.42]. A double-beam
instrument is used in some cases. By using the double-beam arrangement on can divide the radiation from a source into two beams. Each of these beams passes a sample and a reference path, respectively. The information from these beams is then ratioed to obtain the sample spectrum. In actual instruments, the optical path and principal components are much more complex. The Littrow mounting has the advantage of compactness compared with the classical mounting having the same path length. This mounting is employed much less frequently with a prism than with a grating for the dispersing element.

Prism-based spectrometers show a relatively low performance, which results from the low luminosity, so that it is inadequate to satisfy most of the
requirements of spectroscopy, despite its advantage of being more simple [2.10]. In the next section, the grating-based spectrometer is introduced which has higher resolving power at equivalent luminosities.

2.2.6.2 Grating-based Spectrometer

The optics of a grating monochromator in a typical mounting is shown in Figure 2-24. Light from the source passes through a narrow entrance slit and is made parallel by the collimating mirror. Dispersion occurs when radiation falls on the grating through the entrance slit and the collimating optics. The grating then reflects the beam back to the exit slit. The dispersed spectrum is scanned across the exit slit by rotating the mirror or grating, depending on the instrument. The width of the entrance and exit slits can be varied and programmed to compensate for any variation of the source energy with wave number.

The general characteristics of the grating spectrometers using plane grating are similar to those of the prism-based spectrometer. Nevertheless, gratings usually give a considerably better resolution than that obtained with prisms. A grating spectrometer has a luminosity at least 10 times greater than that of a prism spectrometer of the same effective resolving power if the dispersing elements are of the same size [2.10]. The angular dispersion of gratings is 3 to 30 times greater than that of prisms in the infrared region [2.11]. A very high-resolution grating spectrometer has been realised by Lindblom et al. [2.43].

![Fig. 2-24. The basic optical configuration of a grating spectrometer.](image-url)
Rowland demonstrated a grating ruled on a spherical concave surface that can form a spectrum without any auxiliary focusing optics. The spherical grating shows a combined function of dispersion and focusing. The Rowland concave gratings therefore are beneficial in that they will focus a spectrum without the aid of lenses, however, one disadvantage is that they suffer from astigmatism [2.10]. There are many different types of gratings and monochromator mounts designed to take advantage of the unique property of the Rowland circle.

2.2.6.3 Fourier-transform Infrared (FT-IR) Spectrometer

The most significant advances in infrared spectroscopy have come with the introduction of Fourier-transform infrared (FT-IR) spectrometers. The FT-IR spectrometers have replaced the dispersive instruments, as computers have become much more compact, powerful and cheaper since 1980. The FT-IR spectroscopy has dramatically improved the quality of infrared spectra and has minimized the time required to obtain data. In an experimental measurement characterized by random noise, the signal-to-noise ratio (SNR) can be improved by monitoring a larger number of resolution elements simultaneously. The signal will then increase in proportion to the number of measurements, but noise increases only with the square root of the number of resolution elements being monitored [2.10][2.42]. In addition, because the FT-IR spectrometer does not require the use of a slit or other restricting device, the total source output can be passed through the sample continuously. This results in a higher substantial gain. The popularity of FT-IR is primarily due to the multiplex and rapid scanning features of interferometer. The term multiplex is adopted from communication theory, where it describes the system of transmitting many sets of information simultaneously over the same channel. Multiplex devices in spectroscopy therefore are those systems in which a single detector receives many signals simultaneously from different elements of the spectrum.

The difference, for example in resolution, between the grating spectrum used in the 1960s and the Fourier transform spectra recorded in 1973 using the Hale telescope at Mount Palomar is a factor of 500 [2.12], as shown in Figure 2-25.
Fig. 2-25. Improvements in the near infrared Venus spectrum due to Fourier transform spectroscopy [2.12].
Infrared spectrometers

Figure 2-26 shows a schematic diagram of a two-beam Michelson interferometer used in most of the recent FT-IR instruments. The movable mirror creates an optical-path difference between the two beams reflected from the fixed and moving mirrors, respectively. Depending on the path difference created by the displacement of the movable mirror, the two beams interfere constructively or destructively. As the movable mirror moves, the intensity falling on the detector varies as the cosine wave shown in Figure 2-27 (a) for the case where the radiation source is monochromatic. The intensity $I(x)$ on the detector can be represented as a function of frequency $\nu$ in cm$^{-1}$ as follows

$$I(x) = B(\nu) \cos(2\pi \nu x) ,$$  \hspace{1cm} (2.30)

where $x$ is the mirror displacement in centimetres and $B(\nu)$ represents the intensity of the source.

When the source is polychromatic the intensity at the detector can be written as

$$I(x) = \int_{-\infty}^{\infty} B(\nu) \cos(2\pi \nu x) d\nu ,$$  \hspace{1cm} (2.31)

which is one half of the cosine Fourier transform pair [2.42]. The other half is

Fig. 2-26. A Michelson interferometer
Eq (2.31) represents the variation in intensity as a function of path-length difference and Eq (2.32) shows that as a function of wave number. These two equations can be converted into each other by Fourier transformation. The transformation is done by a computer. The spectrum of the source can be subsequently derived by carrying out a Fourier analysis of the output signal.

Recently, a miniaturized Fourier transform spectrometer (FTS) based on optical MEMS technology has been reported by Manzano et al. [2.44]. The device is basically a Michelson interferometer with one scanning mirror actuated by an electrostatic comb drive actuator and claims a resolution of 6 nm at 633 nm.
2.3 Conventional macro-spectrometer vs. integrated micro-spectrometer

Infrared spectroscopy finds widespread application in industry, and in biological and pharmaceutical research. Moreover, spectral analysis with high performance is possible as computers have become more compact and much more powerful at less cost. However, conventional devices involve complex lens systems and moving parts, and are thus bulky and expensive. Many consumers and industrial applications do not require the nm-range resolution that such an instrument offers, and the advantages of a miniaturized micro-spectrometer, such as compact size, reduced weight, simplified assembly, more stable alignment, and low cost prevail. Therefore, microfabricated spectrometers can achieve acceptable performance in low-cost applications [2.45].

Many different types of miniaturized spectrometers have been recently realised. Jerman et al. have reported a miniaturized spectrometer based on a Fabry-Perot interferometer [2.16]. Two wafers are independently processed and bonded. One wafer has a silicon mass, on which a highly reflective dielectric mirror and metallic electrodes are prepared, surrounded by a corrugated diaphragm suspension. The other wafer has a dielectric mirror and a set of electrodes that electrostatically control the distance between the parallel mirrors (see Figure 2-28). The space between the two mirrors is an

![Diagram](image)

**Fig. 2-28. Fabry-Perot interferometer-based spectrometer [2.16].**
optical resonance cavity and results in a transmission of the wavelength that is tuned by the cavity. Electrostatic actuation makes it possible to tune the wavelength. Such a device can be operated in the near-infrared region, at wavelengths greater than approximately 1μm where the silicon is highly transparent. Voltages between 0-70V are required to tune the device over the entire operation range (about 37 nm).

A Fabry-Perot interferometer working in the visible region of the spectrum has been reported by Raley et al. [2.17] (see Figure 2-29). The device has a limited spectral range (about 50 nm) and requires a high operating voltage (100V).

A miniaturized planar waveguide spectrometer, which is a few millimetres thick, has been presented for the application of chemical analysis by Goldman et al. [2.46] (see Figure 2-30). The device consists of two diffraction gratings etched into the surface of substrate and a planar waveguide. The incident light focused through the substrate onto one of the gratings is coupled into the waveguide and propagated to the second grating. A silicon photodiode array is used to detect the spectral data. Since a silicon substrate restricts the spectral range to the NIR, a glass slide is used as the substrate.

A planar grating spectrometer has been realized by Mohr et al. [2.47] (see Figure 2-31). Incident light is introduced by an optical fibre into the device and then dispersed at a reflection grating formed by the deep X-ray lithography technique.

![Diagram](image.png)

**Fig. 2-29.** Fabry-Perot interferometer based spectrometer for visible wavelengths [2.17].
An integrated silicon spectrometer for the visible range has been reported by Kwa et al. [2.48] (see Figure 2-32). The device consists of two silicon chips prepared by bulk micromachining technology after the IC processing. The top part has both optical and electronic components, a grating to disperse the incident light into its spectral components and a photodetector array to detect the dispersed spectrum. The bottom part consists of mirrors to extend the optical-path length. The multiple reflection at the lower wafer
Conventional macro-spectrometer vs. integrated micro-spectrometer

![Image of spectrometer components]

**Fig. 2-32.** Integrated silicon spectrometer for visible spectral range [2.48].

is critical so that the surface of the mirror was smoothed to reduce the scattering at the etched surface. The spectral resolution was obtained as R=10 for the wavelengths between 380 and 730 nm.

Yee *et al.* presented a low-cost micro-spectrometer based on a micromachined grating mounted above a CCD imaging device [2.49] (see Figure 2-33). The measured resolution is 69.8 at 632.8 nm.

All these studies demonstrated the feasibility of realizing miniaturized spectrometers in silicon by applying micromachining technology. Fabry-

![Image of spectrometer components]

**Fig. 2-33.** A spectrometer based on a grating above a CCD array detector for biochemical analysis [2.49].
Perot resonance cavities and gratings have been applied successfully as dispersive elements in micro-spectrometers. Also the fabrication of arrays of integrated silicon bolometers and thermoelectric detectors has been demonstrated.

Spectrometers are basically composed of three components: (1) a dispersive element, (2) a detector or array of detectors and (3) an optical path in between, with lens systems for optical signal conditioning (collimating and focusing) of the incident and refracted light beam. Scaling limitations of each of these components should be considered when downscaling a spectrometer concept to the microsystem level. The main impediment for a fully integrated silicon optical micro-spectrometer is the definition of a sufficiently long optical path and the implementation of lenses.

An array-type micro-spectrometer based on an array of Fabry-Perot optical thin-film filters (etalons) of different resonance width, resulting in 16 channels, has been developed and characterised by Correia et al. [2.20] (see Figure 2-34). Both standard bipolar and CMOS technology were used to

**Fig. 2-34.** A CMOS optical microsystem based on an array of Fabry-Perot etalons integrated with photodiodes underneath the 16 etalons, read-out electronics and a bus interface.
fabricate this single-chip spectrometer. The device has fixed cavities with different widths. Each of the 16 channels is sensitive to a single peak. Such a device is extremely suitable for applications in microsystems such as μ-TAS (μ-Total Analysis Systems), because of its small size, high spectral selectivity and low cost.

In the infrared micro-spectrometer presented here, a simple lensless configuration is pursued with a silicon wafer to serve as the optical path. Such a device yields sufficient spectral resolution (typically 10-20 bands within a selected spectral operating range) for use in a moderately demanding application. And it is rather simple to fabricate and remains compatible with standard IC processing.
References

[2.9] B. Stuart, Modern infrared spectroscopy, John Wiley & Sons Ltd. 1996.
Conventional macro-spectrometer vs. integrated micro-spectrometer


Infrared spectrometers


Infrared Micro-spectrometer based on a Multi-slit Grating


Infrared spectrometers
3.1 Introduction

As mentioned in the introductory chapter, a micro-spectrometer could open the way to a range of new applications. Its performance as compared to that of the conventional macroscopic infrared spectrometer is limited, at least with the present micromachining technology, due to a few technological constraints. The micro-spectrometer has the following most significant technological constraints:

1) lateral and vertical dimensions of the device are restricted by the standard die size and wafer thickness;
2) the shape of optical or mechanical components used for the device is limited to more or less flat and thin;
3) its fabrication process must be compatible with standard IC processing.

It is especially difficult to integrate lens systems into the standard wafer dimension with standard IC processing. With the aid of an externally placed
Design of the multi-slit grating-based infrared micro-spectrometer

slit and collimating lens, however, the light source can be conditioned to radiate a parallel beam of light on the dispersive element. This is in principle already a departure from the fully integrated concept. There are some spectrometers that are realised by micromachining technology without focusing lenses, as introduced in Chapter 2 [3.4][3.5]. These micro-spectrometers can be used in a particular type of industrial application, namely that which does not require the nm-range resolution offered by conventional instruments.

3.2 Basic design considerations

3.2.1 Device structure and principle

The device is functionally divided into two parts, as schematically shown in Figure 3-1. The first part is the top wafer, which contains the dispersing element, a multi-slit grating, realised by a standard metallisation step of IC processing. The second part is the bottom wafer, which has the thermopile-based infrared detector array fabricated on a silicon nitride membrane using bulk micromachining. The hot junctions of the thermopile are located inside the membrane and cold ones at the rim area of the membrane for thermal isolation between them.

These two independently processed wafers can be aligned and bonded during: (1) the final step, or, as an alternative, (2) prior to KOH etching of the detector array wafer. The first method requires a more careful treatment of the detector array wafer with the thin membrane, especially during the cleaning and drying steps, but it gives more independent process control between the two wafers, grating and detector array wafers. The second method fully preserves the fabrication compatibility by keeping the KOH etching step away from the standard processing. A failure during the KOH etching by buckling of the membrane or something else, however, causes the failure of the entire device at the final phase of fabrication.

The multi-slit-grating-based dispersive element deflects the spectral components that are present in incident radiation, depending on the wavelength. The diffracted spectrum is projected onto the infrared detector
array after traversal of the bulk silicon optical path. Every infrared detector in the array corresponds to a single wavelength, in practice, to a range of wavelengths. Consequently, a specific spectral component or a range of wavelengths is measured at a certain infrared detector among the array with respect to a defined optical path length, as can be seen in Figure 3-1. Therefore, it is essential that the infrared detector array is positioned exactly right relative to the grating. In the next sections, a few complications caused by this are discussed.

3.2.2 Optical path

The optical-path length is the most decisive parameter in the design of the micro-spectrometer with respect to the achievable resolution. As will be discussed later, bulk silicon can be used as an optical path medium for infrared light since it is highly transparent at wavelengths beyond 1 μm. By
using the thickness of bulk silicon wafer as it is, one can well define the optical path length. The variation in the wafer thickness, however, can cause a deviation in the optical path length, resulting a mismatch between the dispersed distance that we aim at and its corresponding IR detector. In addition, the higher refractive index of bulk silicon comparing to that of air would reduce the refraction angle for a given grating constant and optical-path length, as will be discussed later.

When the precise thickness of the wafer is known, the deviation in the optical path length can also be calibrated by controlling the etching depth during KOH etching, which is used for the recessed part at the back side of the grating wafer.

In the design one can triple the optical-path length or extend it even more by using the multi-reflections of the incident radiation between the dispersing element and the detector array. In this thesis, though, we mainly study direct projection types.

**Fraunhofer and Fresnel diffraction**

When an aperture is illuminated by plane waves and there is an opaque screen behind the aperture with a distance, the image of the aperture projected on the screen shows a different pattern depending on the distance between the aperture and the screen as shown in Figure 3-2.

With Fresnel or near-field diffraction the distance between the aperture and the screen is so close that the image of the aperture has some fringing around its periphery but is still easily recognizable. A continuous change in the fringes occurs when the distance becomes greater. At a great distance, the projected pattern will spread out considerably with little or no resemblance to the actual shape of the aperture. This is Fraunhofer or far-field diffraction.

In practice the optical behaviour is described by Fraunhofer diffraction in case of a sufficiently large distance between the slit and screen, or when the slit width \( d \) is much smaller than the wavelength \( \lambda \) of the illuminating light on the slit, that is, [3.7]

\[
R > \frac{d^2}{\lambda},
\]

(3.1)
**Basic design considerations**

*Fig. 3-2. A succession of diffraction patterns by a single slit with different distance between the slit and screen; Fresnel diffraction at the bottom and Fraunhofer diffraction at the top [3.6].*

where $R$ is the Rayleigh distance, defined as the smaller one of the two distances from light source to slit and slit to screen. The design of our IR micro-spectrometer, however, assumes that the light source is conditioned to illuminate a parallel beam of light on the dispersive element with the aid of externally placed optical components such as slit and collimating lens. It means that $R$ can be considered as the distance between the dispersive element and the screen, where the IR detectors are located, because the light source radiating a plane wave is at infinity (even though it is with the aid of a lens).

Consequently, the thickness of one single wafer (525 μm) used as the optical path medium is sufficient, in spite of the 3-4 times higher refraction index of silicon, to result in Fraunhofer diffraction (see equation (3.1)), although multi-reflection can result in a better resolving power.

*Infrared Micro-spectrometer based on a Multi-slit Grating* 65
Surface smoothness for less scattering

The surface roughness at the opened slit of the diffraction grating on the front side and at the recessed silicon surface by KOH on the back side of the grating wafer causes scattering of the incident light and diffracted light, respectively. In order to minimize scattering at the surfaces, we chose a double-side polished wafer for the grating wafer. The smoothness on the back side of the grating wafer is necessary not only for less scattering but also for the wafer-level bonding with the detector array wafer.

The measured initial roughness of the polished wafer surface was less than 10 nm peak to peak and the roughness of the etched silicon surface in KOH solution was about 50 nm at optimised condition: i.e., concentrations between 25 and 50 wt% and temperatures between 50 and 60 °C. Figure 3-3 shows the surface roughness of the etched silicon membrane as a function of KOH solution concentration at different temperatures [3.8]. Findler et al. [3.9] also reported a minimum surface roughness with a concentration and temperature of the KOH solution during etching of 30 wt% and 50-60 °C, respectively. An increase in the temperature does not affect surface roughness significantly, but does increase the etch rate [3.10]. KOH etching of the recessed part is thus done in 33 wt% KOH solution at 85 °C for less scattering, as will be discussed in Chapter 4.

*Fig. 3-3. Surface roughness of etched silicon in KOH solution as a function of the concentration [3.8].*
Once the surface roughness is known, the scattering can be calculated using the relation between the root-mean-square (rms) roughness and the total integrated scattering (TIS), or vice versa. It is expressed as [3.11]

\[ TIS = \left( \frac{4\pi \delta}{\lambda} \right)^2 \quad \text{or} \quad \delta = \frac{\lambda}{4\pi \sqrt{TIS}} , \tag{3.2} \]

where \( \delta \) denotes the rms roughness and \( \lambda \) the wavelength. As can be seen in equation (3.2), the longer the wavelength of infrared light, the less scattering. If the surface roughness \( \delta \leq 50 \, \text{nm} \), the TIS will be \( \leq 0.044 \) for 3 \( \mu \text{m} \) of wavelength. A shorter wavelength, for example 1 \( \mu \text{m} \), results in TIS \( \leq 0.4 \) for the same roughness. Consequently, the scattering problem is much less significant in the infrared spectral range than in the visible range.

### 3.2.3 Alignment

A lateral misalignment of the detector array with respect to the grating would result in the erroneous projection of the radiation onto another detector cell in the array than the expected one. However, some lateral misalignment and deviation in optical-path length can be accommodated, through an initial single-wavelength calibration, even after completion of the fabrication.

---

**Fig. 3-4.** Tilted projection of the incident radiation on the detector array due to angular misalignment.
Angular misalignment causes more serious problems. A large angular misalignment would result in a simultaneous illumination of several neighbouring detectors of the array by one single spectral component or range, causing a decreased resolution, as shown in Figure 3-4.

Sufficiently precise lateral and angular alignments are easily achieved with the present single-wafer IC technology. More care is required when two separately processed wafers are bonded. However, even the inaccuracy caused by the double-sided aligner used for the bonding (less than 2 μm) does not result in a serious misalignment, since the infrared detector size is relatively large (100 μm), as shown later.

3.2.4 Lensless spectrometer

In general, a spectrometer consists of components such as an entrance slit, a collimating lens system, dispersive element, focusing lens system and detector. The first two components, the entrance slit and collimating lens, transform spherical waves from an external light source into a parallel one, which enables light to enter the dispersive element without the influence of wavelength-independent variables of the source. In the design of a lensless micro-spectrometer, the entrance slit and collimating lens system are mounted externally.

The dispersive element that deflects the spectral components of the incident light at an angle depending on the wavelength is the most principal part in the spectrometer. A particular wavelength present in the incoming radiation results in an image formed with a corresponding diffraction angle. A multi-slit type of amplitude diffraction grating can easily be realised by a regular metallisation step in the standard IC processing, as will be discussed later.

Unfortunately it is very difficult to integrate the focusing lens system into the micro-spectrometer that is placed between the dispersive element and detector. As an alternative method, spectrometers can be designed having a reasonably long optical path to reduce the affection of the focusing lens, since then the focusing lens is less important as the distance between the dispersive element and the detector is increased. Nevertheless, the spectral resolution remains inferior to what can be achieved with a spectrometer that is equipped with a focusing lens system. As mentioned before, we decided.
to go for a lensless system.

The detector, finally, measures the intensity of the spectral components in the diffracted light as a function of the position, corresponding to a designated wavelength.

3.3 The multi-slit grating

3.3.1 Diffraction by a multi-slit grating

An obstacle placed between a point source and a screen casts an intricate shadow consisting of bright and dark regions quite unlike anything expected from the aspects of geometrical optics. This deviation of light from rectilinear propagation is called as diffraction. The effect is a general characteristic of wave phenomena occurring whenever a portion of a wavefront is obstructed in some way by either a transparent or opaque obstacle. The various segments of the wavefront that propagate beyond the obstacle interfere, causing the particular energy-density distribution, so-called diffraction pattern.

As introduced in Chapter 2, a multi-slit grating that consists of a series of identical parallel and equally spaced slits with a period $a$ (i.e. the grating constant) formed on a plane surface creates the diffraction pattern on the screen by the result of the constructive or destructive interferences of the diffracted incident light. Figure 3-5 schematically shows the multi-slit diffraction grating and its functioning as a dispersive element. The flux intensity distribution as a function of dispersion angle, $\theta$, can be calculated by combining the contributions of all slits (see equation (2.23)). The absolute resultant intensity is proportional to $N^2$. The relative intensity of the diffraction pattern is calculated for $N=4$ and $a=2d$, and shown in Figure 3-5. When $a=2d$ is selected, all even-order maxima ($m=2, 4 \ldots)$ are eliminated in the diffraction pattern as discussed in Chapter 2.

The ratio of the intensity $I$ at a position on the screen to the intensity $I_0$ on the screen for $\theta=0$ is as follows.
**Fig. 3-5.** A diffraction grating which disperses the incident light depending on the wavelength and simulated diffraction pattern in the case of $N=4$ and $a=2d$.

\[
\frac{I}{I_0} = \left( \frac{\sin \beta}{\beta} \right)^2 \left( \frac{\sin N \gamma}{\gamma} \right)^2
\]

with

\[
\beta = n \frac{a \pi}{\lambda} \sin \theta ,
\]

\[
\gamma = n \frac{a \pi}{\lambda} \sin \theta ,
\]

where $\theta$ denotes the angle of diffraction with wavelength $\lambda$, $n$ the index of refraction of the optical-path medium, $a$ the grating constant, and $d$ the slit width. Each slit is considered as a point source, equally spaced with a distance $a$, and all of them generating precisely the same flux intensity distribution. Superimposed, the various distributions yield a multiple-wave
interference system, modulated by the single-slit diffraction envelope as discussed in 2.2.3.2. A larger number of slits results in narrowing of the peak. As can be seen in Figure 3-5, the maximum intensity occurs within the envelope when the following condition is satisfied

\[ n \sin \theta = m \lambda , \]  

(3.6)

where \( m = 0, 1, 2, \ldots \). This is called the grating equation. The integer \( m \) denotes the order of diffraction.

A smaller grating constant results in a higher angle of dispersion, as shown in equation (3.6). However, we cannot make a multi-slit grating with an infinitively small constant, due to the limited feature size in photolithography. Moreover, diffraction occurs only when the grating dimensions are in the order of the wavelength of the incident light, i.e., \( \lambda = a \) [3.12]. This fact indicates the advantage of realising a spectrometer working in the infrared spectral range instead of in the visible range. The longer wavelengths of infrared make it easier to realise an effective grating at a given photolithographic limit. The longer wavelengths result in a larger dispersion angle with respect to a given grating constant and this gives a better resolution for a given optical-path length.

It is not possible to cover a very wide spectral range with one single grating, because the diffraction will only occur if \( \lambda = a \). Thus, several gratings with different grating constants are designed on the same chip to split the detectable wavelength range into a few sub-bands. The grating has 30 or 60 slits with different grating constants from 4 to 20 \( \mu \text{m} \), with the length of the strips equal to 400 \( \mu \text{m} \).

### 3.3.2 Optical property of grating material

A perfect conductor has an infinitive refractive index and no penetration of electromagnetic radiation to any depth, resulting in a perfect reflectivity [3.13]. Conductors such as silver, gold and aluminium in the real world, however, do not have perfect reflectivities, but they show quite high reflectivities over rather broad spectral range, as shown in Figure 3-6. These optical properties can be very useful in radiometric and photometric applications. Especially the high reflectivity of aluminium at visible and
Fig. 3-6. Spectral reflectance of silver, gold and aluminium.

Infrared wavelengths merits the realisation of optical devices in a compatible way with standard IC processing. This is based on the fact that aluminium is the standard material for the metallisation process both in integrated circuits (ICs) and in micro-electro-mechanical systems (MEMS).

Figure 3-7 shows the measured reflection and transmission of an aluminium layer deposited on silicon as a function of the layer thickness. The values measured using a wavelength of 4 μm for a 10 nm layer thickness, for instance, are $R = 85\%$ and $T = 5\%$. A much thicker aluminium layer (600 nm) used in the standard metallisation process of DlMES01 would show a better reflectance.

Consequently, aluminium is the most suitable grating material in terms of fabrication compatibility and simplicity, while it shows high reflectance at the thickness used in the standard metallisation process. The reflectance is even higher than that of silver or gold in the near-infrared spectral region [3.15][3.16].

The boundary condition that the electric field in a thick metallic film is parallel to the surface due to the high mobility of electrons in the film causes metallic light reflection. In a very thin film, however, there are not enough electrons to interact with the electromagnetic wave, resulting in
transparency to light. With a critical thickness between those two cases, the electrons interact with the wave, but remain in the state where they are not free to move. The movement of electrons takes energy from the wave and absorption occurs in such a way. Based on this mechanism, thin metallic films can also be used as IR absorbers [3.17].

3.3.3 Resolving power of multi-slit grating

The angular dispersion of the grating between two spectral components with wavelength difference \( d\lambda \) is derived by differentiating the grating equation with respect to the diffraction angle \( \theta \).

\[
d\theta = \frac{m d\lambda}{acos\theta}
\]  

(3.7)

Selecting a higher order of diffraction results in a larger dispersion angle, as can be seen equation (3.7). However, the intensity decreases rapidly with such an increase of diffraction order. For example, using the third diffraction order would increase the resolving power by a factor of three, but the intensity, in the case of \( N=4 \) and \( a=2d=4 \mu m \), falls to one ninth.
Design of the multi-slit grating-based infrared micro-spectrometer

compared to that of the first diffraction order.

As discussed in 2.2.2.2, two spectral components illuminated on the diffraction grating are resolved if the central maximum of one spectral component is superimposed with first minimum of the other. When the angular separation of two spectral components is less than that, it is impossible to distinguish whether the incident beam is monochromatic or not.

In the case of a double slit, the separation between the peak and the adjacent minimum of the relative intensity is more easily found, as can be seen in Figure 2-9. The angular width between the peak and the adjacent minimum of either spectral line is defined as $d \theta'$ to distinguish these from the angular dispersion of the grating, $d \theta$ shown in equation (3.7). When $\frac{\pi}{\lambda} \sin \theta'$ in the first factor of equation (2.18), $\left[ \cos \left( \frac{\pi}{\lambda} \sin \theta' \right) \right]^2$ changes from 0 to $\pi$, the relative interfered intensity of one of the incident spectral components changes from the maximum, one, to the adjacent minimum, zero. For an arbitrary value of $N$, this can be expressed as

$$d \left( N \frac{\pi}{\lambda} \sin \theta' \right) = N \frac{\pi}{\lambda} \cos \theta' \ d \theta' = \pi.$$  

(3.8)

The angular width $d \theta'$ can thus be obtained:

$$d \theta' = \frac{\lambda}{N \cos \theta'}.$$  

(3.9)

From the Rayleigh criterion, two spectral lines are resolved if the angular separation of the dispersion by a grating, $d \theta$ (equation (3.7)) is equal to or larger than the angular width, $d \theta'$ (equation (3.9)), that is,

$$\frac{\lambda}{d \lambda} = mN.$$  

(3.10)

The resolving power of a grating is defined as the ratio of the wavelength of light illuminating the grating to the narrowest spectral width that can be distinguished (see 2.2.2.2). Therefore, equation (3.10) describes the
resolving power, or

\[ R = \frac{\lambda}{d\lambda} = mN. \]  \hspace{1cm} (3.11)

It is obvious that a large number of slits and a high-order diffraction pattern should be used for a high resolving power since it is fully determined by the number of slits and the diffraction order. However, increasing the number of slits results in a larger grating size. It is then difficult to illuminate the large grating with a uniform intensity.

Because of the rapid decrease in the intensity of the diffraction pattern with the diffraction order, the infrared micro-spectrometer is designed to use the first diffraction order \( (m=1) \). In this case, the number of grating slits directly represents the resolving power.

### 3.4 The silicon optical path

Figure 3-8 shows the wavelength-dependent optical property of silicon that has an indirect bandgap structure. For wavelengths larger than those corresponding to the bandgap energy of silicon (1.12 eV), the free-carrier absorption of light can be disregarded. The upper limit of the wavelength

![Absorption coefficient as a function of wavelength](image)

**Fig. 3-8.** Absorption coefficient of bulk silicon as a function of wavelength.
corresponding to the bandgap energy is thus about 1.1 μm ($\lambda = \frac{hc}{qE_g}$). The absorption increases with the photon energy due to the reduced change in momentum required to generate an electron-hole pair. However, transitions cannot occur even with photon energies larger than 1.12 eV, unless a sufficiently large momentum carried by lattice vibrations or phonons is supplied with the provided photon energy [3.18]. When incident photons with energies higher than 3.4 eV are absorbed, silicon allows a direct transition of electrons from the valence band to the conduction band. This bandgap energy value is equivalent to a wavelength of about 360 nm. At wavelengths shorter than 360 nm, photons are already absorbed at the surface of silicon without extra phonon energy. This wavelength dependence permits red light to penetrate deeply into the silicon.

The above implies that silicon is highly transparent for wavelengths exceeding 1.1 μm [3.19]. Therefore, the bulk silicon can be used to define the optical path for the infrared spectral range. However, the dispersion distance in silicon media for the same optical-path length would be smaller than that in air, due to the larger refractive index of silicon (for $\lambda > 1.1\mu m$, $n = 3.4$, see Figure 3-9).

![Graph showing temperature-dependent refractive index of silicon][1]

**Fig. 3-9.** Temperature-dependent refractive index of silicon [3.20].
As discussed in 3.2.4, the configuration of the simplest lensless infrared spectrometer would involve a grating and an array of infrared detectors at a distance $L$ from the grating (Figure 3-10). Using the grating equation (3.6) and the trigonometric function relation, we can describe the diffraction angles corresponding to the short-end and long-end wavelengths, $\lambda_{\text{min}}$ and $\lambda_{\text{max}}$

\[
\theta_{\text{min}} = \arcsin\left(\frac{m\lambda_{\text{min}}}{na}\right) \tag{3.12}
\]

and

\[
\theta_{\text{max}} = \arcsin\left(\frac{m\lambda_{\text{max}}}{na}\right), \tag{3.13}
\]

respectively.

The size of the projected spectrum on the detector array is then
Design of the multi-slit grating-based infrared micro-spectrometer

\[ P = L(\tan \theta_{\text{max}} - \tan \theta_{\text{min}}) \quad \text{or} \]

\[ \arctan \frac{P}{L} = \arcsin \left( \frac{m \lambda_{\text{max}}}{na} \right) - \arcsin \left( \frac{m \lambda_{\text{min}}}{na} \right). \quad (3.14) \]

If the material of the optical path media, the optical path length, and the grating constant for the specified minimum and maximum wavelengths to be detected are known, the size of projected spectrum \( P \) can be calculated using the equation (3.14). The resolution of the spectrometer defined as \( P/W \) for a detector width \( W \) in the array cannot be very large because the thermopile-based infrared detector, which consists of a few thermocouples, occupies a rather large space. For example, the width of one infrared detector with 12 thermocouples in the array is 100 \( \mu \)m due to the limits set by the photolithography. Evidently, the spectrometer must be designed such that the resolving power of the grating, which is directly proportional to the number of slits in the grating (equation (3.11)), should be large enough to support this spectrometer resolution.

The equation (3.14) can be normalized as

\[ \arctan \frac{x}{L} = \arcsin \frac{m \lambda}{na}, \quad (3.15) \]

where \( x \) is the diffracted position of the spectrum projected on the detector array. It gives the dispersed distance as a function of wavelength for different optical path media, plotted in Figure 3-11 for a 4 \( \mu \)m grating constant and for the first diffraction order. Based on this calculation, we can design the detector array for different parameters such as the grating constant, optical path length, composition of optical path media and so on. Only the first diffraction order of the dispersed IR light after transmission through the combined optical path media, bulk silicon and air is detected in the thermopile-based IR detector array fabricated on a silicon nitride membrane at the bottom wafer, as discussed in the next section.
3.5 The thermopile-based infrared detector array

The infrared detector is the last part to be considered in the design of the infrared micro-spectrometer. An array consisting of a number of identical infrared detectors is formed to detect the different spectral components ending at different locations, as a result of the wavelength-dependent diffraction angle of the incident infrared light. Each detector in the array measures the intensity distribution of the diffracted spectral components from the incident light illuminating the multi-slit grating.

3.5.1 Construction

The infrared detectors are prepared using p- and n-type polysilicon as the thermoelectric material on a thin membrane. Although Bi$_2$Te$_3$ is reported as one of the best thermoelectric materials [3.21], polySi is preferred, due to the anticipated fabrication difficulties and compatibility problems of the alternatives. Silicon nitride was selected as the membrane material because of its high mechanical strength and its ability to resist the etchant used to
remove the bulk silicon. Hot junctions of the thermopile are located at the centre of the membrane and cold junctions are at the rim area of the bulk silicon, which acts as heat sink. Each detector in the infrared detector array is thermally isolated to minimize cross talk between neighbouring detectors by cutting the interval membrane areas using maskless Reactive Ion Etching (RIE) after releasing the membrane. For the maskless RIE sequence, the cut area on the membrane is thinned in advance of metallisation.

The location of the first detector in the detector array is shifted with respect to the centre of the grating in order to avoid the interference of the 0th diffraction order. The entire area on top of the grating wafer except the opened slit regions is shielded with an aluminium layer to prevent unnecessary illumination with infrared light via the grating wafer to the detector array.

The width of one single IR detector and the spacing between the detectors are 100 μm and 4 μm, respectively, resulting in a pitch of 104 μm. Each infrared detector has 12 thermocouples connected electrically in series in the array. The number of thermocouples in the 100 μm-sized IR detector is limited by the photolithography. The number of detectors in one array ranges between 6 and 16 depending on the dispersion distance for a given grating constant, which gives the detectable wavelength range.

3.5.2 Performance - Optimisations of IR detector

The major factor limiting the performance of thermal-type IR detectors is heat loss due to conduction, convection and radiation. In the next sections, the thermal characteristics of the designed IR detector are discussed and ways to optimise the device performance.

3.5.2.1 Thermal conductance

The detector parameters, like for instance the detector’s sensitivity, detectivity, thermal time constant and electrical/thermal resistance, can be optimised by means of a one-dimensional temperature distribution analysis [3.22]. Actually, the heat transfer elements of a thermoelectric sensor are mostly composed of a very thin film, in which the thickness effect can be ignored and heat transfer in the elements occurs only in one direction.
The thermopile-based infrared detector array

Three main mechanisms of heat transfer are conduction, convection and radiation. The total thermal conductance $G$ is conveniently divided into three independent parts; thermal conductance due to radiation $G_{rad}$, convection $G_{gas}$ at the absorbing area and the thermal conductance through the beam $G_{beam}$ [3.23], that is,

$$G = G_{rad} + G_{gas} + G_{beam} = G_l + G_{beam}[W/K] .$$  \hspace{1cm} (3.16)

Figure 3-12 shows the structure of a single pixel of the designed infrared detector array. In order to reduce parasitic thermal conduction through the membrane, we applied RIE etching after KOH etching to remove parts of the Si$_3$N$_4$ membrane to isolate each pixel. It is assumed that the absorbing area is isothermal, that is, that the temperature increase by infrared irradiation on the whole absorbing area is uniform. Furthermore, the temperature increase by infrared absorption on the supporting beam, where the thermopiles are located, is considered negligible compared to that of the absorbing area. Because of the symmetrical structure and temperature uniformity on the absorbing area, only half of the structure needs to be modelled, as can be seen in Figure 3-12.

Fig. 3-12. Heat transfer on bridge and cantilever structures of thermopile-based infrared detector.
Design of the multi-slit grating-based infrared micro-spectrometer

The radiant flux absorbed on the detector $E$ must be equal to the sum of the thermal losses by radiation at the absorbing area $E_i$ and the power $E_{beam}$ transferred into the beam so that

$$E = E_i + E_{beam} . \quad (3.17)$$

If we define $T_a$ as the ambient temperature of the heat sink and $T_i$ as the temperature of the absorbing area, the thermal loss at the absorbing area $E_i$ can be expressed as

$$E_i = G_i (T_i - T_a) \quad (3.18)$$

and the heat transfer energy into the beam can be written by

$$E_{beam} = G_{beam} (T_i - T_a) , \quad (3.19)$$

where $G_i$ and $G_{beam}$ are the thermal conductance of the absorbing area and the supporting beam, respectively. So the equation (3.17) is the following

$$E = E_i + E_{beam} = (G_i + G_{beam})(T_i + T_a) = G(T_i + T_a) . \quad (3.20)$$

**Thermal conductance on the absorbing area $G_i$**

The thermal conductance on the absorbing area $G_i$ consists of two components, thermal conductance due to the radiation $G_{rad}$ and thermal conductance due to the surrounding gas $G_{gas}$.

**Due to radiation**

The power radiating from a body depends on the fourth power of its absolute temperature. The total energy emitted by the unit area of a blackbody surface (emissivity=1) depends only on the temperature. This is known as the Stefan-Boltzmann law for radiation. Therefore, the total radiant emittance from the absorbing area $E_i$ can be written as

$$E_i = \varepsilon \sigma T_i^4 [W] , \quad (3.21)$$
where \( \varepsilon \), \( \sigma \) and \( A \) are the absorptance of the absorbing area, Stefan-Boltzmann constant \( (=5.67 \times 10^{-8} \, W/ \, m^2 \, K^4) \) and the absorbing area \( (= W(L_i - L)) \), respectively. Assuming that the absorbing area has an isothermal collector layer and therefore shows a uniform temperature increase by infrared irradiation gives

\[
G_{rad} = \frac{dE_i}{dT} = 4 \varepsilon \sigma AT_i^3 .
\]  \( (3.22) \)

**Due to convection**

For a quiet gas environment in a small package, the thermal conductance due to the surrounding gas at the absorbing area is \([3.23][3.24]\)

\[
G_{gas} = \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) A
\]  \( (3.23) \)

where \( \lambda_{gas} \) is the thermal conductivity of the gas atmosphere, and \( d_1 \) and \( d_2 \) are the distances between the beam and the substrate, and the upper heat sink, respectively. The upper heat sink corresponds to the grating wafer. The total thermal conductance at absorbing area \( G_i \) is thus

\[
G_i = G_{rad} + G_{gas} = \left[ 4 \varepsilon \sigma T_i^3 + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) \right] A .
\]  \( (3.24) \)

**Thermal conductance through the supporting beam** \( G_{beam} \)

The thermal conductance between the absorbing area and the heat sink can be determined by calculating the temperature distribution \( \Delta T(x) = T(x) - T_a \) at any point \( 0 \leq x \leq L \) along the beam \([3.25][3.26]\). The gradient of temperature is equal to the heat flux at that point divided by the thermal conductivity of the beam, that is,

\[
d\Delta T(x) = d[T(x) - T_a] = \frac{E_{beam}(x)}{tW} dx \lambda ,
\]  \( (3.25) \)

where \( t \) is the thickness of the beam and \( \lambda \) is the thermal conductivity of the
Design of the multi-slit grating-based infrared micro-spectrometer

beam material. And the change in heat flux at any \( x \) is given by the local temperature change with respect to the \( \Delta T(x) \) times the thermal conductance by radiation (Eq. (3.22)) and convection (Eq. (3.23))

\[
dE_{beam}(x) = (T(x) - T_a) \left[ 4\varepsilon \sigma T_i^3 + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) \right] W dx .
\] (3.26)

Eq. (3.25) and Eq. (3.26) result in the following differential equation with the condition of \( \Delta T / T_a \ll 1 \), that is, \( T = T_a \)

\[
d^2 [T(x) - T_a] \frac{dx^2}{dx^2} = \frac{4\varepsilon \sigma W[T^4(x) - T_a^4] + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) W[T(x) - T_a]}{\lambda_{lw}}
\] (3.27)

If the beam is composed of several layers to form a thermopile, the above equation can be expressed by

\[
d^2 [T(x) - T_a] \frac{dx^2}{dx^2} = \frac{4\varepsilon \sigma W[T^4(x) - T_a^4] + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) LW[T(x) - T_a]}{\sum_s \lambda_s f_s LW}
\] (3.28)

where \( s \) represents the different films ( \( s = 1, 2 \ldots \)).

The right-hand term of equation (3.28) represents the heat loss power normalised to the total volume of the beam. This equation is simplified again as follows, with \( (T - T_a) / T_a \ll 1 \),

\[
d^2 [T(x) - T_a] \frac{dx^2}{dx^2} = \frac{4\varepsilon \sigma T_i^3 + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) \sum_s \lambda_s f_s}{[T(x) - T_a]} = 0 .
\] (3.29)

Infrared Micro-spectrometer based on a Multi-slit Grating
The temperature gradient along the supporting beam follows from Eq. (3.17) (3.18) (3.24) and (3.25) as:

\[
W \sum_s \lambda_s t_s \frac{dT(x) - T_a}{dx} \bigg|_{0 \leq x \leq L} = E_{beam} = E - E_i = E - G_i (T_i - T_a)
\]

\[
= E - \left[ 4 \varepsilon \sigma T_i^3 + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) \right] A (T_i - T_a)
\]

The solution of the heat transport equation (3.29) with the boundary condition \( \Delta T(x = 0) = 0 \) is

\[
T(x) - T_a = \frac{\sinh [\delta x] E_{beam}}{W \sum_s \lambda_s t_s \delta \cosh [\delta x]} = \frac{E_{beam}}{W \sum_s \lambda_s t_s \delta \coth [\delta x]},
\]

where \( \delta \) refers to the heat transfer coefficient

\[
\delta = \left( \frac{4 \varepsilon \sigma T_a^3 + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right)}{\sum_s \lambda_s t_s} \right)^{1/2}.
\]

\[
\therefore [T(x) - T_a]_{x = L} = T_i - T_a = \frac{E_{beam}}{W \sum_s \lambda_s t_s \delta \coth [\delta L]}
\]

From equation (3.19) and (3.33), the thermal conductance on the supporting beam \( G_{beam} \) is

\[
G_{beam} = W \sum_s \lambda_s t_s \delta \coth [\delta L].
\]
Consequently, the total thermal conductance $G$ is derived

$$G = \left[ 4\varepsilon \sigma T_i^3 + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) \right] A + W \sum_s \lambda_s t_s \delta \coth(\delta L) . \tag{3.35}$$

### 3.5.2.2 $\Delta T$ as a function of beam length

The temperature difference between the hot and cold junctions of the thermopile, which are located on the absorbing area and the bulk substrate heat sink, respectively, can be expressed as $T_i - T_a$. It is assumed that the cold junction has ambient temperature. This can be written again from equation (3.30), (3.33) and (3.35) as

$$T_i - T_a = \frac{E - \left[ 4\varepsilon \sigma T_i^3 + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) \right] A (T_i - T_a)}{W \sum_s \lambda_s t_s \delta \coth(\delta L)} \tag{3.36}$$

or

$$T_i - T_a = \frac{E}{\left[ 4\varepsilon \sigma T_i^3 + \lambda_{gas} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) \right] A + W \sum_s \lambda_s t_s \delta \coth(\delta L)} = \frac{E}{G} .$$

If the geometries of the IR detector components and several properties used as thermocouples are known, the achievable temperature difference between hot and cold junctions as a function of beam length can be plotted using equation (3.36).

Figure 3-13 shows the calculated temperature difference per absorbed unit power on the absorbing area as functions of beam length and recessed etching depth on the back side of the grating wafer, $d_2$. The latter parameter, $d_2$, is in fact the distance between the silicon nitride beam supporting the infrared detector and the upper heat sink. Table 3-1 shows the values used for this calculation.
The thermopile-based infrared detector array

**Fig. 3-13.** Calculated temperature difference between hot and cold junctions as a function of beam length.

**TABLE 3-1.** Geometric parameters and constants used for the calculation of temperature difference between the absorbing area and heat sink resulted from an infrared irradiance.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absorbing area, A</td>
<td>100 x 100 ([\mu m^2])</td>
</tr>
<tr>
<td>Beam width, W</td>
<td>100 [\mu m]</td>
</tr>
<tr>
<td>Absorptance of SiN, (\varepsilon_{SiN})</td>
<td>0.1</td>
</tr>
<tr>
<td>Emissivity of polySi, (\varepsilon_{Si})</td>
<td>0.19</td>
</tr>
<tr>
<td>Thermal conductivity of surrounding gas (still), (\lambda_{gas})</td>
<td>0.03 [W/m K]</td>
</tr>
<tr>
<td>Thermal conductivity of polySi, (\lambda_{Si})</td>
<td>29 [W/m K]</td>
</tr>
<tr>
<td>Thermal conductivity of SiN, (\lambda_{SiN})</td>
<td>16 [W/m K]</td>
</tr>
<tr>
<td>Thickness of SiN and polySi, (t_1) and (t_2)</td>
<td>0.3 [\mu m]</td>
</tr>
<tr>
<td>Stefan-Boltzmann constant, (\sigma)</td>
<td>5.67 x 10^{-8} [W/m²K⁴]</td>
</tr>
<tr>
<td>The distances between the beam and substrate, (d_1)</td>
<td>525 [\mu m]</td>
</tr>
<tr>
<td>The distances between the beam and upper cap, (d_2)</td>
<td>100, 300, 500 [\mu m]</td>
</tr>
</tbody>
</table>

Infrared Micro-spectrometer based on a Multi-slit Grating 87
Due to the symmetrical structure of the designed infrared detector and the assumed temperature uniformity over the absorbing area (see Figure 3-12), half of the absorbing area ($50 \times 100 \mu m^2$) is used in the calculation. As can be seen in Figure 3-13, the temperature difference remains almost constant for a relatively long beam, where the heat loss by radiation is dominant. Based on this result, the beam length should be chosen in a range where heat loss is dominated by radiation rather than by conduction. In achieving a radiation dominant mode the parameter $\delta$ is very important, as can be seen in equation (3.32) and (3.36). Only by using a thin membrane which has low thermal conductivity, one can obtain a relatively large value of $\delta$ that results in a radiation dominant mode with a feasible beam length. As can be seen in Figure 3-14, using a Al/polySi thermocouple that has a much higher thermal conductivity as compared to that of a polySi/polySi thermocouple yields a much longer beam length in which radiation related losses are dominant. Very long beams are impractical, as the thin membrane becomes fragile.

![Diagram](image)

**Fig. 3-14.** Difference in $\Delta T$ caused by a different thermal conductivity of used thermoelectric materials.
3.5.2.3 Optimization of beam length

The sensitivity of the thermopile is defined as the ratio of the thermoelectric signal voltage $U$ to infrared radiant flux absorbed on the absorbing area $E$, and can be obtained from equation (3.36) as

$$ S = \frac{U}{E} = \frac{N \alpha (T_i - T_a)}{E} = \frac{N \alpha}{G} $$  \hspace{1cm} (3.37)

$$ = \frac{N \alpha}{G + W \sum \lambda_n \delta \coth[\delta L]} $$

where $N$ and $\alpha$ are the number of thermocouples and the Seebeck coefficient of one junction pair, respectively.

However, the sensitivity gives no indication of the minimum radiant power that can be detected. Detectivity is the reciprocal of noise equivalent power (NEP), which is defined as the rms incident radiant power required to produce an output signal equal to the detector noise level.

$$ D = \frac{1}{\text{NEP}} = \frac{\text{sensitivity}}{\text{noise}} \left[ \frac{\text{W}^{-1}}{} \right] $$  \hspace{1cm} (3.38)

When the detectivity is used to characterize a detector, it is necessary to specify the wavelength of the incident radiation, the detector temperature, the chopping frequency, any bias current applied to the device, the area of the detector and the bandwidth of the amplifier used to measure the detector noise. The detectivity is not an ideal parameter for comparing different detectors, as it varies inversely as the square root of both the sensitive area and the bandwidth. The specific detectivity, expressed as $D^*$, is a more useful parameter for comparing detectors of different sizes, as the device noise is generally proportional to the square root of the absorbing area. It is defined as

$$ D^* = D \sqrt{\Delta f} = \frac{S \sqrt{\Delta f}}{\text{noise}} \left[ \frac{\text{cm} \sqrt{\text{Hz}} \text{W}^{-1}}{} \right], $$  \hspace{1cm} (3.39)

where $S$ is the sensitivity of the detector and $\Delta f$ is bandwidth of the
amplifier used to measure the detector signal.

As can be seen in equation (3.34) and (3.37), the thermal conductance and sensitivity will reach their minimum and maximum values at a large length of the beam, respectively.

\[
G_{\text{min}} = G_i + W \sum_s \lambda_s t_s \delta \coth(\delta L) \bigg|_{L \to \infty} = G_i + W \sum_s \lambda_s t_s \delta \\
S_{\text{max}} = \frac{N \alpha}{G_i + W \sum_s \lambda_s t_s \delta \coth(\delta L)} \bigg|_{L \to \infty} = \frac{N \alpha}{G_i + W \sum_s \lambda_s t_s \delta}
\]  

(3.40)

If the absorbing area, the beam material and its heat transfer coefficient are known, \(G_i\) and \(G_{\text{beam min}}\) can be determined for a given sensor. From these, the ratio \(S/S_{\text{max}}\) can be expressed as

\[
\frac{S}{S_{\text{max}}} = \frac{G_i + W \sum_s \lambda_s t_s \delta}{G_i + W \sum_s \lambda_s t_s \delta \coth(\delta L)} = \frac{m + 1}{m + \coth(\delta L)}
\]

(3.41)

where \(m = G_i / G_{\text{beam min}}\). Using this equation, we can plot the sensitivity normalized to the maximum sensitivity, \(S/S_{\text{max}}\) as a function of beam length, \(L\). Figure 3-16 shows \(S/S_{\text{max}}\) versus beam length using the parameters shown in Table 3-1.

In a similar way, the detectivity and time constant can also be described as a function of beam length. These can give a good guide to decide the beam length resulting in optimised performance, but it is necessary to make a proper trade-off among sensitivity, detectivity and thermal time constant.

When only thermal noise is generated by the thermopile with an electrical resistance \(R\), the specific detectivity \(D^*\) is given by equation (3.39)
Fig. 3-15. \( S/S_{\text{max}} \) as a function of beam length.

\[
D^* = \frac{S\sqrt{A}}{\sqrt{4kTR}}, \tag{3.42}
\]

where the electrical resistance \( R \) of the thermopile depends on the beam length \( L \), so it can be written as

\[
R = N\left(\frac{L\rho_n}{A_n} + \frac{L\rho_p}{A_p}\right). \tag{3.43}
\]

\( N \): the number of thermocouples
\( \rho_n, \rho_p \): specific electrical resistivities of the thermopile
\( A_n, A_p \): cross-sectional area of the thermocouple

Consequently, there would be a maximum detectivity value at a certain beam length due to the fact that the specific detectivity, \( D^* \), is proportional to \( S/\sqrt{L} \) and the sensitivity, \( S \), saturates with the increase in beam length. To get the maximum detectivity value we determine the first derivative by the beam length:

\[
\frac{dD^*}{dL} = \frac{d}{dL} \left( \frac{S\sqrt{A}}{\sqrt{4kTM\left(\frac{L\rho_n}{A_n} + \frac{L\rho_p}{A_p}\right)}} \right) = 0 \tag{3.44}
\]
resulting in

\[
\delta L_{\text{max}} = \frac{m + \coth[\delta L_{\text{max}}]}{2 \frac{d}{dL} \coth[\delta L_{\text{max}}]} = \frac{m + \coth[\delta L_{\text{max}}]}{2(1 - \coth^2[\delta L_{\text{max}}])},
\]

(3.45)

where \(L_{\text{max}}\) means the beam length associated with maximum detectivity. The ratio \(D^*/D^*_{\text{max}}\) can be expressed as:

\[
D^*/D^*_{\text{max}} = \frac{S/\sqrt{L}}{S_{\text{max}}/\sqrt{L_{\text{max}}}} = \frac{S}{S_{\text{max}}} \frac{\sqrt{L_{\text{max}}}}{\sqrt{L}}
\]

(3.46)

\[
G_i + W \sum_s \lambda_s t_s \delta \coth[\delta L_{\text{max}}] \sqrt{L_{\text{max}}} = G_i + W \sum_s \lambda_s t_s \delta \coth[\delta L] \sqrt{L}
\]

\[
= \frac{m + \coth[\delta L_{\text{max}}]}{m + \coth[\delta L]} \frac{\sqrt{\delta L_{\text{max}}}}{\sqrt{\delta L}}
\]

Using the parameters from Table 3-1 (page 87) results in a ratio \(m = G_i / G_{\text{beam}_{\text{min}}} = 6.6347 \times 10^{-4}\). The normalized beam length associated with the maximum detectivity, \(\delta L_{\text{max}}\), is 1.09. Based on these values, we plot the ratio \(D^*/D^*_{\text{max}}\) as a function of the beam length, as shown in Figure 3-16. The beam length corresponding to the maximum detectivity \(L_{\text{max}}\) is about 320 \(\mu\)m. The proper beam length can be found more easily by means of Figure 3-16.

Finally, the thermal time constant \(\tau\) of the infrared detector can be expressed as: \(\tau = \frac{C}{G} = \frac{C_i + C_{\text{beam}}}{G}\), where \(C_i\) and \(C_{\text{beam}}\) are the heat capacitances of the absorbing area and supporting beam, respectively. The thermal time constant increases with the beam length because the heat
capacity is proportional to the length. The ratio \( S/\tau \) can be used for maximizing the sensitivity and for minimizing the thermal time constant. In the case of thermopile-based infrared sensor, \( S/\tau \) can be written from equation (3.37)

\[
S/\tau = \frac{N\alpha}{C_i + C_{\text{beam}}},
\]

(3.47)

where the heat capacity is expressed as \( C = V\rho c = aL\rho c \) (\( V \): volume, \( \rho \): density, \( c \): specific heat, \( a \): cross-sectional area, \( L \): length). This ratio would have a maximum value \( (S/\tau)_{\text{max}} \) as the limiting case of zero beam length, that is,

\[
(S/\tau)_{\text{max}} = \left. \frac{N\alpha}{C_i + C_{\text{beam}}} \right|_{L \to 0} = \frac{N\alpha}{C_i}.
\]

(3.48)

The ratio \( \frac{S/\tau}{(S/\tau)_{\text{max}}} \) is plotted as a function of the beam length (see Figure 3-17) using the parameters shown in Table 3-1 and 3-2.

\[
\frac{S/\tau}{(S/\tau)_{\text{max}}} = \frac{C_i}{C_i + C_{\text{beam}}} = \frac{1}{1 + \frac{a\rho c L}{C_i}}
\]

(3.49)
Design of the multi-slit grating-based infrared micro-spectrometer

Fig. 3-17. The ratio $\frac{S/\tau}{(S/\tau)_{max}}$ as a function of beam length.

<table>
<thead>
<tr>
<th>TABLE 3-2. Density and specific heat of the membrane and thermopile.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membrane material</td>
</tr>
<tr>
<td>(Si$_3$N$_4$)</td>
</tr>
<tr>
<td>Density [g/cm$^3$]</td>
</tr>
<tr>
<td>Specific heat [J/g K]</td>
</tr>
</tbody>
</table>

The term $\alpha p c / C_p$ is a constant if the materials are known. When the beam length is, for instance, 320 $\mu$m that associated with the maximum detectivity in Figure 3-16, the ratio $\frac{S/\tau}{(S/\tau)_{max}}$ is about 0.115, as can be seen in Figure 3-17.

A proper beam length could be chosen for the best performance during the design of thermopile-based IR detector in such a way as that discussed above: by taking the detectivity, sensitivity and thermal time constant of the device into account.
3.5.2.4 Integrated Peltier cooler

As a method to enhance the performance of thermopile-based IR detector, we have considered cooling of the IR detectors using an integrated Peltier cooler, which is shown below.

A major portion of the studies on thermoelectric devices focuses on thermopile-based sensors using the Seebeck effect. As a result, various sensors have been developed [3.27][3.28]. More recently, integrated thermoelectric coolers operating based on the Peltier effect and thermoelectric generators [3.29] have started to attract attention. In fact, an integrated Peltier element can be applied to thermally control the cold-junction temperature of thermopile-based IR detectors due to the fact that it has the same structure and is composed of the same material as the Seebeck element so that both elements can be fabricated simultaneously.

The maximum temperature difference between the junction at ambient temperature and the thermally controlled junction, $\Delta T_{\text{max}}$, is the most important parameter determining the performance of a Peltier device. Figure 3-18 shows the plot of such an analysis. Taking into account the heat loss by conduction through the thermoelectric material as well as through the membrane and the parasitic effect due to the contact resistance at the metal-semiconductor interface we can express $\Delta T_{\text{max}}$ as [3.30][3.31]

$$\Delta T_{\text{max}} = (T_a - T_c)_{\text{max}} = \frac{zT_c^2}{2\left(1 + \frac{K'}{K} + \frac{2R_{eo}}{R} + \frac{2K' R_{eo}}{KR}\right)}$$

(3.50)

where

$$K = \frac{\lambda_n A_n}{L_n} + \frac{\lambda_p A_p}{L_p}$$

$$K' = \frac{\lambda_m A_m}{L_m} + \frac{4\varepsilon \sigma T_a^3 A}{4} (\text{vacuum})$$

$$R = \frac{L_n \rho_n}{A_n} + \frac{L_p \rho_p}{A_p}$$

$R_{eo}$ is the contact resistance and the subscripts $a$, $c$, $m$, $n$ and $p$ express the ambient, cold junction, membrane, $n$-type and $p$-type properties. We
analysed the maximum achievable temperature difference for both polySi and polySiGe as the Peltier elements; the results are shown in Figure 3-18. The geometries were taken from a constructed test device that has a cooled

![Graph showing temperature difference and contact resistance for polySi and polySiGe Peltier elements](image)

**Fig. 3-18.** Maximum temperature difference between the hot and cold junctions of (a) polySi and (b) polySiGe Peltier elements, as a function of the element thickness (solid line), and as a function of the contact resistance (dashed line).
area with a suspended epi mass \((A=60\times500 \, \mu m^2)\), Peltier elements with a length \(L\) of 75 \(\mu m\), a \(Si_3N_4\) membrane with a length and width of 100 and 450 \(\mu m\) and a thickness of 0.3 \(\mu m\). Half of the membrane is covered with the Peltier elements. The material parameters were taken from literature [3.30][3.31]. For calculating the temperature difference as a function of the thickness of thermoelectric element, we took the contact resistance, \(5 \, \Omega\) from the data measured in [3.32]. The thicknesses of both types of Peltier elements, \(d_n\) and \(d_p\) which are used to discover the influence of the contact resistance on the maximum temperature difference, is 0.3 \(\mu m\), as it is in the constructed test device. The maximum temperature achieved with a 0.3 \(\mu m\)-thick polySiGe Peltier element is 3.26 K, while it is only 1.34 K with an

\[\Delta T [K]\]

\(\lambda_m [W/Km]\)

\(I [A]\)

\(\Delta T [K]\)

\(R_c [\Omega]\)

\(I [A]\)

**Fig. 3-19.** Comparison of the influence on the maximum temperature difference \(\Delta T\), with respect to the current \(I\) applied, by (a) thermal conduction through the supporting membrane and (b) contact resistance [3.31].
equally thick polySi, as can be seen in Figure 3-18. This difference in performance is primarily due to the much higher thermal conductivity of polySiGe, resulting in a higher figure of merit [3.32].

The influence of thermal conduction has been studied thoroughly [3.33] and the radiation loss can be reduced by downscaling. However, the effect of the contact resistance on the performance of the Peltier device is much more severe when the Peltier device is downscaled, since this reduces the contact area, resulting in an increasing in the contact resistance. The contact resistance does not only degrade device performance, but also causes a change in the optimum device current as well, as can be seen in Figure 3-19 [3.31].

A thermopile-based IR detector with an integrated Peltier cooler

Thermally stabilised silicon micro-platforms have been presented for various applications [3.34][3.35][3.36]. For those devices that work at temperatures far above the ambient temperature, active heating using an integrated resistor and passive cooling by heat conduction to the ambient is sufficient for a good time response. However, in applications where the operating temperature is near ambient temperature, active cooling is necessary as passive cooling is too slow and it is limited to temperatures above the ambient temperature [3.37].

Figure 3-20 shows a schematic representation of a fabricated thermopile IR detector with an integrated Peltier element, together with its corresponding (simplified) lumped-element model. The Peltier element is used to transport heat between the ambient and the side of the thermopile which is the farthest away from the IR absorber, referred to as the cold junction, which is at temperature $T_c$. In the simulation, both the Peltier and Seebeck elements are assumed to be made of polySiGe, connected by Al. These elements are situated on top of a $\text{Si}_3\text{N}_4$ membrane, which is covered by an absorbing area of black gold at the tip.

The capacitors $C_1$ and $C_2$ correspond to the thermal capacity of the cooled thermal mass (at $T_c$) and the thermal capacity of the absorbing area (at $T_i$), respectively, including the supporting membranes. The substrate temperature is assumed constant and identical to the ambient temperature.
The thermopile-based infrared detector array

Fig. 3-20. Schematic view of the Peltier-cooled IR detector and its corresponding lumped-element model.

\( T_a \) Resistors \( R_1 \) and \( R_2 \) correspond to the thermal resistance of the Peltier elements and Seebeck elements, respectively. Current sources \( q_p \) and \( q_i \) indicate heat transportation by the Peltier elements and heat generation by incident radiation on the absorbing area, respectively. Based on the principle of superposition, \( q_i \) does not directly influence the behaviour of the Peltier element, and is therefore ignored. The temperature distribution due to Peltier cooling can be determined using this model and is expressed as

\[
T_c = T_a - \frac{R_1}{Z_c + R_1} T_a + q_i Z_c ,
\]

where

\[
Z_c = \frac{(j \omega C_1)^{-1} \cdot (R_2 + (j \omega C_2)^{-1})}{(j \omega C_1)^{-1} + (R_2 + (j \omega C_2)^{-1})} .
\]

The geometric parameters were taken from the constructed test device (see Figure 3-20), as in the calculation of the achievable maximum temperature difference with the integrated Peltier device. The test device consists of a 4.5 \( \mu \)m thick black-gold absorbing area (450\times450 \( \mu \)m\(^2\)), a 0.3 \( \mu \)m thick and 450 \( \mu \)m wide Si\(_3\)N\(_4\) supporting membrane 50\% of which is covered with
700 μm long Seebeck and 100 μm long Peltier elements, and the suspended epi mass underneath the membrane is 60×500×4 μm³. Using material parameters from the literature [3.30][3.32], we obtained the resulting values for $C_1$, $C_2$, $R_1$ and $R_2$ of 11.33 nF, 1.64 μF, 31.89 kΩ and 223 kΩ respectively. Figure 3-21 shows photographs of the fabricated test device and its schematic fabrication sequence.

![Diagram of IR detector with Seebeck and Peltier elements](image)

**(a)**: Silicon nitride membrane and Seebeck elements. **(b)**: Etched side wall. **(c)**: Schematic fabrication sequence.

**Fig. 3-21.** An IR detector with a 4 μm thick suspended epi layer and integrated Peltier cooler (a) front side photograph (b) back side view (c) schematic fabrication sequence.
In the analysis, both the step response and frequency response are of interest. In the case of steady-state analysis, \( q_p \) is non-varying, so \( C_1 \) and \( C_2 \) can be ignored. In this case, the temperature difference \( (T_c - T_{e}) \) will not change, while the temperature difference \( (T_{e} - T_{a}) \) and thus \( (T_{e} - T_{o}) \) can be influenced directly by the Peltier effect.

The next step is the calculation of the step response of the device. Figure 3-22 (a) shows the result for \( T_c \) and \( T_i \) when a 100 \( \mu \)W Peltier device is turned on at \( t = 0 \), modelled as a step in \( q_p \) at \( t = 0 \). Both values will approach the steady-state value, but the time constants are significantly different, due to the large difference in time constants \( t_1 = R_1C_1 \) and \( t_2 = R_2C_2 \). Examining the frequency response of the system, Figure 3-22 (b) shows the result at an

![Fig. 3-22. SPICE simulation showing the transient response of \( T_c \) and \( T_i \) (a) with a step load and (b) with 50Hz AC driving, both using a single SiGe Peltier element with a cooling power of 100 \( \mu \)W.](image)
actuation frequency of 50Hz, with a 100 μW peak cooling power. Clearly, the temperature at the absorbing area \((T_i)\) is unable to follow the temperature change of the cold junction \((T_c)\).

A thermal mass (constructed out of the 4 μm epi layer underneath the membrane) was incorporated in the model of Figure 3-20, as a design option to choose the value of \(C_1\). If one leaves out this thermal mass, which reduces \(C_1\) to a minimum and at the same time increases the actuation frequency, the ratio \(t_1/t_2\) will be decreased to a minimum. In practice this means that for properly chosen frequencies, \(T_i\) will not be able to follow changes in \(q_p\), while \(T_c\) can. As an IR detector is sensitive to variations in the ambient or substrate temperature, co-integrating the IR detector with the integrated Peltier cooler could reduce its susceptibility to thermal noise and fluctuation by dynamic control of the device temperature. However, the implementation of dynamic background noise reduction techniques must be investigated further.

Background fluctuation noise is the temperature fluctuation noise occurring when radiative exchange is the dominant mode of heat exchange between the thermal detector and its surroundings. When the detector and the background are at the same temperature, the background fluctuation noise limited detectivity is given by [3.26]

\[
D^*_{BF} = \sqrt{\frac{\varepsilon}{8k\sigma(T_D^5 + T_B^5)}},
\]

(3.53)

where \(k\), \(T_D\) and \(T_B\) are Boltzmann's constant \((1.38 \times 10^{-23} [J/K])\) and the temperatures of the detector and the background, respectively. Figure 3-23 shows the background fluctuation noise limited detectivity as a function of \(T_B\) for different \(T_D\) assuming the emissivity of one. Cooling the detector, even to liquid-helium temperatures, will not dramatically increase the detectivity \(D^*\) when it is used to view 300K background, as can be seen in Figure 3-23. On the other hand, truly dramatic \(D^*\) increases are observed when both the detector and the background are at very low temperatures. Consequently, it is important to realise that the detectivity of an IR detector would not be improved significantly by thermoelectric cooling, unless the background temperature varies in accordance with the cold-junction.
temperature.

![Graph showing background fluctuation noise limited detectivity as a function of background temperature, $T_B$, for different detector temperatures, $T_D$.](image)

**Fig. 3-23.** Background fluctuation noise limited detectivity as a function of background temperature, $T_B$, for different detector temperature, $T_D$.

### 3.5.3 Choice of thermoelectric material

In general, telluride compounds have the highest figure of merit $z$, which is defined as $\alpha^2 / (\rho \cdot \kappa)$, where $\alpha$, $\rho$, and $\kappa$ represent the Seebeck coefficient, the electrical resistivity, and the thermal conductivity, respectively [3.38]. However, the incompatibility of telluride compounds with standard IC fabrication technology severely limits its applications if the fabrication compatibility must be maintained.

Two suitable thermoelectric materials for IC compatible processing are polycrystalline silicon germanium (polySiGe) and polycrystalline silicon (polySi). Table 3-3 shows the thermoelectric properties of the different materials. PolySiGe gives a better figure of merit, as can be seen in Table 3-3, mainly due to its lower thermal conductivity. However, this material requires a more complicated deposition sequence, such as epitaxy, and a more difficult etching process.

PolySi is selected as the thermoelectric material for the thermopile-based IR
**TABLE 3-3. Thermoelectric properties of Bi$_2$Te$_3$, polySiGe and polySi [3.31].**

<table>
<thead>
<tr>
<th>Material</th>
<th>Seebeck Coeff. ($\alpha$[\mu V/K])</th>
<th>Resistivity $\rho$[\Omega m]</th>
<th>Conductivity $\kappa$[Wm$^{-1}$K$^{-1}$]</th>
<th>Figure of merit $Z$[10$^{-3}$K$^{-1}$]</th>
<th>Dop. conc. [10$^{20}$ cm$^{-3}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bi$_2$Te$_3$</td>
<td>n -240</td>
<td>10</td>
<td>2.02</td>
<td>2.89</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>p 162</td>
<td>5.5</td>
<td>2.06</td>
<td>2.32</td>
<td>2.55</td>
</tr>
<tr>
<td>PolySiGe</td>
<td>n -136</td>
<td>10.1</td>
<td>4.45</td>
<td>0.328</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>p 144</td>
<td>13.2</td>
<td>4.80</td>
<td>0.413</td>
<td>2.4</td>
</tr>
<tr>
<td>PolySi</td>
<td>n -120</td>
<td>8.5</td>
<td>24</td>
<td>0.071</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>p 190</td>
<td>58</td>
<td>17</td>
<td>0.037</td>
<td>1.6</td>
</tr>
</tbody>
</table>

detector array because its the fabrication process is simple and compatible with the standard IC processing. PolySi thermoelectric material is deposited using Low-Pressure Chemical Vapour Deposition (LPCVD) and n- or p-type areas are defined by ion implantation.

Simulations of the doping profile as functions of the dopant and doping energy have been performed for polysilicon, which is used as thermoelectric material. These simulation results give an indication of what implantation parameters such as the doping concentration and energy result in a better performance for a given thermopiles’ geometry.

The distribution of implanted ions as a function of the depth to which an ion penetrates a silicon substrate can be calculated as follows. As the impurity atoms enter the crystal, they lose their energy to the lattice of the crystal by collisions and finally stop at some average penetration depth. The penetration depth along the axis of incidence is called the projected range, $R_p$ which is shown in Figure 3-24. As expected, the projected ranges increase proportionally to the ion energy. The statistical fluctuations in the projected range are called the projected straggle, $\Delta R_p$ that is the half-width of the distribution at $e^{-1/2}$ of the peak. There is also a statistical fluctuation along an axis perpendicular to the axis of incidence, called the lateral straggle, $\Delta R_\perp$. The projected and lateral stragglers are shown in Figure 3-25. Along the axis of incidence, the implanted impurity profile can be expressed
by [3.39],

\[ n(x) = \frac{\Phi}{\sqrt{2\pi\Delta R_p}} \exp\left[ \frac{(x - R_p)^2}{2\Delta R_p^2} \right] \]  (3.54)

where \( \Phi \) is the ion dose per unit area. \( R_p \) and \( \Delta R_p \) can be derived for different dopants with a specific implantation energy from Figure 3-24 and Figure 3-25. Based on these projected range and straggle values, we calculate the distribution of implanted ions. Figure 3-26 shows the calculation results. Based on the simulated doping profiles, phosphorus and boron are chosen as n- and p-type dopants, respectively. The simulation showed that the majority of the (n-type) arsenic dopant is located near the surface with the applied energy of 40 keV. Boron and phosphorus are thus chosen for p- and n-type thermoelectric materials, respectively. Ion implantation is used instead of diffusion for better control of the dopant depth and concentration since diffusion puts the majority of the dopant atoms near the surface, where contaminants are also located [3.40].

![Projected range for boron, phosphorus and arsenic in Si and SiO₂ as a function of implant energy [3.41].](image)

*Fig. 3-24. Projected range for boron, phosphorus and arsenic in Si and SiO₂ as a function of implant energy [3.41].*
Fig. 3-25. Projected and lateral stragglers for boron, phosphorus and arsenic ions in silicon [3.41].
The thermopile-based infrared detector array

Fig. 3-26. Doping profile simulations for different dopants, (a) arsenic, (b) boron and (c) phosphorus.
References


The thermopile-based infrared detector array


[3.32] D.D.L. Wijngaards, S.H. Kong, P.M. Sarro and R.F. Wolffenbuttel, Thermophysical Characterisation of PolySi0.7Ge0.3 for Use in
Thermoelectric Devices, Transducers '01, June 10-14, 2001, Munich, Germany, pp.1010-1013.


Fabrication of the multi-slit grating based infrared micro-spectrometer

4.1 Technologies used for fabrication

4.1.1 Introduction

The ability of silicon microelectronic technology to fabricate structures with an ever decreasing feature size is the key to the success in this field [4.1]. These advances of microtechnology have enabled an impressive integration density of circuitry and the co-integration of mechanical, chemical, magnetic, thermal or optical structures in addition to the circuitry. The technology to realise on chips miniaturised components such as masses, springs, membranes, chemical sensing cell, mirrors, lenses is called micromachining.

Silicon micromachining is commonly used to fabricate low-cost sensors and actuators for high-volume markets, since silicon is the standard material for microelectronics fabrication and, at the same time, shows excellent
mechanical characteristics that are favourable for micro-scale structures [4.2][4.3][4.4]. Figure 4-1 shows a scanning electron microscope (SEM) photo of a monocrystalline silicon cantilever pressed by a tip. It shows an almost ideal elastic deformation.

The silicon-based pressure sensors, using a micromachined silicon membrane, and silicon-based accelerometers are already commercialised on high-volume markets with low prices. The demand for low-cost, accurate, and reliable sensors is drastically increasing for industrial and consumer product applications. Currently, enormous efforts are being dedicated to the development of the next generation of vehicle dynamic control micro-systems such as the gyroscope, and to the commercialisation of micro total analysis systems (μ-TAS) and biosensors [4.5][4.6][4.7][4.8][4.9]. The versatility of semiconductor materials and the miniaturisation of VLSI patterning techniques promise new sensors with better capabilities and an improved performance-to-cost ratio [4.10]. Moreover, the improved IC technology promises the realisation of more complicated micro-structures, not only for mechanical components but also for chemical and optical ones. The compatibility issues with the standard IC processing, however, restrict the material that can be chosen and the process that can be selected.

There are two major groups of micro-sensor technologies. One is bulk micromachining and the other surface micromachining. Bulk-micromachined sensors are primarily made by the accurate machining of a

![Fig. 4-1. A monocrystalline silicon cantilever deformed by a tip [4.2].](image-url)
relatively thick substrate by removal of selected parts using chemical attack. Surface-micromachined sensors are primarily constructed from stacked thin films and by selective removal of some of these films. These technologies use materials and processes borrowed from VLSI technology. The three processes of deposition, lithography and etching are mainly used to construct structures required for specific sensors.

In this chapter, a brief description of the micro-sensor technologies that have been used for realising the multi-slit transmission grating based IR micro-spectrometer are introduced followed by the description of the actual fabrication sequence.

4.1.2 Silicon IC technology

Standard integrated circuit silicon process technologies have been used with a few compatible post-processing steps to realise multi-slit-type aluminium-grating-based IR micro-spectrometers. The simplicity and compatibility with the standard IC processing had the highest priority in the design of the device.

4.1.2.1 Oxidation

Oxidation is the most common reactive growth process. In reactive growth, the chemically reactive species combines with the substrate to form a new thin film. When oxygen is in contact with a silicon surface at high temperatures, typically above 900 °C, it forms silicon dioxide [4.11]. The controlled growth of the silicon dioxide film on a silicon substrate is essential in the fabrication of the semiconductor device. Thermally grown silicon dioxides have few defects, neither in the bulk nor at the interface with the underlying silicon, resulting in a quality that is sufficient for use in device fabrication.

Since the oxide layer is a good electrical insulator and shows a low heat transfer coefficient, it is widely used as an insulator both in active devices such as MOSFETs for gate oxide and in the region between the active devices, known as the field. The other common use of silicon dioxide is as a mask, or barrier, against dopant diffusion into the substrate [4.12]. The silicon dioxide film grown on the silicon substrate is patterned by
photolithography (see 4.1.2.2 on page 118) for further process steps such as etching, doping and so on. Silicon dioxide is sometimes used as a bulk micromachined thin-film cantilever or membrane material combined with silicon nitride [4.13][4.14]. In addition, a thin oxide film can form the dielectric layer of a capacitor [4.15][4.16]. Because silicon dioxide is dense and hard enough to protect the surface of the substrate from dirt or physical scratching it is also used as a passivation layer. In the following sub-section some of the most common uses of silicon dioxide will be discussed.

**Doping barrier**

Doping is the introduction of a dopant into the exposed wafer surface through an opened window in a surface layer. In silicon technology, the surface layer acting as doping barrier is mostly silicon dioxide. The silicon dioxide patterned on the wafer blocks the dopant, preventing it from reaching the silicon. Diffusion or ion implantation is used for the doping. Dopants used in silicon technology have a very slow rate of movement in silicon dioxide as compared to silicon. That is why only a relatively thin silicon dioxide layer can be used to block the dopants. Moreover, silicon dioxide has a thermal expansion coefficient that is rather similar to that of silicon. Bulk silicon has $2.5 \times 10^{-6}$ and $4.3 \times 10^{-6}$ /K as the thermal expansion coefficient at 300 and 800 K, respectively, while the silicon dioxide film has a coefficient of $4 \times 10^{-7}$ /K [4.17]. This similarity makes high-temperature processes such as oxidation and diffusion possible without warping the wafer caused by a difference in the thermal expansion coefficient. The experimental measurements of the stress and its dependence on the temperature at Si-SiO$_2$ interface were reported by Jaccodine et al. [4.18].

**Align marks formation using differential oxidation rates**

For every 1 μm of SiO$_2$ growth, about 0.45 μm of silicon is consumed during the oxidation (see Figure 4-2) [4.10]. This consumption of silicon is due to the stoichiometric relationships in the chemical reactions and the difference between the densities of Si and SiO$_2$.

During the IC fabrication process, the wafer surface shows a variety of conditions. For instance, some areas may be covered with the field oxide,
Fig. 4-2. Silicon surface consumed during the oxidation.

Fig. 4-3. Making a step using differential oxidation of silicon.

some are doped, some are polysilicon regions and so on. Each of these areas has a different oxidation rate resulting in different oxidation thicknesses depending on the surface conditions. With these differential oxidation rates, alignment marks for stepper are made at the edges of the wafer. Figure 4-3 shows a step created by the oxidation of an exposed area next to a relatively thick field oxide. The oxide grows faster in the exposed area as additional oxide growth on the field oxide is limited by the thick field oxide. The
Fabrication of the multi-slit grating based infrared micro-spectrometer

faster-growing oxide in the exposed area consumes more silicon than that under the field oxide. After removal of the oxide from the whole wafer surface in hydrofluoric acid (HF), these silicon steps can be used as align marks.

Surface or inter-layer dielectric

Silicon dioxide classified as dielectric material plays an important role as an insulator with conductive layer(s) of metal deposited on top and/or the bottom of it. The oxide prevents shorting of the metal layer to underlying layers or to another metal layers. An insufficient quality and step coverage capability of the dielectric inter-layer can cause serious problems, that is, when the spacing between the metal lines becomes less than 0.35 μm [4.19]. The oxide used as inter-layer must be continuous without any voids and thick enough to prevent induction. Induction can occur when the oxide layer is too thin to avoid a build-up of charge in the underlying layer caused by an electrical charge in the metal layer. The accumulated charge on the wafer surface can cause undesirable electrical side effects. An oxide layer thick enough to avoid the induction from the metal layers is called a field oxide.

On the other hand, Metal Oxide Semiconductor (MOS) technology uses this induction phenomenon with a specifically controlled oxide thickness to allow the gate region under the oxide to be inductively charged [4.20]. Figure 4-4 shows an NMOS enhancement transistor in which the source and substrate are grounded and the drain-to-source voltage is set to zero. The positive voltage applied to the gate establishes an electric field directed perpendicularly through the oxide. This field induces negative charges at the surface of the gate region under the oxide layer. The induced charges

![NMOS enhancement transistor showing induced channel under the gate oxide.](image)

Fig. 4-4. NMOS enhancement transistor showing induced channel under the gate oxide.
beneath the oxide creates an n-channel. As the voltage on the gate increases beyond a threshold level, the number of induced negative charges in the channel that represents the conductivity of the channel increases.

Silicon dioxide dielectric layers are also used as the dielectric material formed between two conduction layers for a capacitor, and as an inter-layer in multi-level metallisation structures such as multi-chip modules. For those applications, the silicon dioxide layers are deposited with Chemical Vapour Deposition rather than thermal oxidation.

**Surface passivation**

On every surface of a solid semiconductor the surface atoms are incompletely coordinated; there are dangling bonds that are not shared with neighbours. The vacancies of the surface bind with available elements and usually form an oxide. Most often the surface is passivated intentionally with insulator materials like Si$_3$N$_4$ or SiO$_2$. The main goal of passivation is to minimize the influence of the semiconductor surface on the bulk properties of the semiconductor.

Silicon dioxide passivation layers play an important role in protecting the semiconductor surface and underlying devices from physical, chemical or electrical contamination as follows. Silicon dioxide is very dense (nonporous) and hard. The hardness of the layer therefore protects the wafer surface from scratching during the fabrication process. The silicon dioxide layer also acts as the contamination barrier for the sensitive wafer surface by preventing dirt from the processing environment. One dust particle on a sensitive area has a high probability to destroy the functionality of the whole device. The importance of cleanliness becomes more and more acute with shrinkage of the feature size, since fewer and smaller particles can be tolerated with smaller features.

Regardless of the cleanliness of the processing environment, some electrically active contaminants are gathered on the wafer surface. The oxide layers have been studied extensively to find a way to protect the electrically sensitive device from the mobile ionic contaminants. The alkali elements are easily ionized and have a relatively low mass [4.21]. These ions are typically very mobile in device structures. They are trapped at interfaces and can set up enormous charge fields, shifting the electrical
operating parameters of devices into an unintended range. Despite a lot of effort and progress in the study of the Si-SiO₂ interface, the interaction between the physical, chemical, structural and electronic properties of the interface and the conditions of its formation are still unclear. The lack of such a general understanding makes it difficult to develop surface passivation processes.

4.1.2.2 Photolithography

The patterns on each layer are defined by a process sequence called photolithography. Photolithography transfers the patterns from a mask onto a wafer by exposing, normally with an ultraviolet (UV) light source, photosensitive organic material (photoresist), which is usually spin-coated on the wafer. Photolithography is the most complicated, expensive, critical and time-consuming process step in microelectronic fabrication processing.

For positive resists, the resist is exposed with UV light wherever the underlying material is to be removed. In these resists, exposure to the UV light changes the chemical structure of the resist so that it becomes more soluble in the developer [4.22]. The exposed resist is then washed away by the developer solution, leaving windows of the bare underlying material. The mask, therefore, contains an exact copy of the pattern which is to remain on the wafer. Negative resists behave in the opposite manner. Exposure to the UV light causes the negative resist to become polymerized, and more difficult to dissolve. Therefore, the negative resist remains on the surface wherever it is exposed, and the developer solution removes only the unexposed portions. Masks used for negative photoresists, therefore, contain the inverse of the pattern to be transferred. Negative resists were popular in the early history of integrated-circuit processing, but positive resists gradually became more widely used since they offer better process controllability for small geometry features. Positive resists are now the dominant type of resist used in VLSI fabrication processes. The mask consists of transparent and opaque regions corresponding to the patterns. Figure 4-5 shows a cross-sectional schematic of a simple shadow printing system for optically exposing a wafer.

The key parameters of photolithography are the ability to achieve a high enough resolution and accurate alignment. For a good resolution, the dimensions of a pattern need to be as close to the original design as possible.
Alignment is to match the mask features to the structures on a wafer. The alignment capability becomes more critical, especially in a multi-layered planar process. The entire pattern of circuits or micro-structures must be correctly placed on the wafer surface and the individual parts of the patterns must be in the correct positions relative to each other. Moreover, as the final pattern results from several photo-mask steps applied to the wafer in a sequential manner, one misaligned mask layer can cause a failure in the performance of an entire device.

To avoid any mask damage problem associated with the shadow printing technique, such as contact or proximity printing method shown in Figure 4-5, projection printing exposure tools have been developed. These project an image of the mask patterns onto a wafer at a distance away from the mask. Figure 4-6 shows the image partitioning techniques by projection printing. M:1 represents the demagnification ratio of step-and-repeat projection (e.g., 10:1 for 10 times reduction on the wafer). After the exposure of one chip site, the wafer is moved to the next chip site and the process is repeated to cover the entire wafer surface. The 1:1 optical systems are easier to design and fabricate than M:1 reduction systems, but it is much more difficult to produce defect-free masks.

Fig. 4-5. Schematic of a simple lithographic exposure system.
Fabrication of the multi-slit grating based infrared micro-spectrometer

Fig. 4-6. *Image partitioning by projection printing (a) reduction step-and-repeat (b) 1:1 step-and-repeat.*

4.1.2.3 Patterning

After a photoresist image has been formed on the surface of a wafer, the next process step often involves transferring that image onto a layer under the photoresist. The patterned photoresist acts as a resistant mask when the underlying layer is processed with chemical treatment, ion implantation, or other appropriate processes. Faithful reproduction of the mask patterns on a substrate is the main requirement for patterning structures. Actually, it determines the accuracy of the device characteristic. The device structures are usually patterned by selective chemical etching of opened areas using photoresist as a mask. Etching involves physical attack or some combination of a chemical and physical attack. Finally, the photoresist is stripped, leaving the layer with transferred patterns.

Lift-off

Lift-off is one of the simplest patterning techniques. This technique is commonly used for patterning of thin metal interconnects which have a low
melting point. The sequence for lift-off is shown in Figure 4-7. A thick resist as an intermediate layer is spun and patterned on a substrate. Next, a thin layer of metal is deposited as a structural film on the whole substrate area using evaporation. The metal layer is selectively patterned by immersing the wafer in a solution that can resolve the photoresist used as intermediate layer. Finally, only the structural metal layer remain. With the lift-off technique, etch damage to the substrate is avoided and the patterning is done with infinite selectivity and without undercut. The lift-off process is usually done with an ultrasonic bath to avoid redeposition [4.23]. A multi-layer resist or a single-layer resist with a negative slope as the intermediate layer provides more accurate results, even for thicker metal films [4.24].

Wet chemical etching

Wet etching is a purely chemical process that entails dipping the substrate into an etching bath or spraying it with etchants that may be acid or alkaline. The result of wet etching can either be isotropic or anisotropic, depending on the materials to be etched, or the etchants used. If the material has an amorphous or polycrystalline structure, the wet etching always results in isotropy. In this section, only isotropic etching is introduced. There are also wet etching techniques that result in a high anisotropy due to different etch rates in different crystal directions. These techniques are discussed in more details in section 4.1.3.

Wet chemical etching has some drawbacks such as poor process control and excessive particle contamination. The major disadvantage of wet chemical etching for pattern transfer is the undercutting of the layer underneath the mask, resulting in a loss of resolution in the etched pattern as can be seen in Figure 4-8. This problem becomes more severe for narrow windows since the degree of lateral overetch becomes a higher fraction of the desired window width.
However, wet chemical etching can be highly selective and often does not damage the substrate. These are the main reasons why wet chemical etching is used extensively in semiconductor processing.

Wet chemical etching involves three essential steps: (1) the reactive species are moved (e.g., by diffusion) to the wafer surface, (2) chemical reactions occur at the surface, and (3) the reaction products from the etched surface are transported away [4.25]. Since it is generally desirable to have a large, uniform, well-controlled etch rate, the wet etch solution is often agitated in some manner to assist in the movement of etchant to the surface and the removal of the etch product.

The most common layer etched with wet chemical etching method is a thermally grown silicon dioxide. The basic etchant is hydrofluoric acid (HF). HF solutions have the advantage of dissolving silicon dioxide without attacking silicon, showing a extremely good selectivity: commonly more than 100:1 [4.12]. But a concentrated HF has an etch rate that is too fast (about 30 Å/s) for a controllable process [4.26]. In practice, the HF is mixed with water or ammonium fluoride (NH₄F) and water. Ammonium fluoride acts as a buffer to the unwanted generation of hydrogen ions, which accelerates the etch rate. The concentration and temperature of the buffered HF (BHF) are controlled to realize a reasonable etch rate for a particular oxide thickness. In order to reduce the surface tension of the etchant, a surfactant such as Triton X-100 can be added to allow the etchant to penetrate uniformly into smaller openings. Deposited oxides tend to be etched much faster. Often the etch rate of the deposited oxide film in HF is compared to that of thermal oxide to measure the density of the deposited oxide.
Dry etching

In order to achieve sufficient anisotropy, dry etching methods have been developed. In dry etching, the chemically reactive species is in a gaseous state, an ionised plasma in the form of low-pressure discharges. Dry etching is synonymous with plasma (assisted) etching. The principal mechanism of plasma etching is chemical reaction between the vapour-phase free radicals in the plasma and the solid-phase material of the film enhanced by the positively ionised energetic species. The ionised species in the plasma physically impact the etching film by bombardment due to the strong electric field stimulating the chemical reaction. The ion bombardment with a high directionality, which has a direction perpendicular to the surface of target wafer as can be seen in Figure 4-8, results in sufficient anisotropy. The reaction product is volatile and easily removed from the low-pressure reaction chamber.

Uniformity and etch rate are the key issues in both wet and dry etching. In dry etching however, there are extra matters to be taken into account, such as anisotropy and the etching selectivity between the mask and target material [4.27]. The anisotropic etching profile is a function of local temperature of the target and the directionality of ions related directly to ion energy. To control the etching profile, mixture gas is often used to form a passivation layer at the side wall. It is very difficult to get both a high etch rate and good selectivity at the same time because of the physical bombardment of the masking material. Recently, Inductively Coupled Plasma (ICP) sources have been widely used to achieve a higher plasma density at lower pressure, which results in more uniformity, low contamination and higher selectivity with lower ion energy [4.28][4.29].

There are also some plasma systems that are isotropic and rely only on the diffusion of ionised species to etch the substrate films. To avoid stiction problems, for example when a long and narrow beam structure is released in wet chemical etchant, isotropic dry etching systems are commonly used. Fluorine atom etching of silicon is a good example of this mechanism. The key factor of the process is to produce the reactive etchant species in plasma. Xenon difluoride (XeF₂) etchant has been applied to etch silicon selectively in a vapour phase by Winter et al. [4.30] and pulse-etching of silicon with xenon difluoride has also been reported by Chu et al. [4.31].
As the semiconductor technology moves in the direction of smaller features the advantages associated with anisotropic etching become increasingly important.

4.1.2.4 Chemical Vapour Deposition (CVD)

Two deposition techniques, sputtering and Chemical Vapour Deposition (CVD) have dominated the IC industry during the last 10 years. Sputtering is used mostly for metallisation films, whereas CVD is used for dielectric and polycrystalline silicon thin films [4.32]. CVD has become extremely popular and is the deposition method preferred for a wide range of materials. In general, CVD reactions require the addition of energy to the system, such as heating of the chamber or the wafer. The chemicals are then supplied through the use of gaseous compounds. Gaseous species are thermally broken down into their components in a deposition chamber. The process of thermal breakdown of molecules is known as pyrolysis. When the thermally dissociated gas species arrive at the surface of a substrate, some of these components nucleate on the surface and eventually form the thin film. Unlike evaporated or sputtered films, CVD films can be deposited uniformly with a better step coverage. CVD films have an amorphous or polycrystalline structure. The stress of the films can be reduced by control of the deposition conditions and subsequent annealing.

Presently, Low-Pressure CVD (LPCVD) is used most commonly for thin-film deposition, because it yields superior quality especially in mechanical characteristics. A recent innovation in the LPCVD area is the introduction of vertical chambers [4.12]. Similar to vertical oxidation/diffusion tubes, these systems have several advantages over standard tubes, such as better uniformity, easier installation, reduced particle counts and so on. The cost of these systems is considerably higher than that of conventional LPCVD systems, however.

Plasma-induced reactions are commonly used for the deposition of microsensor materials. One of such processes is known as Plasma Enhanced CVD (PECVD). In a PECVD reactor, gases are decomposed and ionised in a gas discharge to be sufficiently reactive to the substrate surface. PECVD films are deposited at a higher deposition rate and require a lower deposition temperature than LPCVD films, and thus are more suitable as IC-compatible post process step. However, PECVD films show inferior
characteristics in dielectric breakdown strength, pinhole density, impermeability against impurities and so on. This is mainly due to the large amount of hydrogen incorporated in the deposited films [4.32].

A special CVD process, so-called epitaxy, can grow monocrystalline films on crystalline substrates. Epitaxy can also be used to grow films with the same crystal structure as the substrate but with a totally different doping concentration. Furthermore, using epitaxy, it is possible to deposit lightly doped material on a heavily doped substrate resulting in a very sharp junction, which can not be obtained with the diffusion process [4.33].

4.1.2.5 Ion implantation

Ion beams are extensively used either to process or to characterise materials. In the keV to MeV energy range, ion beams are used to implant dopant atoms into a semiconducting wafer to alter its electronic properties. A p-n junction, for example, can be produced by implanting B into an n-type silicon substrate.

Ion implantation is a physical process, while diffusion is a chemical one. This means that implantation does not rely on chemical interaction between the dopant and the wafer material. In an ion implantor, dopant atoms are ionized and isolated, accelerated, formed into a beam and then swept across the wafer. The dose can be tightly controlled by measuring the ion current. By controlling the electrostatic field, the penetration depth of the dopant ion can be also controlled.

For high-density circuits or structures with small feature sizes and closer spacing, ion implantation is preferred to diffusion. An essential parameter of such devices is the lateral spacing between adjacent regions necessary to avoid shorting due to lateral diffusion. The ion implant process provides no remarkable side diffusion and a well-controllable doping profile, and the process takes place at close to room temperature. The lateral and vertical diffusion can be limited more by using rapid thermal annealing (RTA) [4.34]. RTA is the optimal method to activate the implanted dopants in a very shallow junction. The effects of different parameters for ultra-shallow implanted junctions with RTA has been investigated by Downey et al. [4.35].
4.1.2.6 Metallisation

Metal films are most commonly used in semiconductor technology for metallisation. Metallisation processes involve the deposition of one or more layers of conducting metal on the wafer surface in combination with the patterning that defines the electrical connections between the circuit components.

In order to get an Ohmic aluminium-silicon contact with pure aluminium leads, it is necessary to anneal the Al-Si interface at a temperature at which the aluminium atoms move around and spread more uniformly over the silicon surface. However, whenever aluminium contacts silicon, the silicon will dissolve into the aluminium during the annealing. The use of aluminium in integrated circuits at shallow junctions often creates problems such as spiking and electromigration [4.36]. Moreover, the Si-Al system has unfortunate eutectic characteristics, that is, if these two materials are heated when in contact with each other, they melt at a temperature of 576 °C, which is much lower than their individual melting temperatures. The melting points of pure aluminium and pure silicon are 660 and 1412 °C, respectively. The problem is that the temperature of the Si-Al contact must be limited to less than 576 °C because of the eutectic characteristics. If the surface has shallow junctions, the alloy region can extend partly or completely through the junction as shown in Figure 4-9(a). There are two solutions employed. One is using barrier metal layers such as titanium tungsten (TiW) and titanium nitride (TiN), that separate the aluminium and silicon to prevent the formation of the eutectic alloy (Figure 4-9(b)) [4.37]. The other solution is to add 1-2% silicon to the aluminium film [4.38]. Then the aluminium alloys with the added silicon instead of with the silicon substrate (Figure 4-9(c)). The patterning of this layer requires an etch step to remove the aluminium and a dip in a polySi etch bath to remove polysilicon grains [4.39].

![Fig. 4-9. Aluminium-silicon contacts.](image-url)
4.1.3 Bulk micromachining technology

Bulk micromachining is one of the earliest micromachining techniques [4.40][4.41] and it is still the most popular sensor fabrication technology. This technique removes the bulk crystalline silicon selectively using a mask on the front or back side of the wafer, resulting in precise three-dimensional structures. Most anisotropic etchings of crystalline silicon are based on liquid-phase chemical reactions using chemical etchants such as potassium hydroxide (KOH), ethylenediamine and pyrocatechol (EDP), tetra-methyl ammonium hydroxide (TMAH) and hydrazine water. A few techniques have been developed using vapour-phase and plasma etching [4.42][4.43]. These anisotropic etchants have different etch rates for the different crystallographic orientations of silicon.

In general, the etch rate is slowest in the \(<111>\) direction and fastest in the \(<100>\) and \(<110>\) directions. This orientation-dependent etching of \(<100>\) oriented silicon substrate through a patterned mask forms V-shaped grooves in an alkaline etching solution. The etched side edges are \((111)\) planes at an angle of 54.74° with respect to the \((100)\) surface as can be seen in Figure 4-10(a). The etching stops, in this case when the opening is small, since only \((111)\) planes are left. When the etching time is short or the opened window in the mask is large enough, a U-shaped groove will be formed (Figure 4-10(b)). A sharp corner of bulk silicon in KOH solution, shown in Figure 4-10(c), can be undercut because there are many uncoupled bonds at the sharp corner. In order to reduce the corner undercutting, compensation patterns are used [4.44]. Through mask patterns on \(<110>\) oriented silicon, deep vertical-walled etch profiles can be formed due to \((111)\) planes at the side walls, as can be seen in Figure 4-10(d) [4.45].

**Fig. 4-10.** Anisotropic etchings of \((100)\) and \((110)\) planes of silicon.
Some silicon etchants show a reduced etch rate in heavily doped ($\sim 10^{20} \text{ cm}^{-3}$) regions with boron [4.46]. Furthermore, an electrical potential externally applied on one side of p-n junction within the bulk silicon can also be used to influence the etch rate in the direction perpendicular to the junction. A small voltage of 0.6 V is applied to one side of the p-n junction of a wafer in an etch bath of EDP or KOH [4.47]. The etching stops when the exposed side of the p-n junction is completely etched away (see Figure 4-11). This electrochemical etch stop is due to a SiO$_2$ passivation layer that is formed by anodic oxidation when the etching solution reaches the junction.

This method is commonly used to fabricate structures with a reproducible thickness control [4.48]. A p-type silicon wafer with an n-type epi-layer is normally used. The n-type layer is biased positive with respect to a platinum counter electrode in the etching solution. The electrical metal contact to the n-type layer is protected from the etchant by a specially designed wafer holder. Etching stops at the epi-layer-substrate junction and thus the epi-layer thickness, which is very well controllable, determines the microstructure thickness. In contrast to the etch stop technique based on heavily doped areas, which can cause mechanical stress problems, the electrochemical etch stop technique makes it possible to fabricate electronic devices within the etch stop area [4.39].

Bulk micromachining with alkaline solutions is widely used to realise simple mechanical devices such as diaphragm pressure sensors and cantilever-beam piezoresistive accelerometers. It is also used to provide a

![Fig. 4-11. The electrochemically controlled etch stop (ECE) technique.](image-url)
thermal isolation by forming a thin membrane on a chip for applications such as thermal type infrared sensors. The etched surface of bulk silicon achieved by alkaline solutions is also used as an optical component such as mirror in optical sensors [4.49], since a smooth etched-surface can be obtained with KOH etching. An additional method has been studied to decrease the surface roughness by the reflow of deposited glass [4.50]. It is also possible to realise membranes using silicon dioxide or silicon nitride thin films as the membrane material. However, a silicon dioxide film is slightly soluble in KOH solution, so a silicon nitride mask should be used for etching silicon thicker than 200 μm [4.51]. In general, TMAH solution is used to realise a silicon oxide membrane despite its lower etch rate (typically a factor of 2) and lower anisotropy ratio (typically a factor of 4) than those of KOH solution. When silicon dioxide is used as the mask, hydrazine water, EDP and TMAH can be used as etchants.

Thermally grown silicon oxide has an etch rate of 12 nm/hr and a selectivity of 2500 with silicon in a 25% TMAH solution at 80 °C, and has an etch rate of 1000 nm/hr and a selectivity of 120 with silicon in a 33% KOH solution at the same temperature. On the other hand, KOH solution gives a serious contamination problem to the standard IC process due to mobile K⁺ ion contamination. In order to solve this problem KOH etching is done as the final step after the cleanroom processes are completed. Recently, TMAH etchant has received more attention due to its compatibility with standard IC processing. This is due to the fact that the tetra methyl ammonium ion in TMAH is rather large and does not diffuse into the silicon lattice [4.48].

**Dielectric material for thin-film membrane**

Silicon oxide is used as a structural or sacrificial layer in many micromachining processes as well as for the functions discussed in section 4.1.2.1. Oxide films are easily prepared on a silicon substrate by thermal oxidation, LPCVD, PECVD and so on. Oxide films always have a compressive residual stress. Therefore, a design for very small structures or thin membrane of oxide films can cause a mechanical buckling from the substrate. In general, silicon dioxide shows good mechanical strength and very low thermal conductivity (1.4 W/m-K).

Silicon nitride, which has a tensile residual stress, is also very suitable as a structural material due to its favourable mechanical properties. And it is
stable in an oxidizing atmosphere, making it useful as a mask for diffusion and for oxidation processes. Silicon nitride films are characterized by great mechanical strength and a relatively low thermal coefficient of expansion (4 \( \sim 6 \times 10^{-6} / \text{K} \)). The low heat conductivity (16–33 W/m-K) and the strong mechanical property of silicon nitride are ideal to realise micromachined thermal sensors, which need good thermal isolation on a membrane. Moreover, the thermal expansion coefficient of silicon nitride is almost the same as that of bulk silicon (2.5 \( \times \) \( 10^{-6} \) and 4.3 \( \times \) \( 10^{-6} \)/K at 300 and 800 K, respectively) so that high-temperature processes do not deform the device by a difference in thermal expansion. As a result, a low-stress silicon nitride allows one to make a very thin and large free-standing membrane [4.52]. In order to compensate for the residual stress, one can use multi-layer or sandwiched membranes containing both silicon dioxide and nitride layers [4.14][4.53][4.54].

In general, the movement of electrons can be blocked less easily in silicon nitride films than in silicon dioxide [4.32]. To provide a barrier to both electron and ion transport, silicon nitride films are often used in conjunction with silicon dioxide films. The composition of this silicon oxynitride \( (\text{Si}_x\text{O}_y\text{N}_z) \) is adjusted to provide near-zero stress. Other important applications of this film are the following:

- As final passivation layer in VLSI devices to avoid penetration of alkali and moisture [4.55],
- As inter-layers for double-metal technologies [4.56],
- As selective diffusion barriers [4.57],
- As gate films for MOSFET [4.58],
- As storage material in MNOS non-volatile memories [4.59][4.60],
- As oxidation masks in improved local oxidation of silicon (LOCOS) processes [4.61],
- As insulator films in irradiation-resistant devices [4.62][4.63]
- As sensitive insulator surface for an ISFET-based PH sensor [4.64] and
- As waveguide material for optical sensors [4.65][4.66].
4.1.4 Wafer level low-temperature silicon fusion bonding

More complicated sensors require more than one wafer process. Bulk micromachined, separately processed wafers can be bonded together at wafer level to form those complicated devices. Many wafer bonding techniques are currently available [4.67]. However, when a processed wafer contains temperature-sensitive components such as aluminium patterns, high-temperature processes adversely affect the performance of the integrated components. Therefore, direct fusion bonding cannot be used here. An alternative is Si-Au eutectic bonding [4.68]. However, the deposition of additional metal layers (Ti and Au) and reliability concerns limit the applicability of this technique as well. Much effort has been dedicated to developing direct low-temperature silicon wafer bonding techniques.

There are a few low-temperature fusion bonding processes available such as ultra-high vacuum bonding, treatments by plasma, sulphuric or nitric acid to increase bond strength, HF bonding and so forth [4.69]. A remarkable bond strength of 35 MPa can be achieved by an ultra-high vacuum bonding technique that is applied after the removal of native oxide and hydrogen passivation layers from the wafer surface by HF-dip and heating up the wafer to 800 °C before the bonding. This technique has a limited alignment because of the need of high vacuum. Moreover, this process is not applicable to the bonding of wafers that contain integrated circuits or metal patterns because of the high temperature necessary for the removal of hydrogen from the surface.

Several low-temperature annealing methods using plasma treatment have been reported, mostly using oxygen plasma and taking place at a lower temperature than 400 °C, followed by RCA treatment or water rinse before bonding. Despite the low temperature necessary to increase the bond strength, this technique is not reliable yet and needs more research on the influence of the pre-process cleaning methods [4.69]. The bond strength can also be enhanced by dipping the wafers in strong oxidizing acids such as sulphuric or nitric acid instead of RCA cleaning. This technique has been used for the fabrication of the IR micro-spectrometer.

The pre-treatment of the wafer is essential for the final bonding quality. Bare silicon and oxidized wafers are immersed in an ammonium fluoride
etch mixture [4.69] for 10 and 1-2 seconds to remove the native oxide and to create open bonds at the oxide surface, respectively. The wafers are rinsed in deionized (DI) water for 5 minutes and then immersed in 100% nitric acid for 10 minutes. The 100% nitric acid does not affect the metal patterns on the wafer. Finally, the wafers are rinsed in DI water for 1 hour. Immediately after rinsing and spin-drying, the two wafers to be bonded are aligned and brought into contact with each other. In order to get a closer contact between the wafer surfaces, the wafers are pressed with a pressure of 1.5 kg/cm² for 10 minutes at room temperature. For a strong bonding, the prebonded wafers are annealed in a vacuum furnace at 100 mbar. The temperature range required for sufficient bonding is 120 - 400 °C. This range is suitable for bonding fully processed wafers that contain integrated circuits and metal interconnections. The annealing time is about 2 hours. The bond strength obtained with this low-temperature fusion bonding technique is 15 - 20 MPa.

4.2 Fabrication of the IR micro-spectrometer

When micromachined devices are merged with typical standard integrated electronics, compatibility issues arise because of restrictions on the thermal budget of the wafer and because of restrictions on the use of etchant or processing materials in the cleanroom environment. In the case of bulk micromachining, generally the processing itself is incompatible with the cleanroom environment. By postponing incompatible processes to the end of the processing, this problem is solved as no cleanroom re-entrance is required. Metallisation, which is the last step in the cleanroom processes, is quite often a problem in this approach, since the metal layer can easily be attacked by chemical etchants. This problem is solved by using a specially-designed wafer holder to protect the metal layer formed on the front side (see Figure 4-11).

In the case of surface micromachining, there are also several compatibility concerns such as the thermal budget caused by annealing, non-planarity of the wafer surface by the deposition of multiple layers and the attack of electrical components by the etchant used to remove the sacrificial layer.

Several works discuss the development of strategies for combining the
micromachining processes with the standard process in a compatible way, namely [4.70][4.71] and [4.72]. In this section, the fabrication scheme of an infrared micro-spectrometer is described. A multi-slit grating-based infrared micro-spectrometer was designed and fabricated in compliance with the constraints of a standard IC process [4.73]. Aluminium, which is the standard interconnection metal, has a very good and consistent reflectance in the infrared region. Note that the infrared optical parameters are quite different from those in the visible wavelength region [4.74]. Moreover, unlike other micro-optical components, multi-slit diffraction grating can be fabricated easily with high accuracy using the standard IC metallisation process [4.75]. As a result, an effective grating can be achieved, while the fabrication process remains compatible.

Silicon is known as a highly transparent material for wavelengths, exceeding 1 μm, beyond which free-carrier absorption can be disregarded [4.76][4.77]. The optical absorption coefficient strongly depends on the wavelength. For a wavelength shorter than the value that corresponds to the band-gap energy, the absorption coefficients increase rapidly. The steepness depends on the type of band-to-band transition. The bulk silicon can therefore directly be used as optical path for infrared light; the removal of bulk silicon using the bulk micromachining technique is not necessary, as is the case for a spectrometer operating in the visible spectral range [4.75].

The choice of a thermopile-based infrared detector array fabricated from polysilicon thermoelectric material is prompted by its compatibility with IC processing and its wavelength-independent sensitivity in a wide spectral range [4.78]. When compared with resistive bolometers, thermoelectric detectors have a low responsivity. On the other hand, bolometer-based detectors are susceptible to self-heating effects, due to the bias current. This can be a major issue in resistive bolometers incorporating semiconductor layers with a non-linear or negative resistive temperature dependency [4.79]. As thermoelectric-based infrared detectors fundamentally measure temperature difference, the heating of the absorber by the radiation of incident infrared is measured, irrespective of the ambient temperature. Moreover, thermoelectric detectors do not require electrical biasing, show no excess current-induced noise, and are relatively simple to fabricate.

The design of the multi-slit based IR micro-spectrometer is functionally divided into two parts. Consequently, these two functions are realised in
Fig. 4-12. Functional structure of a multi-slit based infrared micro-spectrometer.

separate wafers that are bonded as a final step. The first wafer contains the grating and the other the infrared detector array, as shown in Figure 4-12.

4.2.1 Multi-slit grating

An aluminium multi-slit grating is fabricated on a double-side polished wafer to reduce scattering of the diffracted infrared light at the surfaces and to enhance the adhesion to the detector array wafer during the bonding.

Figure 4-13 shows the schematic fabrication sequence of the aluminium multi-slit grating. Firstly, 100 nm thermal oxide is grown to keep the wafer surface as clean as possible during the processing, until low-temperature fusion bonding with the detector array wafer takes place. The oxide at the front side is removed prior to the deposition of the aluminium layer used for the multi-slit grating, as this oxide layer at the opened area might cause interference or otherwise affect the effectiveness of the grating. This thermal oxide layer on the back side is also used as a mask for a short KOH etch, to form recessed parts. A thicker thermal oxide or an extra silicon nitride layer would be necessary for a deeper KOH etching to enlarge the dispersing distance of the incident infrared light. The etch rates of thermally
grown oxide and bulk Si of (100) plane are about 360 nm/hour and 1 μm/min, respectively, in 33 wt% KOH solution at 85 °C. The 0.6 μm thick aluminium grating material is deposited by sputtering and patterned by dry etching.

Subsequently, the thermal oxide on the back side, or the SiO₂/Si₃N₄ composite layer required for a deeper recess by longer KOH etching, is patterned by means of dry etching. The back side of this grating wafer is
Fabrication of the multi-slit grating based infrared micro-spectrometer

recessed using 33 wt% KOH solution at 85 °C to reduce the heat loss from the infrared detector to the bulk of the grating wafer, while avoiding mechanical contact to the detector array. Multiple reflections of the infrared in bulk silicon can be considered to lengthen the optical path by at least 3 times using an aluminium mirror deposited on the back side of the wafer. In the case of no reflection type, some amount of recess is indispensable since the fixed optical path length corresponding to the wafer thickness is not long enough to achieve a reasonable dispersion distance due to the high refractive index of bulk silicon.

In order to get access to the bond pads located on the lower wafer (the detector array wafer) the bottom of the grating wafer is cut in 300 µm prior to the wafer-to-wafer bonding step, using a dicing saw. The remained thickness is cut to reveal the bond pads for wire bonding after the wafer-to-wafer bonding is completed. To facilitate this cutting step, aluminium line marks are patterned simultaneously with the patterning process of the multi-slit grating to indicate the saw lane from the top.

Figure 4-14 shows magnified images of the fabricated multi-slit gratings. Various gratings, with a pitch (=grating constant) varying from 4 to 20 µm, have been fabricated to split the detectable wavelength range into a few sub-bands. The grating with the constant of 4 µm clearly shows small undercuts at the opened slits. This is due to the fact that the multi-slit grating is

![Fabricated multi-slit gratings with different grating constants](image)

**Fig. 4-14.** Fabricated multi-slit gratings with different grating constants (a) and SEM photographs of the grating (b).
fabricated by the IC standard metallisation process, which was originally set for a small over-etch originally to avoid short circuit. It is interesting to note that these apparently trivial details of a process complicate the design of an on-chip microsystem. This undercut has less optical effect in case of a larger grating constant as can be seen in Figure 4-14. Nevertheless, the undercut grating affects the optical diffraction results, which will be discussed later.

4.2.2 IR detector array

Another wafer is processed to form silicon nitride membranes that contain a polysilicon thermopile-based infrared-detector array to detect the incoming incident infrared light after it has been dispersed through a multi-slit grating and travelled an optical path. Figure 4-15 shows the schematic fabrication sequence.

The process also starts with thermal oxidation (Figure 4-15 (a)). The purpose of the thermal oxide layer is to protect the front-side surface of the detector array wafer until the wafer is bonded with the prepared grating wafer. The oxide on the back side is therefore removed in BHF. The thickness of the oxide layer remaining on the front side is about 100 nm. Subsequently, low-stress silicon nitride and polysilicon are deposited on the oxide by LPCVD with thicknesses of 500 and 300 nm, respectively (Figure 4-15 (a)-(b)). These low-stress silicon nitride and polysilicon layers formed for the membrane and thermoelectric material show a tensile stress of about 400 and 1000 µstrain, respectively [4.39][4.80]. The low-stress silicon nitride on the back side of the wafer is used as a mask layer during KOH etching.

P- and n-type thermopiles are defined by boron and phosphorous ion implantation, respectively (Figure 4-15 (c)). The implant energy and dose for both the n- and p-type regions are chosen as 40 keV and 5x10^{15} ions/cm^2, respectively, based on the simulation results discussed in Chapter 3.

P- and n-type polysilicon thermopiles are patterned by RIE (Reactive Ion Etching) using CF_4, SF_4 and O_2 gas mixture with gas flow rates of 70, 10 and 10 sccm, respectively. Under the condition of 60 W of RF power and 0.05 mbar of pressure, the etch rate of polysilicon is about 120 nm/min. Subsequently, the polysilicon on the back side of the wafer is removed by
wet etching. The front side of the wafer is covered with photoresist to protect it during polysilicon wet etching of the back side. LPCVD TEOS is deposited as the inter-layer instead of PECVD oxide for a better step coverage (Figure 4-15 (d)). For restoration of the crystal damage during the ion implantation and for electrical activation of the dopants, an annealing has been performed in argon gas at 600 and 1000 °C for 20 and 35 minutes, respectively. The reflow associated with this annealing step also reduces the roughness of the surface of LPCVD TEOS. The temperature of annealing (600-1000 °C) is below the diffusion temperature of the dopant to prevent lateral diffusion. Subsequently, contact windows are opened by dry etching.
Fabrication of the IR micro-spectrometer

The contact resistance plays a significant role in the performance of thermoelectric devices [4.81]. The TEOS layer is therefore overetched into polysilicon by dry etching and the wafer is dipped in 0.55% HF for 4 minutes to reduce the contact resistance of the hot and cold junctions of the thermopile (Figure 4-15 (c)).

A 0.6 μm thick 99%Al/1%Si layer is deposited by sputtering at 350 °C and patterned as the metallisation step (Figure 4-15 (f)). After completion of the front side metallisation, the TEOS and nitride layers on the back side are patterned by RIE for KOH etching (Figure 4-15 (g)). The detector array wafer is prepared for bonding with the grating wafer by patterning of the bonding area on the front side using a combination of dry and following wet etching (Figure 4-15 (h)). The TEOS and nitride layers on the bonding area are dry etched. The remaining thermal oxide in the etched area is removed in BHF solution followed by cleaning directly prior to bonding. The reason for wet etching of the oxide layer is that the ion bombardment during the dry etching would make the bonding surface dirty and rough. The photoresist mask is then removed in an oxygen plasma.

Figure 4-16 shows the fabricated thermopile-based infrared detector array on a silicon nitride membrane. The hot and cold junctions of thermopile

![Fabricated IR detector array on a silicon nitride membrane.](image)
located on the inner side and on the rim area of silicon nitride membrane, respectively, are connected by aluminium to avoid a series of p-n junctions.

Figure 4-17 shows a photograph of the back of the IR detector array wafer taken after the KOH etching carried out to release the silicon nitride membrane. The back side align marks to align the front and back sides of the IR detector wafer are located at the rim area.

Fig. 4-17. The back side of the IR detector array wafer, taken after the KOH etching.

4.2.3 Low-temperature wafer-wafer fusion bonding

The next step is the bonding of the detector array wafer and the grating wafer, prepared as explained in the previous sections. The recesses in the grating wafer ensure that there is no mechanical contact between the two wafers. The cut-in parts at the bottom of the grating wafer made in advance using a dicing saw make access to the bond pads possible even after the bonding step. However, it is necessary to treat the fragile grating wafer very carefully especially during spin-dry step, as it has cut-in parts.

Figure 4-18 shows a schematic fabrication sequence of wafer-wafer bonding, KOH and wire-bonding processes with independently processed wafers. Both the grating and the detector array wafers are cleaned and
Fabrication of the IR micro-spectrometer

Fig. 4-18. Schematic fabrication sequence of wafer-wafer bonding, KOH etching and wire-bonding processes with independently processed wafers.

pretreated as follows just before the bonding step. The wafers are immersed for 10 seconds in an ammonium fluoride etch mixture [4.69] in order to remove the native oxide and rinsed in deionized (DI) water for 5 minutes. The silicon wafer surface is pretreated with sulphuric or nitric acid for hydrophilisation. Several hydrophilic silicon fusion bonding techniques are reported mainly for silicon-on-insulator applications [4.82][4.83]. The cleaning step in 100% HNO₃ for 15 minutes results in a chemical oxidation of the silicon surface at the bond area. This layer has a thickness of 0.5 - 1 nm and contains OH-groups which are responsible for the hydrophilicity and act as the media of hydrogen bridges at wafer-to-wafer interfaces. The wafers are rinsed again in DI water for 1 hour. Then they are hydrophilic and ready for pre-bonding.
The two wafers, the grating wafer and the IR detector array wafer, are aligned with a double-side aligner according to the full-wafer alignment marks prepared on both wafers and brought into contact immediately after the cleaning and spin-drying. The wafers are pressed with a pressure of 1.5 kg/cm² for 10 minutes at room temperature. In order to detect unbonded areas, the prebonded-wafer pair is inspected with an infrared-sensitive camera. The bonding strength of prebonding is already sufficient to handle the wafer stack for further processing.

The prebonded wafer stack is annealed in a vacuum chamber for 2 hours at a pressure of 100 mbar and with a temperature of 400 °C. Figure 4-19 shows the measured bond strength of two bare silicon wafers as a function of annealing temperature. In the case of bare silicon, a bond strength of more than 20 MPa is achieved with the conditions mentioned above. A sufficient bond strength can be also achieved at lower temperatures, as can be seen in Figure 4-19.

KOH etching is performed after the bonding, using a special tool to protect the front side of the bonded wafer from the attack by KOH solution (Figure 4-18 (b)). Alternatively KOH etching can be done prior to the fusion bonding. The bonded and etched wafer is sawn, where care is taken that the thin membranes are not damaged and the cut-in part is removed.

![Bond strength vs. annealing temperature](image)

**Fig. 4-19.** Measured bond strength of two bare silicon wafers as a function of annealing temperature (Annealing time=4 hrs).
Fabrication of the IR micro-spectrometer

(Figure 4-18 (c)). Figure 4-20 shows a top view of graphical composition of this device, with the upper half showing the detector array, and the bottom half the aluminium grating.

Finally, the completed chips are mounted and glued on standard 24-pin DIL IC housing and wire bonded (Figure 4-18 (c)). The final result, a fully packaged and wire-bonded IR micro-spectrometer is shown in Figure 4-21.

**Fig. 4-20.** Fabricated IR micro-spectrometer; Aluminium grating (top wafer), thermopile-based detector array (bottom wafer).
Fig. 4.21. Fusion-bonded and packaged IR micro-spectrometer.
References


Fabrication of the multi-slit grating based infrared micro-spectrometer


Fabrication of the multi-slit grating based infrared micro-spectrometer
Measurements

5.1 Introduction

Measurements are required at various levels. Firstly, prior to device design and fabrication, the basic electrical and thermoelectrical characteristics must be known. The measured parameters are compared to data obtained from the literature to verify the correctness of the processing, which is also an important step to realise reproducible devices. The characteristics of LPCVD films, such as the silicon nitride membrane, polycrystalline silicon (polySi) or polycrystalline silicon germanium (polySiGe) thermoelectric material, can vary greatly due to differences in deposition processes and equipment. A few test structures with doped polySi and polySiGe have been developed to determine fundamental material parameters, such as contact resistance, resistivity, Seebeck coefficient and thermal conductivity. The measurement results of fundamental material parameters from the fabricated test structures are briefly described in the next section. Secondly, the optical and thermoelectrical characteristics of the fabricated device components and of the whole devices need to be measured, as is described.
in the subsequent sections.

5.2 Electrical and thermoelectrical measurements

The basic electrical and thermoelectrical measurements have primarily been performed on polySiGe instead of polySi because of its significantly lower thermal conductivity, which yields a superior figure of merit [5.1], while offering a very acceptable level of compatibility with standard integrated circuit technology.

Each thermocouple, which consists of n-type and p-type polySi or polySiGe, is interconnected by aluminium to avoid a p-n junction. As discussed in Chapter 3, contact resistance at the interface between the thermoelectric material and aluminium can have a large influence on the device performance. The contact resistance of the fabricated devices at the interface has been determined using Kelvin test structures [5.2]. Figure 5-1 shows the measured contact resistance as a function of contact window area. The contact resistance of the p-type thermoelectric material is lower and exhibits less variance as compared to that of the n-type material.

Fig. 5-1. Measured contact resistance at the interface between aluminium and polySiGe [5.1].
**TABLE 5-1.** Measured resistivities of polySi and polySiGe thermoelectric layers (300 nm thick).

<table>
<thead>
<tr>
<th>Thermoelectric material</th>
<th>Type</th>
<th>Resistivity [μΩm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>polySiGe</td>
<td>n</td>
<td>38.2±0.6</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>24.5±0.1</td>
</tr>
<tr>
<td>polySi</td>
<td>n</td>
<td>47.9±1.8</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>28.3±1.6</td>
</tr>
</tbody>
</table>

The resistivities of both polySi and polySiGe have been measured using Van der Pauw structures [5.1][5.3][5.4] and are shown in Table 5-1. The thermal and thermoelectric components in the test structures have been designed and fabricated such that they have the same geometric parameters and doping concentration as the device to be used in the actual spectrometer.

In addition to the electrical and thermal parameters, the Seebeck coefficient has been determined using micromachined structures similar to those reported by Arx et al. [5.5]. Figure 5-2 shows one of the structures fabricated to determined the Seebeck coefficient. A 4-wire resistor is formed at the tip of cantilever using p-polySiGe, which is used to simultaneously heat the tip and measure its temperature. This resistor is

![Fig. 5-2. A test structure to determine the Seebeck coefficient [5.1].](image-url)
Measurements

electrically connected to bond pads by long S-shape leads to reduce the heat loss. To measure the temperature, it is necessary to know the resistivity and temperature coefficient of resistance (TCR) of the heater beforehand, which have been determined as $24.5 \pm 0.1 \ \mu \Omega \text{m}$ and $(10.6 \pm 0.7) \cdot 10^{-4} / \text{°C}$, respectively. From those parameters, the temperature at the tip of the cantilever, $T_{\text{high}}$, can be known. A second resistor is formed to measure the potential temperature rise on the substrate, which gives the temperature value, $T_{\text{low}}$. The Seebeck coefficient, $\alpha$, can be derived from the relation

$$\alpha = \frac{(V_{\text{high}} - V_{\text{low}})}{(T_{\text{high}} - T_{\text{low}})}.$$

The remaining missing parameter, $(V_{\text{high}} - V_{\text{low}})$, is determined through the electrical contact at both ends of the polySiGe thermoelectric specimen. These ends are positioned on the substrate and the tip of the cantilever, respectively. They are designed to be close to the two heaters to minimise the thermal resistance.

Figure 5-3 shows the measured result of n-type polySiGe in the temperature range of -25°C to 125°C. The combined Seebeck coefficient of the n- and p-type polySi thermoelectric material was about 240 $\mu \text{V/K}$ at ambient temperature.

![Figure 5-3](image)

*Fig. 5-3. Measured Seebeck coefficient of n-type polySiGe as a function of temperature [5.1].*
5.3 Optical quality of fabricated gratings

The optical quality of the fabricated multi-slit gratings, with grating constants ranging from 4 to 20 μm, has been examined for a single wavelength prior to the wafer-wafer fusion bonding. The grating was illuminated using a wide-band light source through a monochromator. The resulting diffraction pattern that was projected through the bulk silicon grating wafer was recorded by a CCD. Figure 5-4 shows the measurement set-up. The recorded diffraction pattern of a 60-slit grating with $a=4 \mu m$ and $d=2 \mu m$, as an example, is shown in Figure 5-5. Because the sensitivity of

---

**Fig. 5-4.** Measurement set-up to record the diffraction pattern of multi-slit gratings.

---

**Fig. 5-5.** Recorded diffraction pattern of a grating ($N=60, a=2d=4 \mu m$) for different wavelengths (a) 1 μm (b) 1.05 μm and (c) 1.1 μm.
the CCD is limited to radiation in the visible and near-IR range, wavelengths in near-infrared are used ($\lambda=1000$, 1050 and 1100 nm). The images in Figure 5-5 actually represent the intensity distribution as a function of distance on a screen (in this case the CCD). The intensity decreases drastically with increasing diffraction order. It also shows the increase of the dispersion distance with the wavelength of the incident light. Even though the grating constant in the design is twice the slit width, i.e. $a=2d$, the even order diffraction patterns did not disappear completely. This is due to the fact that the practical grating constant is not exactly twice the slit width due to the lateral overetch of aluminium set for the standard IC processing, as discussed and shown in chapter 4. This effect is most profound when the grating constant is 4 $\mu$m, which is the smallest grating constant amongst the fabricated multi-slit gratings.

The dispersion distances as a function of wavelength for the recorded diffraction patterns are in very good agreement with the calculations, as shown in Figure 5-6, considering the optical path length (the thickness of bulk silicon is 520 $\mu$m and the distance between grating wafer and CCD is 1 mm). The dispersion distance has been calculated using equation (3.15).

![Diagram showing dispersion distances](image)

**Fig. 5-6.** Simulated and measured dispersion distances of a fabricated multi-slit grating as a function of wavelength.
Obviously, only a qualitative conclusion can be drawn from this result, due to the very narrow spectral range. More conclusive results are given in the next sections.

5.4 Thermoelectrical response of the thermopile-based IR detectors

Subsequently, the thermoelectric response of an individual infrared detector element is measured with the measurement set-up shown in Figure 5-7. Optical systems often require the use of a chopper to modulate the light beam either because a detector, such as pyroelectric detector, does not respond to continuous radiation or because the signal-to-noise ratio can be improved by modulation and coherent detection techniques. The radiation power of the monochromator was measured in advance as a reference using a pyroelectric radiometer system (Oriel model 7080). The output voltage per unit radiation power of our polysilicon thermopile IR detector varies between about 10 and 16 V/W when radiated by monochromatic light in the wavelength range 1-9 µm, as shown in Figure 5-8. More consistent output voltage per unit radiation power is observed (~14-16 V/W) in the wavelength range 1.5-5 µm, where the spectral responses of the fabricated IR micro-spectrometer are measured and shown in the next section.

![Diagram of measurement set-up](image)

*Fig. 5-7. Measurement set-up to measure the spectral responses of IR detector array.*
5.5 Spectral response of the IR micro-spectrometer

The spectral responses of the fabricated IR micro-spectrometer are measured for grating constants 4 and 6 μm, using incident light 2.4, 3, 3.5, 4, 4.5 and 5 μm. For all cases, the optical-path length is composed of 25 μm of Si and 500 μm of air, which means that 500 μm of bulk silicon is etched away.

When the wavelength of incident light is 2.4 μm (Figure 5-9), the 2nd detector shows a peak that corresponds to the wavelength 2.36 μm. Moreover, 4th and 5th detectors and 7th and 8th detectors, which correspond to the 2.99-3.2 μm and 3.46-3.55 μm wavelength ranges, respectively, show peaks for the wavelengths of 3 and 3.5 μm. For the wavelengths of 3.5, 4, 4.5 and 5 μm (Figure 5-10), the peaks are on the 2nd, 3rd, 4th and 6th detectors, respectively, which correspond to those wavelengths of incident lights. However, these peaks are hardly resolved when the neighbouring wavelengths are illuminated simultaneously on the grating. An improved thermal isolation between the thermopile-based infrared detectors in the
array should enhance the sharpness of the response.

The measured spectral responses of the fabricated IR micro-spectrometers have demonstrated that all peaks were found on the detector number anticipated, corresponding to the wavelength of incident light. The half-power bandwidth is about 0.3 - 0.5 μm.

**Fig. 5-9.** Spectral responses of the IR micro-spectrometer illuminated using a blackbody source and a monochromator set at $\lambda=2.4$, 3, and 3.5 μm ($a=4$ μm, $L=25 \text{ mm} \ [\text{Si}] + 500 \text{ mm \ [air]}$).

**Fig. 5-10.** Spectral responses of the IR micro-spectrometer illuminated using a blackbody source and a monochromator set at $\lambda=3.5$, 4, 4.5 and 5 μm ($a=6$ μm, $L=25 \text{ mm} \ [\text{Si}] + 500 \text{ mm \ [air]}$).
References

[5.1] D.D.L. Wijngaards, S.H. Kong, P.M. Sarro and R.F. Wolffenbuttel, Thermophysical Characterisation of PolySi0.7Ge0.3 for Use in Thermoelectric Devices, Transducers '01, June 10-14, 2001, Munich, Germany, pp.1010-1013.


Conclusions

6.1 The integrated infrared micro-spectrometer

Multi-slit grating-based infrared micro-spectrometers for the spectral analysis of infrared light have been realised. The fabrication process is relatively simple, because separate processes are used for the upper and lower wafers, which contain a multi-slit grating and infrared detector array, respectively. By postponing the incompatible KOH etching process to the end of the sequence with final wafer bonding, the compatibility problem with the clean-room environment is solved. The fabrication is also less demanding than that of spectrometers designed for operating in the visible part of the spectrum. Much less morphological demands are required in the fabrication of optical components, due to the longer infrared wavelengths. The transparent optical property of silicon and the high reflectance of aluminium in the infrared spectral range have shown to be very convenient for the realisation of the integrated infrared micro-spectrometer in a simple and IC compatible fabrication sequence.
Conclusions

The observed optical performance of the fabricated multi-slit grating is in good agreement with diffraction theory. The fundamental material parameters, such as contact resistance, resistivity and Seebeck coefficient, measured with test structures, are in a reasonable agreement with the characteristic values found in literature.

Each IR detector in the array corresponds to the given wavelength of incident light, with the half-power bandwidth of about 0.3 - 0.5 μm. The spectral resolution of the fabricated micro-spectrometer is about 0.6 μm, despite the large dimension of the thermopile-based IR detector (the width of a single detector = 100 μm). By increasing the optical path length using multi-reflection, we could improve the resolving power further for the same detector size since it is proportional to the optical path length [6.1].

The IR detector size limits the resolution in the thermopile-based IR microspectrometer, whereas the photolithography limits the fabrication of the multi-slit grating and is thus the resolution limiting factor in a microspectrometer for visible light.

Combining a conventional microelectronic technology and micro-electromechanical systems (MEMS) processing for optical components has yielded small-size and low-cost solutions for many industrial and consumer applications.

6.2 Applications

The spectral analysis system has found many applications, not only in science but also in industry. The information from the spectral analysis is being used for astronomic observations, chemical analysis, security, quality inspection and agricultural testing.

Miniaturized spectral devices operating at visible wavelengths, such as integrated interferometers [6.2] and micro-spectrometers [6.3], have been demonstrated. Moreover, such devices are already commercially applied for colour measurement [6.4][6.5] as well as for reprography, to achieve true reproduction of colour despite variable environments, or for dentistry, to determine the colour of a patient’s teeth.
Using infrared for spectral analysis adds much to visible analysis, as is explained in Chapter 2, since it is in the infrared part of the electromagnetic spectrum that the emission and absorption lines of almost any type of molecule are found. Recently, analytical applications of infrared have become more important in the field of molecule detection and biotechnology [6.6]. Especially spectral analysis systems using near infrared are in great demand in diagnostics, DNA sequencing, bio-analysis of immunoassays as well as in medical and pharmaceutical applications. Miniaturization of the spectral system continues to bring breakthroughs in those infrared spectral applications.

6.3 Future work

The IR micro-spectrometer described in this thesis is intended to demonstrate the feasibility of its realization remained compatible with IC standard processing by the micromachining technology. The demonstrated device can be further improved. For instance, read-out circuitry could easily be added to the IR detector array, due to the fabrication compatibility. The optical path length could be drastically extended with a simple modification in the layout to obtain multi-reflection (see Figure 6-1(a)). Due to the good transparency of silicon and the excellent reflectivity of aluminium in the infrared spectral range, multi-reflection will improve the resolution of the device without significant loss of incident infrared light.

By adding a more effective absorbing area, we can improve the performance of the IR micro-spectrometer even further.

Finally, the fabrication sequence could be much simplified by using, e.g., a single-chip process with surface micromachining technology (see Figure 6-1(b)). A multi-reflection scheme can also be applied, but then it is necessary to implement back illumination and front shielding.

A major limitation of the device described in this thesis in terms of achievable spectral resolution is the fact that it is a lensless system. A lens in between the grating and the detector array would greatly improve the performance, but at the expense of increased fabrication complexity.
Fig. 6-1. IR micro-spectrometer: (a) Multi-reflection type, and (b) surface-micromachined type.
References


An optical analysis system operating in the infrared spectral range has many applications. It is especially useful in chemical analysis systems, since it is in the infrared part of the electromagnetic spectrum that the emission and absorption lines of almost any type of molecule are located. Miniaturisation of those spectral systems could lead to breakthroughs, not only in science, but also in industry. This thesis describes the design, fabrication and characterisation of a miniaturised spectrometer that operates in the infrared range of the electromagnetic spectrum.

The device described in this thesis is functionally divided into two parts. The first part is the top wafer, which contains the dispersing element, a multi-slit grating, realised by the standard metallisation step of IC processing. The second part is the bottom wafer, which has the thermopile-based infrared detector array fabricated on a silicon nitride membrane for thermal isolation of its sensing area using bulk micromachining. The multi-slit grating-based dispersive element separates the spectral components that comprise the incident radiation. The diffracted spectrum is projected onto the infrared detector array after traversal of the bulk silicon optical path. Consequently, a specific spectral component or a specific range of wavelengths is detected at a specific infrared detector within the array.

The high reflectance of aluminium at infrared wavelengths merits the
realisation of the multi-slit grating in a compatible way with standard IC processing. The bulk silicon is used to define the optical path, as silicon is highly transparent for wavelengths exceeding 1 μm.

The thermopile-based infrared detectors in the array are designed to have optimised geometric values based on the simulation results and the properties of the material used, which result in the best performance in terms of its detectivity. For the IR detector that has an absorbing area of 100×100 μm², as an example, the beam length is 300 μm, which corresponds to the beam length to result in the maximum detectivity (see Figure 3-16).

The observed optical performance of the fabricated multi-slit grating is in good agreement with diffraction theory. Moreover, measurements of the spectral response of the IR micro-spectrometer demonstrated that each IR detector in the array has a peak, which corresponds to the given wavelength of the incident infrared light. The spectral resolution of the fabricated micro-spectrometer is about 0.6 μm, despite the large dimension of the thermopile-based IR detector (the width of a single detector is 100 μm).
Samenvatting

(summary in Dutch)

Een optisch analyse systeem dat werkt in het infrarode spectrage gebied heeft vele toepassingen. Het is vooral bruikbaar in chemische analyse systemen, sinds de emissie- en absorptielijnen van zowel elk type molecuul zich in het infrarode gedeelte van het elektromagnetisch spectrum bevinden. Miniaturisatie van deze spectrale systemen kunnen leiden tot doorbraken, niet alleen in de wetenschap, maar ook in de industrie. Dit proefschrift beschrijft het ontwerp, de fabricage en karakterisatie van een geminiaturiseerde spectrometer die werkt in het infrarode gebied van het elektromagnetisch spectrum.

Het apparaat beschreven in dit proefschrift is functioneel gesplitst in twee delen. Het eerste gedeelte is de bovenste wafel die het dispergerende element draagt, een tralie met meerdere spletten, gerealiseerd uit de standaard metallisatie stap van een IC proces. Het tweede gedeelte is de onderste wafel, welke de reeks met thermokoppel-gebaseerde infrarood detector bevat, gefabriceerd op een silicium nitride membraan middels bulk micromachining, voor een thermische isolatie van het gevoelige gebied. Het tralie-gebaseerde dispergerende element scheidt de spectrale componenten waaruit de invallende straling bestaat. Het uitgewaaierde spectrum wordt geprojecteerd op de reeks infrarood detectoren na passage van het optische pad door het bulk silicium. Dientengevolge worden er een specifieke of een reeks van golflengten gedetecteerd op een bepaalde infrarood detector.
binnen de reeks.

De hoge mate van weerkatsing van het aluminium voor infrarode golflengtes komt de realisatie van een tralie, op een wijze compatibel met standaard IC fabricage, ten goede. Het bulk silicium wordt gebruikt om het optische pad de definiëren, daar silicium in hoge mate transparant is voor golflengtes boven 1 μm.

De thermokoppel-gebaseerde infrarood detectoren zijn ontworpen op optimale geometrische waarden, gebaseerd op de simulatie-resultaten en de eigenschappen van de gebruikte materialen, wat resulteert in de best mogelijke prestatie in term van de detectivity. Voor de IR detector, welke bijvoorbeeld een absorberend oppervlak heeft van 100×100 μm², is de lengte van de balk 300 μm, wat correspondeert met de lengte van de balk om de maximale detectiviteit te bereiken (zie figuur 4-16).

De waargenomen optische prestatie van de gefabriceerde tralie met meerdere spleten is in goede overeenstemming met de theorie over diffractie. Bovendien maakte meting van de spectrale responsie van de IR micro-spectrometer duidelijk dat elke IR detector in de reeks een piek laat zien, die correspondeert met de gegeven golflengte van het invallende licht. De spectrale resolutie van de micro-spectrometer is ongeveer 0.6 μm, ondanks de grote omvang van de thermokoppel-gebaseerde IR detectoren (de breedte van een enkele detector is 100 μm).
List of publications

JOURNAL PUBLICATIONS


PRESENTATIONS

List of publications


S.H. Kong, D.D.L. Wijngaards and R.F. Wollenbuttel, Fabrication of an Integrated Silicon Infrared Micro-spectrometer, Proceedings MME'00, October 1-3, 2000, Uppsala, Sweden, pp.65-68. (Received ‘The Best Poster Award’)


D.D.L. Wijngaards, S.H. Kong, P.M. Sarro and R.F. Woffenbuttel, Thermophysical Characterisation of PolySi0.7Ge0.3 for Use in Thermoelectric Devices, Transducers '01, June 10-14, 2001, Munich, Germany, pp.1010-1013.

D.D.L. Wijngaards, G. de Graaf, S.H. Kong and R.F. Woffenbuttel,
Temperature profile comparison of various substrate geometries for passive heat sinks, Proceedings MME'01, September 16-18, Cork, Ireland, pp.143-146.
Acknowledgements

This work could not have been completed without the help, support and cooperation of many people: the members of the Laboratory for Electronic Instrumentation, Department of Information Technology and Systems, Delft University of Technology (TU Delft) and the staff of the Delft Institute of Microelectronics and Submicron technology (DIMES). I would like to express my deepest gratitude to them.

In particular I would like to thank:

Prof. J.H. Huijsing, my promotor, for giving me the opportunity to study for the Ph.D. degree at the EI Laboratory and for creating a wonderful working environment where it was a pleasure to perform the required research.

Dr. Reinoud F. Wolffensbuttel, my project supervisor, for his inspiring ideas, continuous support and guidance not only in the research but also in real life. Also his courtesy to allow me all freedom in this research should be mentioned.

Dr. Marian Bartek for his practical guidance during the training course on the Cleanroom Procedures and Silicon Processing organized by the DIMES technology centre.
Acknowledgements

Ger de Graaf for his technical support and helpful cooperation during the thermoelectric and optical measurements of the devices.

Prof. C.I.M. Beenakker for giving me the opportunity to work in the DIMES cleanroom and also for his warm encouragement.

Prof. P. French for his many pieces of scientific advice and fruitful talks about daily life.

Prof. L. Sarro, for her helpful discussions and technical guidance to ensure that the cleanroom processes would be successful during the design and fabrication steps of the devices.

Wim van der Vlist, for his dicing and packaging work of the devices, as well as his tutorial and practical guidance on the CMOS and bipolar processes in the cleanroom of DIMES.

The members of DIMES for giving technical assistance when the devices were fabricated. Bert Goudena, Mario Laros, Hugo Schellevis, Henk van Zeijl, Jan Groeneweg, Cassan Visser, John Slabekoorn and the other staff members of DIMES for their technical assistance and support.

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

My former roommate, Dr. Jose Higino G. Correia at University of Minho, Portugal, for his helpful advice in theoretical and practical fields during his stay in the Netherlands.

My fellow Ph.D. students at the Laboratory for their support, advice and help, and for creating such a great working atmosphere. In particular Edmond Cretu and Davey Wijngaards for sharing their experiences and scientific backgrounds with me. As Dutch is not my language, I wish to thank Davey again for his help in translating the summary.

The Dutch Technology Foundation (STW) that provided the financial support to perform the research (project DEL 55.3733).
Rev. C.K. Lee in the Korean Church of The Hague for his moral support, as well as for his warm prayer and blessing.

Prof. Jong-Sung Kim at An-dong National University, Korea, for his enthusiastic encouragement and advice to do my Ph.D. at TU Delft.

I cannot forget Prof. M. Esashi at Tohoku University, Japan, and Prof. Byoung-Ki Son, Prof. Yong-Hyun Lee, Prof. Si-Young Choi and Prof Jang-Kyu Shin at Kyungpook National University, Korea. I would like to express my appreciation for their continuous encouragement and support.

I also wish to thank my dear friends, Phong M.K.P. Dang, Peter Gallast, Leon Perlot, Cheon-Hwan Rho, Yong-Hwan Seok and Bon-Sam Ku, for their continuous encouragement and helpful discussions.

My parents, brothers and sister have provided the infinitive encouragement from the far-off place, Korea. This work could have never been completed without them.

Finally, I would like to thank my beloved wife, Eun-jeong and son, Jin-seok for their love, support and patience during the past years.
Acknowledgements
About the author

Seong Ho Kong was born in Youngcheon, South Korea in 1967. He received the B.Sc. degree in Electronic Engineering from Kyungpook National University, Korea in 1993 and his M.Sc. degree in Mechatronics and Precision Engineering from Tohoku University, Japan in 1996 with his thesis work concerning the fabrication of reactive ion etching systems for deep silicon etching. In February 1997 he joined the Laboratory for Electronic Instrumentation, Department of Information Technology and Systems, Delft University of Technology, as a Ph.D. student, where he performed research on the application of silicon micromachining for the infrared micro-spectrometer.