Design and fabrication of a surface micromachined positioning device

Delft University Press / 1996
Design and fabrication of a surface micromachined positioning device

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus
Prof. ir. K.F. Wakker,
in het openbaar te verdedigen
ten overstaan van een commissie,
door het College van Dekanen aangewezen,
on maandag 11 november 1996 te 10:30 uur.

door

Johannes Frans Lodewijk GOOSEN
elektrotechnisch ingenieur
geboren te Rotterdam.
Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. S. Middelhoek

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

Samenstelling promotiecommissie:
Rector Magnificus, voorzitter
Prof. dr. ir. S. Middelhoek, promotor
Dr. ir. R.F. Wolffembutte, toegevoegd promotor
Prof. dr. ir. P.P.L. Regtien
Prof. dr. W. Mokwa
Prof. dr. K. Najafi
Dr. ir. P.M. Sarro
Dr. ir. P.J. French

Technische Universiteit Delft
Technische Universiteit Delft
Technische Universiteit Delft
Universiteit Twente
RWTH Aachen
University of Michigan
Technische Universiteit Delft
Technische Universiteit Delft

Published and distributed by:
Delft University Press
Mekelweg 4
2628 CD Delft, The Netherlands
Phone: (+31) 15 2783254
Fax: (+31) 15 2781661

CIP-DATA KONINKLIJE BIBLIOTHEEK, DEN HAAG

Goosen, J.F.L.

Thesis Delft University of Technology. - With ref. - With summary in Dutch.
ISBN 90-407-1390-1
NUGI 841
Subject headings: Microactuator / Surface micromachining / Stress measurement / Electrostatics

Copyright © 1996 by J.F.L. Goosen

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 permission of the publisher: Delft University Press, Mekelweg 4, 2628 CD Delft, The Netherlands.

PRINTED IN THE NETHERLANDS
In memory of my mother.

For the encouragement and support.
On the cover:
The world of micromachining. SEM of a micromotor with the aluminium interconnect surrounding it peeling from the substrate as described in this thesis.
Table of contents

1. Introduction ............................................................................................................ 1
   1.1 Miniaturization of sensors and actuators ......................................................... 1
   1.2 Motivation and Objectives ............................................................................. 3
   1.3 Organisation of this thesis ............................................................................. 4
   1.4 References ...................................................................................................... 5

2. Micromachined actuators ...................................................................................... 7
   2.1 Introduction ..................................................................................................... 7
   2.2 Actuation principles ....................................................................................... 8
      2.2.1 Thermal .................................................................................................. 8
      2.2.2 Mechanical ............................................................................................ 9
      2.2.3 Magnetic ............................................................................................... 10
      2.2.4 Electrostatic .......................................................................................... 10
      2.2.5 Actuation principle evaluation ............................................................... 11
   2.3 Electrostatic drive principles .......................................................................... 11
      2.3.1 Variable capacitance .............................................................................. 11
      2.3.2 Charge relaxation .................................................................................. 13
   2.4 Distributed systems ........................................................................................ 15
   2.5 Actuator design considerations ...................................................................... 16
      2.5.1 Electrical field strength ......................................................................... 16
      2.5.2 Friction and wear ................................................................................... 17
      2.5.3 Material properties ................................................................................ 17
      2.5.4 Design tolerances .................................................................................. 18
   2.6 References ...................................................................................................... 19

3. Micromotion positioning system ........................................................................... 23
   3.1 Introduction ..................................................................................................... 23
   3.2 Transport principle ......................................................................................... 24
   3.3 Actuator design considerations ...................................................................... 26
3.3.1 Step size ............................................................... 26
3.3.2 Drive couple ....................................................... 27
3.3.3 Load couple ....................................................... 29
3.3.4 Hinges .............................................................. 30
3.3.5 Actuator implementation .................................... 31
3.4 Mechanical modelling .............................................. 34
  3.4.1 Mechanical theory ............................................. 34
  3.4.2 Hinges ............................................................ 37
  3.4.3 Plate bending ................................................... 46
  3.4.4 Step size .......................................................... 48
3.5 Electrostatic modelling ............................................ 49
  3.5.1 Capacitance modelling ....................................... 50
  3.5.2 Electrostatic forces .......................................... 53
  3.5.3 Capacitor plate isolation .................................... 58
3.6 Load restrictions ................................................... 62
3.7 Arrays .................................................................. 65
  3.7.1 One-dimensional array ....................................... 65
  3.7.2 Two-dimensional array ....................................... 67
  3.7.3 Multi-dimensional array ..................................... 68
  3.7.4 Partitioned array ............................................... 69
  3.7.5 Array indexing .................................................. 70
3.8 Electronically driven array ........................................ 71
  3.8.1 Addressing circuit ............................................. 72
  3.8.2 Programmable cell logic ..................................... 73
3.9 References .......................................................... 77

4. Micromechanical system fabrication ................................. 79
  4.1 Introduction ......................................................... 79
  4.2 Micromachining technologies ................................... 80
    4.2.1 Bulk micromachining ...................................... 80
    4.2.2 Surface micromachining ................................... 82
    4.2.3 Wafer-to-Wafer bonding ................................... 83
    4.2.4 SOI and Epi micromachining .............................. 84
    4.2.5 Other micromachining techniques ......................... 85
4.3 Requirements on the surface micromachining process .................................. 87
  4.3.1 Surface micromachining applications .............................................. 87
  4.3.2 Structural requirements ..................................................................... 88
  4.3.3 On-chip electronics .......................................................................... 88
4.4 Surface micromachining process design considerations .......................... 89
  4.4.1 Order of processing ........................................................................... 89
  4.4.2 Structural and material choice .......................................................... 90
  4.4.3 Temperature steps ............................................................................ 93
  4.4.4 Low stress mechanical layers ............................................................. 95
4.5 Surface micromachining module ................................................................. 96
  4.5.1 Electronic processing ......................................................................... 97
  4.5.2 Nitride protective layer ....................................................................... 98
  4.5.3 First sacrificial layer .......................................................................... 99
  4.5.4 Contact holes to polysilicon ............................................................... 102
  4.5.5 First polycrystalline silicon ............................................................... 103
  4.5.6 Second sacrificial layer ...................................................................... 105
  4.5.7 Silicon nitride ..................................................................................... 106
  4.5.8 Second polycrystalline silicon ............................................................ 107
  4.5.9 Protection layer .................................................................................. 108
  4.5.10 Contact holes ................................................................................... 110
  4.5.11 Interconnect ....................................................................................... 111
  4.5.12 Oxide removal ................................................................................... 113
  4.5.13 Sintering ........................................................................................... 113
  4.5.14 Wafer dicing ...................................................................................... 114
  4.5.15 Sacrificial etching ............................................................................. 115
  4.5.16 Rinsing and drying ............................................................................ 122
4.6 References .............................................................................................. 126

5. Stress measurement ...................................................................................... 135
  5.1 Introduction ............................................................................................ 135
  5.2 Stress measurement techniques .............................................................. 135
    5.2.1 Resonance frequency technique ......................................................... 136
    5.2.2 Load deflection ................................................................................ 137
    5.2.3 Wafer curvature ................................................................................ 137
Table of contents

5.2.4 Beam curvature ......................................................... 138
5.2.5 Buckling ........................................................................ 139
5.2.6 Displacement technique .................................................. 145
5.3 New buckling structures ...................................................... 147
  5.3.1 Diamond structure ...................................................... 148
  5.3.2 Push-pull structure ...................................................... 151
5.4 Rotating beam structure ..................................................... 155
  5.4.1 Modelling .................................................................... 155
  5.4.2 Finite element analysis .................................................. 157
  5.4.3 Measurement .............................................................. 160
5.5 Conclusions ..................................................................... 164
5.6 References ....................................................................... 165

6. Realization and measurement ............................................... 169
  6.1 Actuator design ............................................................... 169
  6.2 Fabrication of the actuator ................................................ 170
  6.3 Measurements .................................................................. 178
    6.3.1 Testing of the actuators .............................................. 178
    6.3.2 Electrical measurements ........................................... 179

7. Conclusions ...................................................................... 181

Appendix A: Mechanical calculations ...................................... 183

Summary ............................................................................... 189

Samenvatting ......................................................................... 195

Acknowledgements .................................................................. 201

List of publications and presentations ..................................... 203

About the Author .................................................................... 205
1.1 Miniaturization of sensors and actuators.

In the last decades, miniaturisation has been an important aspect in the development of technology. The most extreme example of this is the field of microelectronics, thanks to silicon integrated-circuit (IC) process technology. Smaller components make the systems faster, more reliable, cheaper and capable of incorporating more complex functions. Along with the electronics, the sensors that are a part of many systems, are also miniaturised. To this end, a number of technologies have been developed, that are similar to the silicon integrated-circuit (IC) process technology, for the fabrication of sensors in silicon [1.1].

The introduction of micromachining has been a major factor in this miniaturisation of sensors, making the structures three dimensional and bringing new principles to the designers. The materials used, such as crystalline silicon, polycrystalline silicon and silicon nitride have very good mechanical properties and are very well suited for the fabrication of mechanical structures [1.2]. This has given rise to the miniaturisation of many mechanical sensors, such as pressure and acceleration sensors [1.3]-[1.5]. This integration of sensors in silicon not only improves reliability, speed and in some cases reduces price, it also makes it possible to combine the sensing element with microelectronics on one
single chip to create a so called smart sensors. Small sensor signals can be amplified, conditioned and transformed into a standard output format, without relatively long interconnect wires. Sensors can be combined with self-test, self-calibration and bus interface circuits, simplifying their use and making them more accurate and reliable.

However, the incorporation of electronics on the same chip as the sensor element is not always necessary or even desirable [1.6]. The cost of integrating the electronics on the same chip with the sensors is high, due to the special processing that is needed, making hybrid sensors a more attractive alternative. In a hybrid sensor the electronics are fabricated using a standard process and the sensor fabrication is not hindered in any way by the presence of electronics. Simple bond wires connect the electronics to the sensor, providing the same level of reliability and ease of use as fully integrated smart sensors. Nevertheless, fully integrated systems are needed in some applications. The sensor signals can be very small and the presence of bondpads and bondwires in the signal path can introduce a loss of signal due to the introduction of noise, series resistance and parasitic capacitance. Another configuration that necessitates the incorporation of electronics on the same chip is that of the matrix sensors. These systems consist of an array of sensors that provide information about the distribution of the signal and find application in such fields as optical imaging and tactile sensors. The number of separate sensors in such a systems make hybrid systems impossible due to the excessive number of bond wires needed to connect all sensors to the electronics. In such systems indexcircuitry is essential and full integration of both the sensors and the electronics on the same chip is necessary.

Apart from the sensors fabricated in silicon, actuators have also been fabricated [1.7]-[1.9], although these have not received the attention given to sensors. Many micromachined actuator structures have been fabricated to demonstrate the possibilities of the fabrication processes but had no practical use. In the last decade, many actuation principles have been developed and some practical actuators have been fabricated., often as an integral part of a sensor. However, the down-scaling of existing systems to sub-millimetre scale creates more problems for actuators than it does for sensors. Two application fields can be distinguished when considering actuators. The first is that of the displays. Modulating displays, which in contrast to self generating displays, only switch, direct or regulate an external light source, can be fabricated using
micromachining and are not loaded mechanically. The small micromachined structures can be used as electrically modulated mirrors or filters taking the place of more common LCD systems, in particular in projection displays [1.10]. The large arrays needed to obtain the desired resolution need electronics to be feasible.

In contrast to sensors and displays, the second actuator category; the mechanical actuators such as micromanipulators, micromotors, etc., are designed to perform work and therefore, are mechanically loaded. However, as the micromachined structures are very small and fragile it is difficult for these actuators to influence the macroscopic world. This also presents the problem, that no ready applications exist for these actuators, as they can not take over the function of their larger counterparts the way sensors and displays can. Therefore, new applications must be found. One way to increase the usefulness of these actuators is by using them in so called, distributed systems in which many small actuators combine their effort in both power and range. Some of these large arrays face the same connection problems to the electronics, as do the large sensor or display arrays and integration of electronics on the same chip becomes necessary.

1.2 Motivation and Objectives

The practical application of micromachined actuators is not only difficult due to the scaling of the power output and range, but also due to the inability of surface micromachining to reproduce the many mechanical components which comprise most macroscopic systems. To find ways around this, new methods must be found to incorporate their function using surface micromachined structures. To illustrate that micromachining in silicon can provide useful actuators, a device was designed for the accurate positioning of small objects. It is configured as a distributed systems to increase its power and range. As will be shown in section 3.8, the functionality can be extended by the incorporation of electronics.

To fabricate this device, a surface micromachining process was designed, which is intended as a general surface micromachining process that can be used for the fabrication of a wide range of sensors and actuators. In addition to the mechanical requirements, it had to be compatible with a standard bipolar electronic process to enable designers
Introduction

to use both electronics and mechanical structures on the same chip, to allow the fabrication of large systems.

1.3 Organisation of this thesis

In this thesis, chapter two gives an overview of the design considerations of micro actuators. Several possible drive principles are described with special attentions given to electrical actuation and the added value of, so called, distributed systems is discussed. In addition some design considerations are given for the miniaturisation of actuators.

In chapter three the design of the micromotion positioning device is discussed. The principle is described, performance is analysed and the ensuing boundary conditions and trade-offs examined. The actuator is modelled mechanically an electrostatically and the implementation of the basic actuator is described and the possible devices discussed. Finally the use of electronics on the same chip is discussed and a number of possible circuits described.

In chapter four an overview is given of the possible micromachining technologies and the requirement on the technology for the fabrication of the structure are discussed. Problems, such as residual stress, the creation of stringers and sacrificial etching are examined. The compatibility between microelectronic circuits and surface micromachining are discussed and the final processing sequence is described.

Chapter five is devoted to the measurement of intrinsic stress in thin films. An overview is given of possible stress measurement techniques and three new stress measurement structures are described.

In chapter six the geometry of the actuator is discussed and the fabrication sequence shown. Some measurements of the actuators and the electronic components are presented.

Finally chapter seven presents some conclusions in addition to some suggestions of future work.
1.4 References

Introduction
2.1 Introduction

Apart from the many sensors designed using micromachining, a number of actuators have also been designed. These actuators use a wide variety of fabrication techniques and excitation principles, which vary from simple attracting capacitor plates to complex rotating micromotors. Even though many actuators have been demonstrated, very few have actually been utilized. Most have been used to demonstrate the capabilities of micromachining processes and were never intended to perform any work, which, due to the small size and low power output, is not easy to achieve [2.1].

When designing an actuator in silicon, we cannot make use of any direct actuation property of silicon, as it does not exhibit any electromechanical effects. Therefore, any conversion of an electrical signal into a mechanical action has to be achieved using other materials and/or system properties. A wide variety of actuation principles is available for micromachined actuators.
2.2 Actuation principles

When we consider the driving mechanism for micromechanical actuators, many possibilities present themselves [2.2]. This overview is limited to methods where the power is supplied by an electrical signal. Other hydraulic, pneumatic, radiant, vibration, thermal and other systems have been tried, or would be feasible, but the possibility of using electricity as an information carrier in addition to the power supplied by that signal, provides us with an enhanced flexibility. In addition, the supply of electrical power to the small micromachined systems is much easier than with other forms of power.

The actuation principles used to actually induce the motion, still varies greatly, as the electrical power can be converted into other forms of energy in the actuator. Sometimes it is possible to distinguish more than one form of energy, in which case the actuation method is mentioned under the heading of the energy domain that induces the motion.

2.2.1 Thermal

Thermal actuation is used often and has been in use for a long time [2.3]. This is done by simply sending a current through a resistance, which then heats up. The most commonly used effect is the expansion of a material at a higher temperature.

By using beams constructed of two materials with different expansion coefficients, the small change in size due to the raised temperature is amplified [2.4]. This is the same mechanism as used in bi-metals in electric irons, toasters, etc. Extensive movements have been realised by the use of organic films such as polyimide, as these are more pliable than metals or ceramics and therefore easier to deform.

The disadvantage of thermal actuation is the use of special materials and the limited switching speed, which is determined by the heat conduction of the materials and the cooling by its surrounding and the small forces that are created. The advantage is that actuation is very simple and movement/deflection is large and can be controlled continuously over the whole range by controlling the temperature.

Examples are self assembling boxes [2.5] and the positioning device in which the 'legs' consist of an organic film sandwich with a chrome heater element [2.6].
2.2.2 Mechanical

Mechanical deformation can be induced by heating a gas which expands, the gas pressure increases, which moves a membrane [2.7]. Heating takes place with a simple resistor. In some actuators, such as inkt-jet nozzles, a liquid is heated above its boiling point and enters the gas phase, creating large pressure differences. Other methods include piezo-electric materials that deform when an electrical field is applied [2.8],[2.9] or the use of shape-memory alloys [2.10].

The advantages and disadvantages of expanding mediums is the same as for direct thermal actuation, in that it is easy to construct, pressure and therefore, motion can be controlled continuously, large forces can be obtained, but speed is limited by the heat capacitance of the gas and surrounding.

For piezo-electric materials movement is limited and special materials are needed, since silicon does not have piezo-electric properties. The amplitude of the motion is very small, however, it is easy to control and the forces involved are large.
2.2.3 Magnetic

Magnetic fields are created to attract or repel magnetic materials in order to induce movement in structures [2.11],[2.12]. Although it is the most commonly used method in large electric actuators, such as electromotors, for integrated actuators some difficulty arises in miniaturization. To attain the desired magnetic field strength, large coils must be constructed, which are very difficult to make in planar technologies, special materials are needed for cores and large currents are needed. However, large field strengths are potentially possible.

![Magnetic field created by planar coil.](image)

Fig. 2-3  Magnetic field created by planar coil.

2.2.4 Electrostatic

Electrostatic fields can easily be formed between conducting surfaces. When this field is present, forces act on the surfaces and they deform or move [2.13],[2.14].

![Excitation using electrostatic attraction between two capacitor plates.](image)

Fig. 2-4  Excitation using electrostatic attraction between two capacitor plates.

No special materials are needed and power is only dissipated dynamically. In static operation no power is used. The forces involved are determined by the electrostatic field strength which is limited by the breakdown of the medium. In air the breakdown field strength is approximately 3-$10^6$ V/m. However, when the gap between the plate
becomes smaller, the breakdown field strength increases to values up to $3 \times 10^8$ V/m (see section 2.5.1). This effect makes it possible to achieve considerable forces without using special materials or having to operate in a vacuum. However, only attracting forces can be applied, limiting the use in certain applications.

### 2.2.5 Actuation principle evaluation

When considering the possible actuation principle for an actuator, other factors must be taken into account. The actuator must be fabricated in a micromachining process in silicon. To preserve compatibility with standard microelectronic processing and processing facilities, no ‘exotic’ materials can be used. This rules out the use of piezo-electricity and magnetic fields. As the technology is also incapable of constructing enclosed spaces in which a gas pressure could be changed, other mechanical excitation principles are unsuited. This leaves electrostatic and thermal bi-morph actuation. Due to the speed restrictions of thermal excitation, the small forces, high energy consumption and the possible harm to objects in contact with the hot actuator, the electrostatic actuation principle is chosen.

### 2.3 Electrostatic drive principles

When using an electrostatic field to drive a microactuator, many possible configurations are possible [2.15]. Although the electrostatic principle is used in all these actuation methods, two different drive principle can be distinguished. The one most commonly used, is known as the variable capacitance drive in which the electrostatic field between the capacitor plates produces a static force on these capacitor plates. The other is the charge relaxation drive. This system uses the phase difference caused by the finite charge relaxation speed in a travelling field.

#### 2.3.1 Variable capacitance

The variable capacitance actuation principle is the most commonly used method of electrostatic actuation and is extensively investigated
[2.16]-[2.18]. Two conducting parallel plates separated by an insulating layer or gap form a capacitance with a value given by:

$$C = \varepsilon_r \varepsilon_o \frac{wl}{d}.$$  \hspace{1cm} (2-1)

where $C$ is the capacitance, $w$ is the width of the plates, $l$ the length of the plate and $d$ the separation between the plates. $\varepsilon_o$ and $\varepsilon_r$ are the free space and relative permittivities of the gap (see Figure 2-5).

Fig. 2-5  Geometry of parallel capacitor plates.

If the voltage $V$ is applied across these two plates, the potential energy, $U$, of the capacitor is:

$$U = \frac{1}{2} CV^2 = \frac{\varepsilon_r \varepsilon_o wlV^2}{2d}.$$  \hspace{1cm} (2-2)

The force in any direction is the negative partial derivative of the potential energy in each direction. To calculate the force along the plates in the direction of $w$, assume the plates are offset so they overlap by $x$, as shown in Figure 2-6.

![Figure 2-6](image)

Fig. 2-6  Lateral electrostatic force due to offset in the $w$-direction.
Electrostatic drive principles

The potential energy is now:

\[
U = \frac{\varepsilon \varepsilon_0 x l V^2}{2d}
\]  \hspace{1cm} (2-3)

The force in the direction of \( x \) is given by:

\[
F_w = \frac{\partial U}{\partial x} = \frac{\varepsilon \varepsilon_0 l V^2}{2d}
\]  \hspace{1cm} (2-4)

The forces in the \( l \) and \( d \) direction can be calculated in a similar manner.

![Diagram](image)

**Fig. 2-7** Electrostatic forces and overlap between two parallel plates.

\[
F_l = \frac{\varepsilon \varepsilon_0 w l V^2}{2d}
\]  \hspace{1cm} (2-5)

\[
F_d = \frac{\varepsilon \varepsilon_0 w l V^2}{2d^2}
\]  \hspace{1cm} (2-6)

The above calculations use a parallel plate capacitor assumption to calculate the potential energy. Any fringe fields are not included. However, when the plates are moved, the potential energy stored in the fringe fields remains nearly constant and therefore, adds very little to the overall electrostatic force.

2.3.2 Charge relaxation

Charge relaxation makes use of a dynamic effect. When a conductor is placed in an electrical field, charges are induced. When there is a potential difference along the surface of the conductor, the induced charges will move along the surface of the conductor towards the highest
potential until the accumulation of charges offsets the potential difference. However, the speed of the charges is finite due to the relaxation in the conductor and the gap. The resulting displacement between the charge and the potential difference introduces a force acting on the conductor. When the conductor is free to move, the flux of charges will push it along the potential gradient. If the field changes constantly, the charges can not ‘overtake’ the potential gradient and a constant force will result.

This force is determined by the number of charges and how far they lag behind the field. Very good conductors have a large number of induced charges, but very short relaxation times causing only a small displacement between the field and the charge. Poor conductors have a large displacement but few charges. As both a large number of charges and a large displacement is needed to achieve a large force, an optimum must be found for the material properties for a particular geometry and excitation.

A constant force can only be generated when the charges have unlimited movement in the direction of the field variation. It can therefore only be used to act on continuous loops, which results in a static torque.

Although the basic principle behind this force is clear, it is not easy to model these forces. A model for a micromotor driven using this principle is proposed by S.F. Bart and J.H. Lang [2.19]. The torque produced by such a micromotor is given by:

\[
\tau = \pi \int_{R_i}^{R_o} \Re \{ \varepsilon_0 \hat{E}_0 \hat{E}_z^* - \varepsilon_1 \hat{E}_0 \hat{E}_z^* \} r^2 dr
\]

(2-7)

where \( \tau \) is the torque, \( R_i \) and \( R_o \) are the inner and outer radii of the rotor, \( \varepsilon_0 \) is the permittivity of the upper gap, \( \varepsilon_1 \) of the lower gap and \( r \) is a dummy variable. \( E \) is the electric field with a complex amplitude. The subscript \( \theta \) denotes the field in the direction of the motor movement and the subscript \( z \) the field perpendicular to the rotor movement. The superscript denotes which surface of the rotor the electric field applies to. \( u \) denotes the upper surface of the rotor and \( l \) the lower surface. The superscript * denotes the complex conjugation and \( \Re \{ \ldots \} \) the real part of its argument. The variables are shown in Figure 2-8.
Because all electrical quantities depend on $r$, numerical integration is usually required. As it is not possible to achieve a static force, but only a static torque, this excitation principle is only useful for a limited number of application.

2.4 Distributed systems

A major disadvantage of all micromechanical actuators is that they are limited in the amount of work they can deliver, due to limited forces and limited ranges over which they can move. Nonetheless, power densities are large (i.e. the amount of force created might not be very large, but it is acting on a small area).

To overcome these problems, it was proposed that several micro actuators could be put in parallel to increase the power and put in series to increase the range. This would create larger systems that would be strong enough to manipulate macroscopic objects. These systems are known as distributed systems [2.20]-[2.23].

Another advantage of most distributed systems is the redundancy of the actuators. In most configurations the work of one actuator can be taken over by one or more similar actuators in the same system, so breakdown of an actuator does not result in a failure of the whole system.

Although this principle cannot be used in all cases, it does provide a new possibility for the micromachined actuator. Distributed systems consist of large arrays of small actuators, either interconnected or independent.
Micromachined actuators

This necessitates batch processing of the individual actuators, something that monolithic micromachining technologies are well suited for.

2.5 Actuator design considerations

When an micromechanical actuator is designed, special issues must be taken into consideration. Some of the design features that are found in macroscopic actuators, like electromotors, cannot be used. The scaling of the design influences the possibilities. Generally, in macroscopic actuators the bulk effects are dominant, in microscopic actuators, surface effects dominate the behaviour. One aspect already mentioned is the increase of the breakdown field strength of air when the gap between the conducting plates is decreased. Other considerations are the friction and wear of surfaces, mechanical properties of materials and tolerances in the design.

2.5.1 Electrical field strength

Many microactuators are driven electrostatically. This is in contrast to the macroscopic actuators, which are almost exclusively electromagnetic. This is caused by the increase in the breakdown field strength of the air when very thin layers of air are involved. This effect occurs as the amount of gas in the gap decreases. Fewer ionising gas collisions take place which are needed to sustain electrical breakdown across the gap. Due to this effect the energy density that can be reached in microscopic actuators using electrostatics is comparable to the maximum obtainable in magnetic actuators.

However, breakdown may occur at lower field strengths due to local increases in the field strength. This can occur at corners, can be due to surface irregularities or due to surface roughness. Variations in the processing may cause gaps to be smaller than anticipated, and mechanical stresses can cause the structures to deform, reducing the gap in some places. As these effects can not always be predicted, a safe margin must be maintained. Especially because the breakdown will not only damage the system, at these microscopic dimensions it will be completely destroyed.
2.5.2 Friction and wear

Another concern causing problems in the application of micromechanical actuators, is the friction between surfaces and the resulting wear. Studies have shown that friction between microscopic surfaces is similar to that of larger surfaces [2.24]-[2.27]. However, initial starting friction, the force that must be overcome to start the motion from rest, is larger and more complicated than that of macroscopic surfaces. The microscopic surfaces are usually rather rough in comparison to their lateral dimensions, in comparison to larger surfaces. This will make more irregularities 'grip' each other, increasing the starting force. In addition, the surfaces are very flat, causing a larger portion of the surface to bind this way. Also, an effect called sticking can occur (see also section 4.5.16.1). This is caused by surface charges, hydrogen bridging, solid and liquid bridging and Van der Waals forces [2.28]-[2.30]. To complicate matters further, lubrication of touching surfaces is difficult if not impossible. Normal lubricants do not function properly in these small gaps and at these high speeds.

To overcome these problems, contact must avoided, or at least minimised. This can be done by the use of ridges, ledges and bumps, known as bushings [2.26],[2.31]. Through special designs, air blown between the surfaces can act as a bearing while the actuator moves [2.25], or the actuation force can be used to break contact during movement [2.24],[2.32]. Other materials have also been studied [2.27] and from these it follows that the friction-coefficient of silicon nitride is lower than that of polycrystalline silicon. Even materials such as diamond-like carbon and silicon carbide films have been tested to reduce the friction-coefficient and resulting wear [2.33]. Instead of a sliding motion, a rolling motion will reduce both friction and wear in some cases [2.34]-[2.36].

Many of these methods can be combined to minimize friction and wear, but special consideration must be given to this aspect.

2.5.3 Material properties

Apart from its electrical properties, silicon has excellent mechanical properties [2.37]-[2.39]. The material is very stable and shows no plastic behaviour at room temperature. Therefore, it shows no drift and hysteresis. Silicon is has a high yield strength, high Young's modulus and is light (as strong as steel and as light as aluminium), although this
Micromachined actuators

strength dependents on the crystal orientation. However, silicon is brittle. It breaks when the strain becomes to high, and the crystalline structures will determine the direction of the break. Furthermore, silicon is a good thermal conductor, has a low but temperature dependent expansion coefficient and is also piezoresistive.

Other silicon based materials used in micromachining, like polycrystalline silicon, silicon nitride and silicon oxide, have similar mechanical characteristics [2.40]-[2.42], although these are strongly dependent on the deposition conditions and resulting internal structure. Silicon nitride is a good insulator and has a Young’s modulus, twice that of silicon.

Although silicon, polysilicon and silicon nitride, seem almost ideal materials for the construction of mechanical devices, differences in thermal expansion coefficients between these materials can cause large internal strains, due to the high deposition temperature (see also section 4.4.5).

2.5.4 Design tolerances

The shape of the structures is determined by several practical constraints. First of all the photolithography that is used to determine the lateral dimensions and shapes. This is a very accurate technique, with a resolution of 0.1 to 1 μm depending on the equipment available. Due to processing, the dimensions can change up to 0.5 μm on all sides, although this will be the same for all edges of the same layer on the same chip.

Secondly the vertical dimensions are determined by deposition of the material. Depending on the deposition technique the thickness is within 10% of the desired thickness and will not vary more than 1% over a whole chip. The thickness of the layers will vary between wafers and different fabrication runs.

These tolerances are much larger than those of fabrication technologies for macroscopic devices and they limit what can be done. Bearings made with these tolerances will not function properly. The axle would have too much freedom of movement. To overcome these problems we can mimic nature and use bending beam and plates, instead of hinges as a more dependable and robust means of anchoring moving parts [2.43].
2.6 References


Micromachined actuators
3.1 Introduction

Of the different types of micromachined actuators presented in the last chapter, distributed systems offer unique features that can be used to increase the applicability of micromachined systems.

One possible distributed system which is presented here, consists of an array of small levers that can move an object resting on the device. This type of system was first proposed by Fujita et al. [3.1] inspired by

Fig. 3-1  Caliary motion of hairs in a biological motion system
the ciliary motion of hairs in nature. This biological system is used for the transportation of particles (e.g. the clearing of the air pipe) and as a propulsion method for single-cell organisms in liquids. The individual hairs oscillate with a phase difference to bordering hairs, creating a wave like motion as shown in Figure 3-1. A simplified version can be implemented as a micro electro mechanical system (MEMS), which can be used to transport and position small objects.

3.2 Transport principle

Each actuator in the transport system consists of a lever or arm that can be rotated around a fixed point by the application of an electrical signal, similar to the hairs in a ciliary motion system. By sequentially driving actuators that rotate in opposite directions, a similar drive motion can be created. Basically changing the phase difference between bordering 'hairs' to 90 degrees. These actuators are placed in a large array to combine their strength and range, and to pass the object from actuator to actuator.

Figure 3-2 shows a cross section of the array to illustrate the operation of the system. Figure 3-3 shows the drive signals for three drive cycles. Figure 3-2 a, shows the system in it's starting position, with an object resting on top of the actuators. In the first phase of the drive sequence, all actuators with the same orientation are rotated and the object is lifted (see Figure 3-2 b). Because of the rotation around a hinge,
Fig. 3.3 Electrical drive signals for the actuators showing three movement cycles.

The object will not only be lifted, but also shifted sideways over a small distance as is shown by the dashed line that indicates the original position of the object. Subsequently, the other levers are rotated upwards to take over support of the object (Figure 3-2 c) and the first set of actuators is rotated again and return to their original position (Figure 3-2 d). The object is still lifted and slightly shifted, but is now resting on the other set of levers, which rotate in the opposite direction to the first set. When these are rotated back, the object is lowered and moved sideways again in the same direction and over the same distance as during the lifting of the object (Figure 3-2 e). The system has returned to its original state and the object is shifted sideways. This cycle can be repeated, lifting and lowering the object, and moving it with every step.

By reversing the switching sequence, the object can similarly be moved in the opposite direction, without the need for any additional hardware.

The positioning resolution possible with this system, is determined by the displacement of the object after completion of one cycle as described above. The distance the object can be moved is only limited by the overall size of the array. As long as the object is supported by actuators, it can be moved. The size of the object is also of little importance. As will be shown in section 3.3.3, the size of the object does have some relevance towards the maximum load the system can handle, as this determines the number of actuators the object rests on, but larger objects are not more difficult to transport. The overall movement speed over the array is determined by the time needed for one cycle. This is not only determined by the system itself, but also by the object placed on top of it. At a certain vertical beam motion, the object will lag behind because of the limited acceleration provided by gravity and the air damping, causing the object not to return in time to the rest position for the next
cycle. As only a fixed horizontal displacement is made after each vertical cycle, the lateral speed is limited and will drop at higher operating frequencies.

3.3 Actuator design considerations

The basic actuator needed to execute the required action is a lever that is strong enough to lift the load. To determine the best design, several aspects of the actuator have to be considered. These are determined by; the technology used to fabricate the devices, the specific application of the system and by the chosen geometry.

The system described here is fabricated using surface micromachining technology [3.2]. This technology is similar to the microelectronic fabrication technology and utilizes the deposition of thin films for the construction of the device. The materials used have excellent mechanical properties and very well documented electrical characteristics (see also section 2.5). Actuators can be fabricated that are small in size (generally between 10 µm and 1 mm), but have a high energy density (in the order of 10^5 J/m^3). This combination makes it possible to have small steps combined with short cycle times, resulting in a system with high positioning precision and reasonable transport speed. The small size of the actuator also limits the drive voltage needed for the actuator, so it can be driven using electronic circuits fabricated in conventional standard processes, without the added complexity of high voltage electronics. Furthermore, the incorporation of electronics with the actuators can make the system more user-friendly and incorporate complex functions into the system. These electronics must be incorporated on the same chip due to the huge number of actuators.

3.3.1 Step size

One important aspect of the actuator is the distance the object, placed on top of it, is moved in one cycle or ‘step’. This step size determines two important aspects of the complete system. First the resolution of the movement. The movement consists of a discrete number of steps, so when assuming no slip to occur, the overall movement is a multiple of the step size, resulting in a resolution of one step. Second, the overall movement speed of the object over the array that can be achieved with
the system is determined by the number of steps that can be made within in a given time interval and the size of these steps.

Therefore, whatever the design, a trade-off is needed between resolution and speed of movement. This has to be considered for each application, and could result in the combination of accurate (small steps) and fast (large steps) areas in one array.

Fig. 3-4  Basic concept of actuator in movement array.

Figure 3-4 shows a schematic representation of the most basic form of the actuator. The movement along the surface of the device is given by:

$$\Delta x = L [\cos(\theta + \varphi) - \cos(\theta)]$$  \hspace{1cm} (3-1)

As can be seen from Eq. (3-1), the step size is linearly dependent on the length of the arm, \( L \), is dependent on the angle of rotation, \( \varphi \), but is also strongly dependent on the initial angle, \( \theta \). The displacement, \( \Delta x \), for a certain angle of rotation \( \varphi \), is at a maximum for a given length \( L \), if the angle \( \theta + \varphi \) equals \( \frac{\pi}{2} \).

3.3.2 Drive couple

As has been shown earlier (see section 2.2 and 2.3), electrostatic actuation is an attractive principle to use for actuation when the gap size is small enough to take advantage of the rise in the breakdown field-strength of air. It can easily be implemented in a surface micromachining process, without the use of materials not commonly used in silicon technology, such as the ferromagnetic materials needed for magnetic actuation and still achieve similar energy densities.

To rotate a plate or beam in order to function as the needed actuator, a couple has to be applied. When using electrostatic forces, only attracting
forces between capacitance plates are possible. This can be transformed into the rotation needed by making use of a hinge. Two possibilities are shown in Figure 3-5. From this figure the equation for the couple can be derived. The electrostatic force is presented in the figure by a force $F_e$. From this follows that the couple on the lever that is used to lift the load during transportation, due to the electrostatic force is given by:

$$M_e = F_e \cdot L_2.$$  

When assuming an electrostatic force resulting from a charged parallel plate capacitance as described in section 2.3.1 [3,3], the couple, $M_e$, is given by:

$$M_e = \frac{1}{2} \cdot \varepsilon \cdot \frac{V^2}{d^2} \cdot A \cdot L_2,$$  

(3-2)

where $\varepsilon$ is the dielectric constant, $V$ is the applied voltage, $d$ is the distance between the plates, $A$ is the area of the plates and the other variables are as indicated in Figure 3-5.

The parallel plate assumption is valid in most designs when determining the electrostatic force, as will be shown in a more thorough evaluation of the electrostatic forces involved, included in section 3.5.

As can be seen from equation Eq. (3-2), this couple can be maximized by maximizing; the area $A$, distance $L_2$, voltage $V$ and minimizing distance and $d$. When the lever rotates, the plates will come closer together, reducing $d$ and increasing the couple $M_e$. 

Fig. 3-5 Possible electrode drive configurations for electrostatic rotation of the plate.
3.3.3 Load couple

Another important design consideration is the size and weight of the object that can be transported. This load presses down onto the actuators and introduces a couple around the hinge that must be overcome by the drive couple of the devices when the object is to be lifted for transportation as shown in Figure 3-6.

Fig. 3-6 Load couple as a result of pressure \( P \) from the object weight.

Because we are considering an array of actuators, it is best to describe this load in terms of pressure. The weight of the object is distributed over an area that presses down on the actuators. When the object occupies more area on the array, it will automatically rest on more actuators, distributing the weight over these actuators. This means that the couple on the individual actuators is proportional to the pressure on the device, not to the weight of the object. This also means that the number of objects that can be placed on the array is not limited. As long as there is space available on the array, objects can be transported unless the maximum pressure is exceeded.

The couple around the hinge point caused by the load on an individual actuator is dependent on the angle of rotation, \( \alpha \), with \( 0 < \alpha < \varphi \) and is given by:

\[
M_f = L_f A_d P \cos(\theta + \alpha)
\]  

(3.3)

where \( M_f \) is the load couple, \( A_d \) is the area taken up by one actuator including all the interconnect, spacing needed and all the actuators not presently actuated. \( P \) is the pressure on the device due to the load and the other variables are as indicated in Figure 3-6.

The load couple is at a minimum when \( L_f \) is at a minimum and the area \( A_d \) is at a minimum. This load couple is the primary couple that must
Micromotion positioning system

be overcome by the actuation, but as will be shown later, not the only one. All the couples acting on the lever must be balanced for it to function properly.

3.3.4 Hinges

To support the lever and make rotation possible, a hinge must be connected to the free moving arm. It serves to anchor the lever on the substrate and to provide the one degree of freedom needed for the rotation.

The hinge must satisfy the following demands;

- It must provide smooth rotation without large opposing forces.
- The lever arms should be held firmly in place with a very small tolerance.
- The hinge should be stiff enough to resist bending due to large loads from either the object to be transported, or the actuation force.
- The complete hinge/connection must include a spring force to return the arm to its original position, when no actuation force is present.
- It must provide good electrical contact to the rotating arm in order to be able to drive the actuator electrically.

The possible hinge constructions based on thin film technology make use of the bending or twisting of beams. Although complex axis-in-bearing

![Possible hinge configurations in a thin film process. a; simple beam hinge, b; double beam hinge and c; torsion beam hinge.](image)

Fig. 3-7
constructions are possible, the tolerances are high, electrical contact questionable and there is no spring force (see also section 2.5.4).

Possible hinges are shown in Figure 3-7. The complete mechanical analysis of these hinges is presented in section 3.4.2.

### 3.3.5 Actuator implementation.

When designing the actuator, i.e. choosing the geometry $L_1$, $L_2$, $d$ and $\varphi$, to determine the drive couple and the step size, several aspects must be considered. As shown in section 3.3.1 the step size, $\Delta x$, is determined by the length of the arm, the angle of rotation and the initial angle. When the drive couple, $M_e$, is considered, it is obvious from Eq. (3-1) and Eq. (3-2) that these two aspects of the actuator have conflicting dimensional requirements. As can be seen from Figure 3-8, the geometric values for $L_2$ and $d$, and the maximum angle $\varphi$ over which the lever can rotate are related through: $\varphi = \sin(d/L_2)$. From this follows that the larger the electrostatic force (large $L_2$ and small $d$), the smaller the step distance, $\Delta x$. Therefore, a trade-off needs to be made between the drive couple and the step size.

![Fig. 3-8 Dependence of the maximum rotation angle $\varphi$, on the geometry of the actuation electrodes.](image)

Another factor that must be considered when determining the maximum load and the step size, is the load couple $M_l$. As can be seen from Eq. (3-3), this is set with the selected step size. To keep the load couple at a minimum for a maximum step the configuration in Figure 3-8 a is the preferred one.

However, one parameter influencing the step size has not been used. This is the initial angle $\theta$, between the hinge point and the point on which the object rests. This initial angle can be used to increase the step size, without influencing either the drive couple or the load couple. When
Micromotion positioning system

Fig. 3-9  *Increase of the initial angle in a thin film process, by the use of a spacer.*

using thin film technology, all deposited films are parallel to the substrate, but the initial angle can be increased by placing a bump underneath the side of the plate that supports the object as shown in Figure 3-9.

The height of the bump is limited by the process used. In the fabrication process, described in chapter 4, the bump is formed using the first sacrificial layer and the first polysilicon layer, which brings the height of the bump up to about 700 nm. The use of polysilicon also provides us with the possibility to use this as an electrode to screen-off the passive side of the plate from electrostatic fields and thus prevents any electrostatic forces that could develop between plate and substrate. When this plate is at the same potential as the moving plate, no forces can act on the passive side of the plate to reduce the actuation couple.

Fig. 3-10  *Step-size increase as a function of the lever length L for three initial angles.*
With such a limited choice in height to increase the initial angle, the optimal length of the arm for maximizing the step size must be determined. Figure 3-10 shows the graph of the step size, due to a rotation over 0.1 rad. as a function of \( L \) for three initial angles. As can be seen from this graph, the length of the arm; \( L \) has the largest influence, but the initial angle does add substantially to the step size.

However, if this initial angle is used, the weight of the object will press on the contact point, causing the plate to rotate as shown in Figure 3-11. This will cause the electrostatic plates to separate, reducing the electrostatic force.

Fig. 3-11 Rotation of the plate due to asymmetrical pressure from the load.

This can be solved by bending the electrostatic plates upward as well. The object then rests on two opposite sides of the hinge, so the plate will not rotate. By bending both the lower and the upper electrode, the electrostatic force is not reduced. The shaping of the electrodes can be done using the same sacrificial layer as was used to form the bump that increases the initial angle, and the device becomes virtually symmetrical. The general shape of the actuator is shown in Figure 3-12.

Fig. 3-12 Electrode bending to avoid rotation of the actuator and separation of electrodes.
3.4 Mechanical modelling

Now that the overall shape of the actuator has been chosen, a more elaborate model of the actuator can be constructed. Additionally, the hinge structure must be chosen from the mechanical analysis.

3.4.1 Mechanical theory

Before analysing the specific mechanics of the actuation plate and possible hinge structures, a short discussion of the general theory of the deformation of beams due to forces, couples and other influences on the mechanical behaviour, like residual stress in the material, is given.

3.4.1.1 Deflection of beams

When a beam is subjected to a transverse load, as shown in Figure 3-13,

\[ \kappa = \frac{1}{\rho} = \frac{d\theta}{ds} = \frac{d^2v}{dx^2} \quad (3-4) \]

where \( \rho \) is the radius of the curvature, \( \theta \) is the slope of the curve and \( s \) is the length along the curve. This approximation is valid when the deflections are small, in which case; \( \theta \approx \tan \theta = dv/dx \), and \( ds \approx dx \).

Fig. 3-13 Deflection curve of a beam subjected to a transverse load.
If the material of the beam is linearly elastic and follows Hook’s law, which is the case for the materials used here (see section 2.5.3), the curvature is:

$$\kappa = \frac{1}{\rho} = \frac{M}{EI_y} \quad (3-5)$$

in which $M$ is the bending moment, $E$ is the Young’s modulus and $I_y$ is the second area momentu with respect to the y axis (the direction of the deflection)

From Eq. (3-4) and (3-5) it follows that:

$$\frac{d\theta}{dx} = \frac{d^2\nu}{dx^2} = \frac{M}{EI_y} \quad (3-6)$$

Because the bending momentum at every point in the beam is determined solely by the loads on the beam, the expression for the deflection curve $\nu(x)$, can be derived.

The second area momentum is determined by the shape of the cross section:

$$I_y = \int_{y^2}dA = \iint_{z^y}y^2dydz \quad (3-7)$$

For a rectangular beam with width $w$ and height $h$, the second area momentum is given by:

$$I_y = \int_{z^y}y^2dydz = \frac{wh^3}{12} \quad (3-8)$$

### 3.4.1.2 Torsion of beams

When a beam is subjected to a twisting couple, $M_t$, the cross-section of the beam will rotate around the longitudinal axis (see Figure 3-14).
Micromotion positioning system

![Image of a beam](image)

Fig. 3-14  *Twisting of a beam.*

The twist angle per unit length $d\phi/dx$, is given by:

$$
\frac{d\phi}{dx} = \frac{M_I}{G I_p},
$$

(3-9)

where $G$ is the shear modulus of elasticity and $I_p$ is the polar moment of inertia.

This polar moment of inertia is determined by the cross-section of the beam and is given by:

$$
I_p = \int_0^b r^2 dA = I_y + I_z,
$$

(3-10)

where $r$ is the distance from the longitudinal axis of area $dA$.

For a rectangular beam with width $w$ and height $h$, this becomes:

$$
I_p = I_y + I_z = \frac{wh^3}{12} + \frac{hw^3}{12} = \frac{wh}{12}(h^2 + w^2)
$$

(3-11)

The shear modulus of elasticity, $G$, and the Young’s modulus are related according to:

$$
G = \frac{E}{2(1 + \nu)},
$$

(3-12)

where $\nu$ is the Poisson’s ratio.
3.4.2 Hinges

As shown in section 3.3.4 the hinges used, utilize the bending and/or twisting of beams. The beam will not only support the plate, but also act as a spring to return the plate to its original position when no other couples act on it. This does, however, cause an additional couple to act on the plate due to the rotation, which has to be overcome by the drive couple.

Apart from the drive couple, $M_e$, that causes the plate to rotate, three other loads act on the hinge. These are the shear force caused by the actuation, $F_e$ and the shear force, $F_l$, and couple, $M_l$, due to the load, as shown in Figure 3-15. When analysing the specific hinge configurations, only half of the symmetrical structure is modelled and the loads indicated in the figures are total force $F_0$ consisting of half the sum of the drive force $F_e$ and the load force $F_l$, and the total couple $M_0$ consisting of half of the sum of drive couple $M_e$ and load couple $M_l$.

![Fig. 3-15 Loads acting on the hinge. The electrostatic force, $F_e$, the electrostatic couple, $M_e$, the load force, $F_l$, and the load couple, $M_l$.](image)

3.4.2.1 Simple beam hinge

The simple beam hinge consists of two simple beams along the side of the plate, which is connected at the end of the beams, as is shown in Figure 3-7 a. Figure 3-16 gives a schematic representation. The plate,
which serves as the electrostatic drive plate on one side and lever on the other, is not shown but is connected at point B and is represented by the loads, $F_0$ and $M_0$, it puts on the hinge beam. Due to these loads the beam will deform as shown.

Fig. 3-17 Free body diagram of a beam for the calculation of the reaction forces.

Because the hinge is free on one end, no internal stress along the longitudinal axis will develop due to the bending of the beam or due to residual stress caused by the processing. Figure 3-17 shows a part of the beam which is in equilibrium, thus the sum of all forces and moments must be zero. This results in:

$$M = R_A x - M_A,$$

(3-13)

where $R_A$ and $M_A$ are the reaction force and moment of the beam at point A. Using the boundary conditions; $v_s(0) = 0$ and $dv_s/dx|_{x=0} = 0$, the deflection curve can be found by integrating Eq. (3-5):

$$v_s(x) = -\frac{1}{2} M_A \cdot \frac{1}{E I_y} \cdot x^2 - \frac{1}{6} R_A \cdot \frac{1}{E I_y} \cdot x^3$$

(3-14)

By substituting the reaction force and moment $R_A$ and $M_A$ the deflection can be determined. The reactions can easily be found using the static equilibrium; $R_A = F_0$, and $M_A = F_0 L - M_0$. This results in:

$$v_s(x) = -\frac{1}{2} \frac{F_0 L - M_0}{E I_y} \cdot \frac{1}{2} \frac{F_0}{E I_y} \cdot x^2 - \frac{1}{6} \frac{F_0}{E I_y} \cdot x^3$$

$$= -\frac{1}{2} \frac{M_0}{2 E I_y} \cdot x^2 - \frac{1}{2} \frac{F_0 L}{2 E I_y} \cdot x^2 + \frac{1}{6} \frac{F_0}{E I_y} \cdot x^3$$

(3-15)
From this we can easily determine both the deflection of the simple beam hinge at the tip, $\delta_s$, and the angle, $\theta_s$ over which the plate will rotate.

$$\delta_s = v_s(L) = \frac{1}{3} \frac{L^3}{EI_y} \cdot F_0 - \frac{1}{2} \frac{L^2}{EI_y} \cdot M_0$$  \hspace{1cm} (3-16)$$

$$\theta_s = \frac{dv_s}{dx}_{x=L} = \frac{1}{2} \frac{L^2}{EI_y} \cdot F_0 - \frac{L}{EI_y} \cdot M_0$$  \hspace{1cm} (3-17)$$

As can be seen from Eq. (3-16) and (3-17) the beam will rotate not only due to the load couple $M_0$, but also due to the load force $F_0$.

3.4.2.2 Double supported beam hinge

Another possible hinge consists of two double supported beams, along the side of the plate, with the plate connected in the middle at point C (see chapter Fig. 3-18). The loads it places on the hinge are indicated again as $F_0$ and $M_0$.

![Diagram of the double beam hinge](image)

**Fig. 3-18** Deflection of the double beam hinge under loads $F_0$ and $M_0$

This double supported beam is statically indeterminate [3.4] and the reactions can not immediately be obtained by solving the equilibrium equations. To circumvent this, we make use of the superposition method. When Hook's law is valid, the deflections are linear with the forces and couples that cause them. Therefore, the total deflection of the sum of two loads equals the sum of the deflections of each load.

By removing the statical redundants the problem becomes statically determined and can be solved in a similar manner as with the simple hinge beam in the previous section. The deflections caused by the redundants is then calculated separately on the same statically determined beam. The boundary conditions, which specify the deflection and angle at the beam ends, can then be used to calculate the redundant reaction forces and couples, after which all reactions can be calculated.
In the double supported beam hinge, the reaction force $F_B$ and the reaction couple $M_B$ in point B are chosen as the redundants. Figure 3-19 shows the deflected hinge structure with the redundants removed, and the deflections caused by the redundants. The resulting structure is a simple beam. The general solution has already been calculated in the previous section and the deflections needed can easily be determined using Eq. (3-15), (3-16) and (3-17).

![Diagram showing deflection caused by load and reaction forces](image)

Fig. 3-19 *Deflection of the double supported hinge caused by, a; the load and b; the reaction forces at point B.*

The boundary condition of the beam being clamped at the right side as well as the left, shows that:

$$\delta_B = \delta_B' + \delta_R = 0, \quad (3-18)$$

where $\delta_B$ is the total deflection at point B, $\delta_B'$ is the deflection caused by the loads $F_0$ and $M_0$ in the released structure, and $\delta_R$ is the deflection caused by the redundant reaction force $R_B$ and the couple $M_B$.

$$\theta_B = \theta_B' + \theta_R = 0, \quad (3-19)$$

where $\theta_B$ is the total angle at point B, $\theta_B'$ is the angle caused by the loads $F_0$ and $M_0$ and $\theta_R$ is the deflection caused by the redundant reaction force $R_B$ and the couple $M_B$. 

40
δ_B' can easily be calculated. From point A to point C, the beam acts as a simple cantilever and the deflection curve is given by Eq. (3-15). From point C to B, the beam has no forces or couples acting on it. Therefore, the curve of the beam will be a straight line. The deflection curve and the deflection and angle at point B is found by substituting the geometry in Eq. (3-15), and by extending the curve from point C.

\[
v_d(x) = v_i(x) \\
= v_i(L) + L \tan(\theta_i(L)) \cdot x \\
0 \leq x \leq L
\]

\[
L \leq x \leq 2L
\] (3-20)

When the angle \( \theta_i(L) \) is small, as was the assumption in section 3.4.1.1, the approximation; \( \tan(\theta) \approx \theta \) is valid and can be used to simplify Eq. (3-20). Substituting the load force \( F_0 \) and load couple \( M_0 \), the deflection becomes:

\[
v(x) = -\frac{1}{2EI_y} \cdot x^2 \frac{1}{2EI_y} \cdot x^2 + \frac{1}{6EI_y} \cdot x^3 \\
0 \leq x \leq L
\]

\[
= -\frac{1}{2EI_y} \cdot L^2 - \frac{1}{3EI_y} \cdot L^3 \left( \frac{1}{2EI_y} \cdot L^2 + \frac{M_0}{EI_y} \cdot L \right) \cdot (x-L) \\
L \leq x \leq 2L
\] (3-21)

The deflection and angle at point B are given by:

\[
\delta'_B = -\frac{3}{2EI_y} \cdot L^2 \cdot M_0 - \frac{5}{6EI_y} \cdot F_0
\]

(3-22)

\[
\theta'_B = -\frac{1}{2EI_y} \cdot F_0 - \frac{L^2}{EI_y} \cdot M_0
\] (3-23)

The deflection curve and the deflection and angle at point B caused by the reactions can be calculated by simply substituting the reaction force
Micromotion positioning system

$F_B$ and couple $M_B$ with a total length $2L$ of the beam instead of $L$ in Eq. (3-15), (3-16) and (3-17):

$$v_R(x) = -\frac{1}{2} \frac{M_B}{EI_y} \cdot x^2 - \frac{F_B L}{EI_y} \cdot x^2 + \frac{1}{6} \frac{F_B}{EI_y} \cdot x^3$$  \hspace{1cm} (3-24)

$$\delta_R = -\frac{8}{3} \frac{L^3}{EI_y} \cdot F_B - 2 \frac{L^2}{EI_y} \cdot M_B$$  \hspace{1cm} (3-25)

$$\theta_R = -2 \frac{L^2}{EI_y} \cdot F_B - 2 \frac{L}{EI_y} \cdot M_B$$  \hspace{1cm} (3-26)

Using Eq. (3-18) and (3-19) we can now calculate the reactions $F_B$ and $M_B$:

$$F_B = \frac{L}{2} \cdot F_0 + \frac{3}{4} \cdot M_0$$  \hspace{1cm} (3-27)

$$M_B = -\frac{L}{4} \cdot F_0 - \frac{1}{4} \cdot M_0$$

By substituting $F_B$ and $M_B$ in Eq. (3-24) and then summing Eq. (3-24) and (3-21) the total deflection of the double supported beam under loads $F_0$ and $M_0$ can be found:

$$v_d(x) = -\frac{L}{8} \cdot F_0 + \frac{1}{8} \cdot M_0 \cdot x^2 + \frac{1}{12} \cdot F_0 - \frac{1}{8L} \cdot M_0 \cdot x^3$$  \hspace{1cm} 0 \leq x \leq L

$$= -\frac{L^3}{6} \cdot F_0 + \frac{L^2}{2} \cdot M_0 - \frac{L^2}{2} \cdot F_0 - L \cdot M_0 \cdot x + \ldots$$  \hspace{1cm} (3-28)

$$+ \frac{3L}{8} \cdot F_0 + \frac{5}{8} \cdot M_0 - \frac{1}{12} \cdot F_0 - \frac{1}{8L} \cdot M_0 \cdot x^3$$  \hspace{1cm} L \leq x \leq 2L

42
The deflection and the angle of rotation at connection point C are given by:

\[
\delta_d = -\frac{1}{24EI_y} \cdot F_0 \tag{3-29}
\]

\[
\theta_d = -\frac{1}{8EI_y} \cdot M_0 \tag{3-30}
\]

Contrary to the simple beam, the deflection is only determined by the force \( F_0 \) and the rotation at the connection point only by \( M_0 \).

3.4.2.3 Torsion beam

The last possible hinge structure is based on the twisting beam as shown in Figure 3-20. The couple \( M_0 \) now acts as a twisting couple, indicated by the double arrow. Again we can make use of the superposition principle to analyse the structure. The structure is statically indeterminate and couple \( M_B \) is chosen as the redundant. The boundary condition is: \( \theta_B = 0 \), where \( \theta_B \) is the total twisting angle at point B.

Fig. 3-20  Deflection and twisting of the torsion beam hinge due to loads \( F_0 \) and \( M_0 \)

Again we can use Eq. (3-15), (3-16) and (3-17) to do the analysis of the deflection and Eq. (3-9) to determine the rotation. The deflection curve is given by:

\[
v_r(x) = -\frac{1}{2EI_y} \cdot x^2 + \frac{F_0L}{EI_y} \cdot x^2 + \frac{1}{6EI_y} \cdot x^3 \tag{3-31}
\]
Micromotion positioning system

Using the boundary conditions to determine \( M_B \), the deflection curve and the deflection at the tip can be found:

\[
\nu_T(x) = -\frac{1}{4} \frac{F_0 L}{EI_y} \cdot x^2 + \frac{1}{6} \frac{F_0}{EI_y} \cdot x^3
\]

(3-32)

\[
\delta_T = -\frac{1}{12} \frac{L^3}{EI_y} \cdot F_0
\]

(3-33)

The rotation is solely determined by the twisting couple \( M_0 \), and is given by:

\[
\theta_T = \frac{L}{Gt_p} \cdot M_0
\]

(3-34)

3.4.2.4 Hinge selection

For the selection of the best possible hinge structure, we must be able to compare them. Many variables are involved, which can be varied when designing the hinge, but the obvious criteria are a minimum couple at a certain angle and maximum load support. To provide a basis for selection, all the hinges have been given a geometry, so that the deflection is the same when the same force \( F_0 \) is applied. Only the length of the hinge beams will be varied to this end. We have assumed that the cross-section of the hinge beams is the same and that the Young’s modulus is the same.

This provides different rotations due to the load couple, \( M_0 \), which can serve as the basis of comparison. The rotations are given below for all four configurations (the first simple beam hinge discussed in section 3.4.2.1, is asymmetrical and can be used with a positive or negative rotation at the end of the beam). When considering the torsion hinge, \( \nu = 0.3 \) was assumed, which is the value for polysilicon from which the hinges will be made [3.5].
**Mechanical modelling**

![Graph showing rotation angle as a function of load force](image)

**Fig. 3-21 Rotation angle of the four possible hinge configurations as a function of the applied load, for which the deflection is the same.**

As can be seen from Figure 3-21, the rotation is largest for the twisting hinge under the same load, than that of the bending beams and thus the best choice when selecting the hinge. As was mentioned in section 3.3.4, bending of the hinge beams is unwanted, but can not be avoided. However, this need not be just a problem. Actually, it does have some advantages. A reduction of the gap size will reduce the step size, but it will also increase the electrostatic force, as will be shown in paragraph 3.3.4.

This effect compensates the increase of the load couple by an increase in the drive couple, thus increasing the maximum load that can be transported. The reduction in net step size for large loads is acceptable in most applications if the continued functioning of the device is guaranteed. Figure 3-22 shows the initial drive couple as a function of the loads. As can be seen from these curves, the increase in electrostatic force due to the sagging is larger than the force needed. The reduction of the gap is linear with the load, while the electrostatic force is approximately quadratic with the gap and thus quadratic with the load. By choosing the proper hinge geometries, the device will be able to transport all loads with an actuation voltage that is solely determined by
Fig. 3-22 *Increase in the initial drive couple as a function of the deflection.*

the device and not by the load. The step size will drop to zero for large loads due to the inability of the plate to rotate. This will ultimately determine the maximum load of the device.

To avoid large loads from damaging the actuators by stretching the hinge beams up to breaking point, it is advisable in most applications to incorporate an overrange protection in the form of a solid bump positioned between the actuators, which is slightly lower than the vertical plane of the unloaded actuator.

### 3.4.3 Plate bending

Due to the mass on the plate, the electrostatic forces and the support by the hinge beams, the plate is undergoing loading, so it will deform. When this bending would become too large, the lifting is less efficient and the step size will decrease. In extreme cases, it can cause the object to get stuck against unloaded and therefore unbent plates, due to excessive bending of the activated plate.

The deflection due to the load changes during the actuation. At first the load rests on the unrotated plate. As the plate is rotated, the weight of the object will move from two support points on either side of the hinge to one side of the hinge. However, the loading of the plate remains virtually the same. The plate remains loaded on both sides of the hinge.
point. If the plate is considered in a static situation, the couple around the hinge point caused by the force acting on one side must be equal to the couple caused by the other force on the other side of the hinge plus the couple caused by the hinge itself. At the moment all the weight of object rests on one side, the electrostatic force is of the same magnitude, and the overall loading of the plate in this case is twice that of the plate under just the load of the object, as the plate has not been rotated and therefore the couple caused by the hinge is absent. Figure 3-23 shows the loading of the plate and the deformation resulting from it.

![Diagram](image)

Fig. 3-23  Loads acting on the plate and the resulting deformation.

A simple plate assumption can be used, as the bending is small in comparison to the thickness of the plate. The plate is assumed to be flat, without any steps. This is a valid assumption, as the step are placed close to the end of the plate to avoid extra flexibility of the plate. The electrostatic force is considered to act as a concentrated force on the very end of the beam. Although this is not quite true, the electrostatic force on the plate changes during rotation and changes as the plate deforms, and no analytical analysis of this force and the resulting deformation would be possible. The assumption of the concentrated electrostatic force enables us to do an analytical analysis, which gives the worst case situation.

The hinge is modelled as a free support that allows free rotation, but not translation. The plate is statically determined, so a straightforward solution of Eq. (3-6) is possible. This is most easily seen when the load forces are modelled by a level plate with two freely rotating contact points on the plate and the hinge a force equal to twice the load force (see Figure 3-24).
Micromotion positioning system

Fig. 3-24 Modelling of the plate as a double, simply supported beam.

The deflection of the plate and the maximum deflection, are given by:

\[ v_p(x) = \frac{1}{6EI_y} \frac{F_l}{L_1 + L_2} \cdot \left( L_2^2 L_2 + 2L_1 L_2^2 \right) \cdot x + \frac{1}{6EI_y} \frac{F_l L_2}{L_1 + L_2} \cdot x^3 \] (3-35)

\[ \delta_p = \frac{1}{48} \frac{3L_1^2 + 6L_1 L_2 - L_2^2}{EI_y} \cdot F_l \] (3-36)

When calculating this bending for a typical actuator geometry, the resulting bending will be in the order of a few tens of nanometers. This bending is small, so the assumption of small deflections was valid.

3.4.4 Step size.

The distance over which an object is transported during one cycle of the drive sequence as described in section 3.2 and section 3.3.1 can now be more specifically determined. Figure 3-25 shows the cross section of the

Fig. 3-25 Cross section of the rotating plate that lifts the object to be transported.
plate which lifts the object. Filling in the geometric variables shown in the figure in Eq. (3-1) gives an expression of the step size

\[ \Delta x = L_1 \left( \frac{L_1}{\sqrt{L_1^2 + h_1^2}} \cdot \cos \left( \arctan \frac{h_1}{L_1} + \arctan \frac{h_2}{L_2} \right) \right) \] (3-37)

When unloaded and not actuated, \( h_f \) is determined by the layer thickness of the first sacrificial layer, the second nitride layer and the first polysilicon layer (see section 4.6). However, when a load is placed on it, the hinge beams will bend, as was shown in the previous section. During actuation, an additional electrostatic force will reduce this even further, and will increase as the plate rotates. The small step size, which is in the order of several hundred nanometers, will be reduced.

3.5 Electrostatic modelling

Figure 3-26 shows the two electrostatic plates and the movement of the plates in relation to one another. The electrostatic forces involved in the actuator are difficult to calculate using analytical methods. Nevertheless, to gain some insight into the factors that influence the electrostatic forces, some analytical analysis is required.

Fig. 3-26 Shape and relative position of the upper and lower electrode and the change in orientation due to rotation around the hinge axis.
3.5.1 Capacitance modelling

To calculate the electrostatic forces acting on the actuator, the capacitance between the actuating plates is used. One approach to calculate the capacitance is the geometric expansion as suggested by M.R. Boyd et. al [3.6].

The capacitor charges are associated with geometric features of the capacitor plates. The overall capacitance can then be expressed as a geometric expansion. The first term is proportional to the surface area of the plates, \( A \), the second term is proportional to the length of the edge of the plates, the perimeter, \( P \). This approach can be refined by including internal and external corners, proximity of edges, etc. As will be shown later, for our purpose the first two factors suffice:

\[
C = A \cdot C_{\text{area}} + P \cdot C_{\text{perim}}
\]  \hspace{1cm} (3-38)

The capacitance of the area per unit area, is given by the parallel plate assumption:

\[
C_{\text{area}} = \frac{\varepsilon}{d},
\]  \hspace{1cm} (3-39)

where \( \varepsilon \) is the permittivity of the medium and \( d \) is the distance between the capacitor plates.

![Fig. 3-27 A strip capacitor with a representation of the electrostatic fields.](image)
The capacitance of the perimeter per unit length, can be determined using the approximate capacitance formula for the capacitance of a pair of infinitely long parallel plates per unit length (see Figure 3-27). Many approximations are available [3.7]-[3.10]. Here we make use of the one derived by A.E.H. Love [3.11]:

$$C_{Love} = \varepsilon \left( \frac{w}{d} + \frac{1}{\pi} \ln \left( 1 + \ln \frac{2\pi w}{d} \right) \right),$$  \hspace{1cm} (3-40)

where the geometric variables are as indicated in Figure 3-27.

This can be broken up into the parallel plate assumption with a correction for the fringe fields. As is shown in [3.6] the error when using this approximation is less then 3% when \( w/d > 3 \). As this is the case with the geometry presented here and the inaccuracies caused by variation of the geometry, thickness of films and uncertainty in the Young's modulus are several times larger, this is accurate enough for our purpose. However, this assumes the thickness of the capacitance plates is much less then the gap.

In the case of the actuator presented here, as is the case for most surface micromachined devices, the thickness of the layers is of the same order as the gap. To correct for this, a capacitance addition by the sides of plates with finite thickness can be included using an effective width for infinitely long parallel plates of thickness \( t \), as derived by Wheeler [3.8]. This models the additional field caused by the side walls, as an increase of the width of the strip capacitance. This effective width can then be used to substitute the width in the model for thin plates. The effective width is given by:

$$w_{eff} = w + \frac{t}{\pi} \ln \left( \frac{4e}{\left( \frac{2t}{d} \right)^2 + \frac{1}{\pi \left( \frac{w}{t} + 1, 10 \right)} \right).$$ \hspace{1cm} (3-41)
The increase in width due to the thickness can then be used as an extra parallel plate capacitance of aforementioned width. This results in a capacitance increase at the perimeter due to the side walls of the plate of:

\[
C_{\text{side}} = \varepsilon \frac{t}{\pi d} \ln \left( \frac{2e}{d} \right) + \frac{1}{\pi \left( \frac{w}{t} + 1, 10 \right)}
\]

When assuming opposite edges do not influence each other, \( w \gg t \), the thickness addition can be simplified to:

\[
C_{\text{side}} = \varepsilon \frac{t}{\pi d} \ln \left( \frac{2ed}{t} \right)
\]

This simplification introduces an error in the side wall capacitance of less than \( 4 \cdot 10^{-6} \) for \( w > t \).

This capacitance can be used to complement the Love model. This results in a capacitance per unit length for the perimeter of:

\[
C_{\text{perim}} = \frac{\varepsilon}{2\pi} \left( 1 + \ln \left( 2\pi \frac{w}{d} \right) + \frac{t}{d} \ln \left( 2e \frac{d}{t} \right) \right) \cdot P
\]

This capacitance is half the correction for fringe fields and side walls in Eq. (3-40) and (3-43), because strip capacitors, for which the equations are valid, include two perimeters.

For perimeters where the edge of one electrode extends from the edge of the other electrode, as shown in Figure 3-28, we assume this to be a ground plane extending into infinity. The ground plane can then be modelled, using a mirror image of the electrode at the same distance of the ground plane as the original electrode. For this situation Eq. (3-44) holds, with \( d \) double the original distance.
3.5.2 Electrostatic forces

The electrostatic forces per area and length of perimeter can be calculated from these capacitances using: \( F = -\frac{\delta U}{\delta d} = 1/2U^2 \frac{\delta C}{\delta d} \):

\[
F_{\text{area}} = \varepsilon \cdot \frac{U^2}{2d^2} \cdot A \tag{3-45}
\]

\[
F_{\text{perim}} = \varepsilon \frac{U^2}{4\pi d} \cdot \left( 1 - \frac{t}{d} + \frac{t}{d} \ln\left( \frac{2ed}{t} \right) \right) \cdot P \tag{3-46}
\]

where \( F \) is the force, \( \varepsilon \) is the permittivity of the surrounding medium, \( U \) is the voltage between the plates, \( d \) is the distance between the plates, \( t \) is the thickness of the top plates, \( A \) is the area of the plates and \( P \) is the perimeter of the plates.

As can be seen from Eq. (3-46), including the thickness complicates the analysis considerably. However, the side wall contribution is as large as 70% of the total perimeter force contribution when \( t = d \).
Micromotion positioning system

This analysis assumes the capacitor plates will remain undeformed when undergoing the electrostatic forces and loading. This is not true as the plates will bend and come closer together. The actual deformation and the electrostatic forces can not be determined analytically as they are mutually dependent. The following analysis is therefore a worst case situation. The actual electrostatic force generated will be larger.

The above analysis assumes the plates to be parallel. However, the capacitor plates will be at an angle when the actuator rotates and the electrostatic force will depend on the angle of rotation, $\alpha$. Because the angle will remain small, as was determined in section 3.4.4, the electrostatic force can be calculated using an integration over the area/length, with a gap depending on $\alpha$. When the axis is at $x = 0$, the distance between the plates is: $d = h_2 \cdot \tan(\alpha)$ (see Figure 3-29). The error introduced by this assumption is less than 1%. The couple is found in a similar way by integrating the force, times the distance from the axis, $x$, over the same area/length with a varying gap.

![Diagram of capacitor plates with angle and distances labeled](image.png)

**Fig. 3-29 Change in plate distance due to rotation of the top plate.**

For an area at distance $a$ to $b$ from the axis and width $w$, the force and couple around the axis can be described as:

\[
F_{area}(\alpha)_{a}^{b} = \frac{1}{2} \tan(\alpha) \left( \frac{1}{h_2 - btan(\alpha)} - \frac{1}{h_2 - atan(\alpha)} \right) \tag{3-47}
\]

\[
M_{area}(\alpha)_{a}^{b} = \frac{1}{2} \tan(\alpha) \left( \frac{1}{h_2 - btan(\alpha)} + \ln(h_2 - btan(\alpha)) - \frac{1}{h_2 - atan(\alpha)} \right) \tag{3-48}
\]
For the perimeter at a distance \( a \) to \( b \) from the axis, the force and couple are given by:

\[
F_{\text{perim}(\alpha)} |_a^b = \frac{1}{4 \pi \tan \alpha} \left[ \frac{\ln \left( \frac{2 \cdot e(h_2 - btan\alpha)}{t} \right)}{h_2 - btan\alpha} - \ln(h_2 - btan\alpha) - \ldots \right]
\]

\[
M_{\text{perim}(\alpha)} |_a^b = \frac{1}{4 \pi \tan(\alpha)^2} \left[ -btan\alpha + \frac{t}{2} \ln \left( 2 \cdot \frac{e(h_2 - btan\alpha)}{t} \right)^2 - \ldots \right]
\]

\[
\left( h_2 + t \cdot \frac{th_2}{h_2 - btan\alpha} \right) \cdot \ln \left( 2 \cdot \frac{e(h_2 - btan\alpha)}{t} \right) + atan\alpha - \ldots
\]

The edges at the end and beginning of the plates do not change angle and are positioned above or below an infinite ground plane. The force and couple for these edges at a distance \( a \) from the axis of rotation are given by:

\[
F_{\text{edge}(\alpha)} |_a = \frac{U^2}{4 \pi (h_2 - atan\alpha)} \cdot \left[ 1 - \frac{t}{h_2 - atan\alpha} \cdot \ldots \right]
\]

\[
\left( 1 + \ln \left( 2 \cdot \frac{e(h_2 - atan\alpha)}{t} \right) \right) \cdot w
\]
Micromotion positioning system

\[ M_{\text{edge}}(\alpha)_{i_0} = \frac{U^2}{4\pi(h_2-\tan\alpha)} \left( 1 - \frac{t}{h_2-\tan\alpha} \right) \ldots \]
\[(1 + \ln \left( \frac{\epsilon(h_2-\tan\alpha)}{t} \right)) w \cdot a \]

(3-52)

Figure 3-30 shows the geometry of the capacitor plates. The areas and perimeters as indicated in the figure are modelled separately as a function of the rotation angle \( \alpha \).

\[ \text{Fig. 3-30 } \text{Geometry and areas of the electrode plates.} \]

For area B, which is not level, we assume that the transition fields between the areas are negligible. Figure 3-31 shows the situation. The vertical electrostatic force in that case is the same as the force of a plane capacitance covering the same horizontal area (i.e. the increase of area compensates for the angle of the force and the vertical force component is the same as the force would be if the area was level). The same applies for the force as a result of the perimeter. We can therefore ignore the step in the electrodes and consider the entire electrode to be flat for force and couple analysis.
Electrostatic modelling

![Diagram of electrode plates at angle \( \varphi \), showing that the vertical force is the same as that of planar electrode plates of the same horizontal size.]

The force and couple around the axis are given by:

\[
F_e = F_{area}(\omega)\frac{L_2}{L_0} + 2F_{perim}(\alpha)\frac{L_2}{L_0} + F_{edge}(\alpha)\frac{L_2}{L_0} + F_{edge}(\alpha)\frac{L_2}{L_2} \tag{3-53}
\]

\[
M_e = M_{area}(\omega)\frac{L_2}{L_0} + 2M_{perim}(\alpha)\frac{L_2}{L_0} + M_{edge}(\alpha)\frac{L_2}{L_0} + M_{edge}(\alpha)\frac{L_2}{L_2} \tag{3-54}
\]

Figure 3-32 and 3-33 show the force and the couple respectively. Both split into contributions of the area and perimeter.

![Graph showing electrostatic force of the area, perimeter, and side walls.]

Fig. 3-32  Electrostatic force of the area, perimeter and side walls.
Micromotion positioning system

![Graph showing the relationship between couple M [Nm] and rotation angle α [rad].](image)

**Fig. 3-33** Electrostatic drive couples of the area, perimeter and side walls

### 3.5.3 Capacitor plate isolation

The plates of the drive electrodes move closer together as the actuator rotates. To avoid a short circuit between the two plates, the rotation must be stopped before the electrodes touch.

![Diagram showing insulation possibilities between the two electrodes.](image)

**Fig. 3-34** Insulation possibilities between the two electrodes. a: an insulation layer, b: a landing electrode at the same potential as the upper electrode.

One possible solution is to include an isolating layer between the two electrodes (see Figure 3-34 a). This dielectric layer has the advantage,
that it will not significantly reduce the maximum rotation angle or reduce the electrostatic force. In addition, the extra layer stiffens the rotating plate, reducing any bending due to the loads.

Although the nitride isolation layer has a breakdown voltage which is higher than that of air, fields at the touching edge must be considered, and here the medium is air. These areas are therefore the most sensitive to breakdown.

The dielectric layer is not included in the capacitance calculations in section 3.5.1. The modelling is only relevant for the area capacitance term, as the fringe field and side wall fields still only have air as a medium. The gap, however, is increased by the thickness of the isolating layer with a different relative permittivity. The capacitance of a parallel plate capacitor with a dielectric of thickness \( h_d \) in the gap \( h \), can be calculated using the effective relative permittivity of [3.9]:

\[
\varepsilon_{\text{eff}} = \frac{h_d \cdot \varepsilon_d + (h - h_d) \cdot \varepsilon_a}{h}
\]  

(3-55)

where \( \varepsilon_d \) is the relative permittivity of the dielectric (nitride) layer and \( \varepsilon_a \) is the relative permittivity of air.

When calculating the force from the capacitance, \( h \) will change as a function of \( x \) and \( \alpha \), and results in a force and couple given by:

\[
F_{\text{area}}(\alpha) = \frac{1}{2} \varepsilon_0 w U^2 \frac{h_d(\varepsilon_r - \varepsilon_a)}{2(h_2-b\tan\alpha)^2} \frac{\varepsilon_a}{h_2-b\tan\alpha} \frac{h_d(\varepsilon_r - \varepsilon_a)}{2(h_2-a\tan\alpha)^2} \frac{\varepsilon_a}{h_2-a\tan\alpha} \]

(3-56)
Micromotion positioning system

\[ M_{area}(\alpha)_{u} = \frac{1}{2 \tan(\alpha)^2} \left( \frac{h_d \varepsilon_a - h_d \varepsilon_r + h_2 \varepsilon_a}{2(h_2 - b \tan \alpha)^2} + \cdots \right) \]

\[ \varepsilon_a \ln(-h_2 + b \tan \alpha) - \frac{h_2 h_d (\varepsilon_r - \varepsilon_a)}{2(h_2 - a \tan \alpha)^2} - \cdots \]  

(3-57)

\[ \frac{h_d \varepsilon_a - h_d \varepsilon_r + h_2 \varepsilon_a}{h_2 - a \tan \alpha} - \varepsilon_a \ln(-h_2 + a \tan \alpha) \]

where \( \varepsilon_0 \) is the vacuum permittivity.

Another advantage of using nitride as an insulating layer between the electrodes, is that it makes the plate on which the object rests, much more rigid. Doubling the thickness of the plate reduces the bending by a factor of 8, as the moment of inertia is inversely proportional to the thickness to the third power. In addition, the Young’s modulus of silicon nitride is more than twice that of polysilicon (3.85 GPa vs. 1.70 GPa resp.) reducing the bending even more. This makes the assumption that the capacitor plates are flat and therefore the electrostatic force analysis, more accurate.

Another possibility is to use landing electrodes as shown in Figure 3-34 b. These are placed at the contact point of the upper electrode when pull-in occurs. They are separated from the lower electrodes by a small gap, and thus the total electrode surface is reduced. This small reduction in electrode size has considerable influence on the resulting drive couple. The part that is removed is the part of the electrode with the longest arm resulting in the highest couple contribution.

When considering the drive couple, the insulator is the most advantageous to avoid contact between the charged electrodes when considering electrical field breakdown. However, there are other effects to consider. One is sticking of the moving plate to the substrate after actuation. When the drive voltage is switched off, the actuator should return to its original position. In other words, the spring force of the hinges must overcome any attracting force that may develop between the moving top plate and the lower electrode.
The force can be caused by sticking effects as it does during the processing stage. These include 'glueing' by dirt and residue of processing chemicals, the forming of hydrogen bonds on the surface of the material (see 4.5.16.1) [3.12]-[3.14] and permanent charging of materials such as silicon nitride and silicon dioxide [3.15].

Sticking due to dirt and the formation of hydrogen bonds is proportional to the surface area of the contact between the touching plates. Because of the rotation of the top electrode, the electrodes will only touch at the outer edge. This surface is small enough to avoid problems of this nature.

The permanent charging is an important effect to consider when using electrostatically driven devices. When using an insulating nitride layer between the electrodes, this layer may get charged during actuation. The permanent charge may cause the movable electrode to remain attracted to the lower electrode, even after the voltage is switched off. This charge need not be very large to withstand the spring force in the hinges and thus obstruct the functioning of the device, due to the very small distances involved when the electrodes touch. To avoid this buildup of charge in the insulating layer over consecutive actuation cycles, the actuation voltage should change polarity every cycle to compensate, resulting in an actuation cycle as shown in Figure 3-35. However, this does not avoid charging that takes place in one cycle, which could inhibit the proper operation of the device if large enough.

When using landing electrodes this permanent charging effect is avoided, due to the absence of an insulating layer that can get charged. Any charging that might occur in surface oxide layers that form on the conducting plates, is avoided as the electrode makes contact with the grounded landing electrode and any charge difference will be removed.

Fig. 3-35 Symmetrical drive cycle to avoid charge build-up in the dielectric layer.
Micromotion positioning system

To avoid any problems, an actuation cycle as shown in Figure 3-35 can also be used for this configuration as well.

As a general conclusion, the use of silicon nitride as an insulating layer offers the best solution. Nevertheless, actuators of both kinds have been fabricated and implemented in a positioning array (see chapter 6).

3.6 Load restrictions

The constraints for actuator design presented in section 3.3 and 3.4 are not the only factors that determine the permissible loads. Of course, the actuators must be capable of lifting the object and moving it, which is limited by the hinge rigidity, but the interaction of the actuators in an array and the overall system also put limitations on the possible objects that can be transported.

The bending of the actuators due to the load, should not be so large, that the actuators can not lift the load above the level of unloaded actuators. This would inhibit movement as the edges of the load would get stuck against unloaded actuators (see Figure 3-36).

Fig. 3-36 Lodging of the object due to bending of the hinges and insufficient lift.

Another important factor is the contact surface of the load. The surface must be sufficiently flat, so that no irregularities catch the actuators,
Load restrictions

Fig. 3-37 Irregular object surface catches the actuator blocking movement.

making it impossible to return to the original position (Figure 3-37). The object would get stuck, or break the actuator in question. As the movement of the object is small during one step, the irregularities must be large in order to cause such a problem. In the process described in chapter 4, it is estimated that surface roughness of up to about 200 nm is allowed, depending on the ration between the depth and surface area of the pits and bumps.

The actuators moving the object do not lift it more than the thickness of the gap, which is 380 nm in the process described in chapter 4. The surface planarity of the object that can be transported, must be better then this in order for all actuators to make contact with the object. If this is not the case, the load will rest only on a small portion of the actuators (Figure 3-38).

Fig. 3-38 Contact with only a part of the available actuators due to non-planarity of the object.

Another problem that arises from surface roughness and non-planarity is the bending of the hinges due to uneven distribution of the weight. Although this will cause all the actuators to be in contact with the load, it causes some actuators to rotate with a different starting angle and results in a different step (Figure 3-39). Some of the actuators will slip, not being able to move the object the full distance, or not able to move that distance itself. This slipping must be possible without the
Micromotion positioning system

Fig. 3-39 Different step sizes of neighbouring actuators due to non-planairity of the object surface.

actuators breaking. The surface roughness of the load must be very low, in order to make this possible.

The problems that arise from the surface roughness, i.e. the catching and the inability to allow slipping of the lever arm, are reduced by a rounding of the contact points due to processing effects. Bevelled sides of the contact point at an angle of about 45° make it impossible to catch the sides of all but the largest bumps, and makes slipping easier (see Figure 3-40).

Fig. 3-40 Rounding of the edges due to processing, reducing the catching of irregularities and enabling slipping.

A further problem might arise if the object that is to be transported, is a conductor. The contact points of the plate are directly connected to the top electrode and are at the same potential. Any charging of the top electrodes could charge the load. To avoid this problem, the top electrodes should be kept at ground potential and the signal is to be applied to the lower electrode.

These observations limit the possible loads that the system is able to transport. The contact surface must be polished, or be a natural crystal
surface to be flat and planar. Planairity problems are larger with bigger objects, so size might be limited. Possible loads include, silicon chips, surface and bulk micromachined components, micro-optical components, such as prisms and beams splitters and natural crystals in material analysis.

Other objects can be transported and positioned when placed on a 'platform' that satisfies the requirements, as shown in Figure 3-41. This makes it possible to distribute the load over a larger area to manageable levels and circumvents any problems with surface roughness. However, the increased lateral dimensions of the platform might limit the possible functions of the arrays as described in section 3.7, such as multiple object positioning.

Fig. 3-41 Platform that can be used to distribute the load of heavy objects, or prevent problems with object with rough surfaces.

3.7 Arrays

The individual actuator described so far in this chapter must be placed in an array to operate as the electrically actuated lever described in section 3.2 and to play its role as part of the positioning device. Several different configurations are possible providing several options in exchange for complexity in both layout and drive signals.

3.7.1 One-dimensional array

The most basic configuration is an array with all actuators in the same direction. Half the actuators rotate one way, the other half rotates the opposite way as described in section 3.2. This type of array is capable of moving an object along one axis. It can only be used for simple transport.
Only three electrical terminals are required for the whole array. One for each set of drive pulses and one ground. A one-dimensional array is shown in Figure 3-42.

Not all areas in an array need to have the same specifications. It is possible that some areas have different step sizes, resulting in different transport speeds and different positioning resolutions.

Fig. 3-42 Schematic representation of a one-dimensional array and a SEM of part of a one-dimensional array.

Another possible configuration for a one-dimensional array, is a rotational array, where the actuators are placed in a circular pattern (see Figure 3-43). Objects on the array can be rotated around the centre of the array in either direction. The step size of the actuators must increase, when further away from the centre, to provide smooth transport without slipping. This can be done by increasing the rotation angle and/or by decreasing the lever arm length as shown in section 3.4.4.
3.7.2 Two-dimensional array

By combining two one-dimensional arrays perpendicular to one another, it is possible to realize an array that can transport an object in both x and y direction. When one orientation is active, the other is in the rest position, and the object is lifted over the inactive actuators. This is possible because the active stroke of the actuators is upwards. A two-dimensional array is shown in Figure 3-44. Such an array can be used not only to transport objects, but also to accurately position them anywhere on the array with high accuracy.

The entire array is driven using five electrical terminals. Two for the drive pulses in each direction and one common ground. By combining the actuation of actuators in different directions, diagonal displacement is possible. For example, the object is lifted in the x-direction and is lowered in the y-direction, resulting in a diagonal displacement.

The disadvantage of a two-dimensional array is the lower density of the active actuators. In a two-dimensional array, only one out of four actuators are active at a time, as opposed to the one-dimensional array,
Micromotion positioning system

where one in two actuators are active. This doubles the load force and couple the actuators have to overcome (see section 3.3.3).

Similar to the one-dimensional array, a radial two-dimensional array is feasible. The circular pattern is combined with a radial pattern, making full $\varphi$-$r$ positioning possible. When an object is moved to radially, the orientation of the actuators that move the object, differs over the contact surface of the object. This is most apparent around the centre. When the object is moved close to the centre, all of the actuators in the radial array will push towards the centre, not along the direction of movement (see Figure 3-45). Parts of the object on opposite sides of the centre are pushed in opposite directions. To function properly, the actuators must be able to slip, as the object cannot move in two different direction at the same time. This requires objects with great surface planarity (see section 3.6). In this case, the object will move towards the centre, until all movement forces cancel out. A problem occurring with this type of array, is that once the object is positioned at the centre of the array, it can not be moved outward again, because all forces are balanced. However, it can still be rotated.

It is also possible to combine arrays that differ in other ways than orientation. In some applications, both speed and accuracy could be required. This can be accomplished by combining a fast array with low accuracy (large step size), with an slow array with large accuracy (small step size) in the same direction.

3.7.3 Multi-dimensional array

The number of combined arrays is not necessarily limited to two. As long as the actuators can lift the load and the actuators that are active are not too far apart to support the load, as many arrays can be combined as
desired. The arrays and combinations can be tailored to a specific need, each array adding two contact pads to the total.

E.g. The two dimensional x-y array could be combined with a circular array to make rotation possible, or even with a full \( \phi-r \) array. This would also solve any problems around the centre of the radial array.

### 3.7.4 Partitioned array

By putting several arrays next to each other, or splitting one up, depending on how you look at it, extra functions can be incorporated. Similar to the situation around the centre of the \( \phi-r \) array described in section 3.7.2, arrays at opposite sides of the parting line can be driven in opposite directions. An object placed on such an array, will move toward this line from any part of the array and stop on this line, when the movement forces on either side cancel each other (see Figure 3-46 a). When the parts are driven in the same direction, the array functions as a normal one-, or two-dimensional array (Figure 3-46 b).

![Fig. 3-46 Forces acting on objects on different part of a partitioned array. a, movement to the dividing line until the forces cancel out and b, both sections in the same direction for normal positioning.](image)

By splitting the array again along a line perpendicular to the first, not only a second line is formed to use as a reference, the intersection of the two lines, presents a reference point toward which an object can be moved, irrespective of its original position (see Figure 3-47 a). In addition to this, the object can be rotated when positioned at this point (see Figure 3-47 b). This is most easily achieved when the partitioning is done at 45\(^\circ\) to the main transportation directions.

The problem around the centre of the \( \phi-r \) array as described in section 3.7.2 can easily be solved by splitting the radial array in two parts at either side of the centre.
Micromotion positioning system

![Diagram](image)

Fig. 3-47 *Partitioning of a two dimensional array to enable a, centring of the object on the meeting point of the sections and b, rotation around this point.*

Splitting the array into more parts provides more possibilities for rotations, reference lines and points. In addition to this added flexibility, several different objects can be positioned on the same array. As long as the different objects are on different parts of the array, they can be moved independently from each other. This makes it possible to position objects relative to each other and then moving them together, which is not possible with other positioning systems.

For every part, however, two, or in the case of two-dimensional arrays, four, extra contact pads are needed, limiting the number of independently driven areas. In addition to this, each area must have contact wires to it. This can cause problems with large arrays.

### 3.7.5 Array indexing

Another way to increase the flexibility of an array is by using x-y indexing when driving the array. In this way, a specific area on the array can be actuated without the rest of the array responding. This way multiple objects can be positioned independently from one another on an array.

This can be done by making use of the minimal voltage, $V_{min}$, that must be applied to an actuator before it rotates. When able to switch the top and lower electrode independently at $\frac{1}{2}V_{min}$ and $-\frac{1}{2}V_{min}$ respectively, only the actuators where the opposite electrodes are connected to a voltage of opposing polarity, will rotate. The other actuators are only at half the minimal actuation voltage and will not rotate.
By connecting the top electrodes to x-index lines and all lower electrodes to y-index lines, a rectangular area can be selected to transport, while other parts remain still. This can be done with both one and multi-dimensional arrays.

In these indexed arrays, only translations can be executed. No reference lines or points and no rotations are possible, because those involve different movement directions on different parts of the array at the same time. In two-dimensional arrays, two different areas can be driven at the same time if the direction of transport is perpendicular to that on the other. This kind of array is especially useful for the absolute and relative positioning of multiple objects on the same array by positioning each one in turn. Of course, this kind of array can be partitioned as well. Creating one or more reference lines, points and possible points of rotation.

The limit is again the number of contact pads that can be used. A number of index lines situated side by side can be connected to the same contact pad to keep the number of pads at an acceptable level. This does increase the lateral dimensions of the areas that can be selected. However, the number of contact pads is much less than required for partitioning to the same resolution. Adding a line or column adds two contact pads for one-dimensional arrays and four for two-dimensional arrays. In addition, there are no problems of lead wires to areas within the array.

### 3.8 Electronically driven array

The arrays described in the previous sections are limited by the number of contact pads that can be used. When the micromechanical structures can be combined with electronic circuits, many of these limitations are lifted.

As will be show in chapter 4, the processing used to fabricate the actuators is compatible with a BiFET electronic process. This allows electronic circuits to be combined with the actuators on the same chip. However, the maximum voltage that can be switched is limited to about 20 V [3.16]. The actuation voltage of the actuators in an application must be below this, if electronic circuits are to be used.

The circuits described below operate on a lower supply voltage and an end stage is used for the level conversion.
3.8.1 Addressing circuit

To overcome the problem of many contact pads, addressing circuits can be used. The circuit consists of a multiplexer driving a number of memory cells, equal to the number of contact points. The drive circuits then provide the sections with the proper signals according to the function set in the memory cells. This makes it possible to program the function of the array, after which the array will execute this function until a new function is programmed. This can be used to drive both partitioned arrays and indexed arrays.

The size of the circuit is not critical. It can be positioned around the outer edge of the array, without influencing the density of the actuators in the array.

![Diagram of actuator array driven using programmable indexing.](image)

Fig. 3-48 Actuator array driven using programmable indexing.

The circuit diagram is show in Figure 3-48. The serial to parallel converter can be simple and need not be very fast. It serves as multiplexer and as the memory. Two outputs lines of the serial to parallel converter are connected to a selection circuit that switches one of three
signals, namely, the overlapping pulse trains, \( \varphi_1 \) or \( \varphi_2 \) as shown in Figure 3-3 or 3-35, or ground to a line of actuators or a section of a partitioned array.

### 3.8.2 Programmable cell logic

The ultimate partitioning that can take place is the individual programming of each actuator. This gives ultimate freedom in determining reference points and lines, rotation around every point, and complete independent positioning of many different objects on the same array. However, this necessitates a circuit at every actuator element, or as will be shown in the next section, each basic cell. Otherwise the number of lead wires becomes prohibitive. To provide the greatest density, the circuit must be positioned underneath the mechanical structures. As this is not possible with the process used here, the circuits would have to be placed next to the cell. In either case, the circuit must be as simple as possible, to minimize in surface area needed.

#### 3.8.2.1 Basic cell

In order to get the flexibility mentioned in the previous section, it is not necessary to drive all actuators completely independently, as only a limited number of drive combinations give desired movements. Therefore, the array can be split up into the smallest possible building block that can perform the necessary functions. Basically the control is transferred to the local level at a reduced number of connections and increased local logic.

![Diagram of basic cell](image)

*Fig. 3-49 The basic cell needed to perform all logical functions.*
Micromotion positioning system

Here we will look only at the two-dimensional x-y array. Although the same can be done for other arrays, this type of array can perform most desired functions. For an x-y array the possible functions are; translate along both x and y axis and the diagonals, all actuators down, and possibly all actuators up. These functions can be performed by a group of four actuators, two for each direction, rotating in opposite directions when actuated as shown in Figure 3-49. This will be considered the basic cell.

3.8.2.2 Drive circuit

The drive circuit is in principle similar to the addressing circuit described in the previous section. To drive this basic cell, the basic drive signals are provided (the overlapping pulse trains, $\varphi_1$ and $\varphi_2$ shown in Figure 3-3 and 3-35), which are switched by the local logic circuit, according to its programming. In addition to this, programming lines, the indexing lines and the supply voltage for the circuit and actuators are needed.

The basic circuit schematic is shown in Figure 3-50. The memory cells are programmed with one of the 10 states the cell can be in, and a drive circuit that switches the drive sequences to the four actuators.

![Drive circuit of the basic cell](image)

Fig. 3-50  Drive circuit of the basic cell.

The memory cells consist of simple flip-flop circuits. The enable lines are connected to an x and y index line by an NOR gate that can be used for selecting the cell to be programmed. The input line is connected to a data line. The four cells can have different index lines and one data line, or one index line and different data lines. The first approach needs a
minimum of five lines per cell, the second a minimum of six lines. The second does however, provide parallel programming of the cell, which needs only a quarter of the time.

The four signals have distinct functions, which are directly reflected in the circuit. The first signal determines whether or not the cell is active (i.e. whether it is stationary in the up or down position, or provides movement by switching). The second signal determines whether the cell is up or down, when the cell is inactive, or changes direction of movement 180°, by switching the pulse trains φ1 and φ2. The third signal changes direction of movement 90° by rotating the signals around the four actuators. The last signal rotates the movement 45° by switching between two actuators. Figure 3-51 shows a schematic representation of the switching done in the drive circuit. Table 3-1 shows the direction of movement and the drive signals on the actuators as a result of the programmed signals. These functions can easily be performed using pass logic in a small area, using the output signals to switch the high drive voltage for each actuator.
Table 3-1. Programming table for basic cell driving circuit.

<table>
<thead>
<tr>
<th>Prog. signals</th>
<th>Up-stroke direction of actuator</th>
</tr>
</thead>
<tbody>
<tr>
<td>d0 d1 d2 d3</td>
<td>➔  ➘  ◄  ➝  ➞</td>
</tr>
<tr>
<td>0 0 0 0</td>
<td>down</td>
</tr>
<tr>
<td>0 0 0 1</td>
<td>down</td>
</tr>
<tr>
<td>0 0 1 0</td>
<td>down</td>
</tr>
<tr>
<td>0 0 1 1</td>
<td>down</td>
</tr>
<tr>
<td>0 1 0 0</td>
<td>up</td>
</tr>
<tr>
<td>0 1 0 1</td>
<td>up</td>
</tr>
<tr>
<td>0 1 1 0</td>
<td>up</td>
</tr>
<tr>
<td>0 1 1 1</td>
<td>up</td>
</tr>
<tr>
<td>1 0 0 0</td>
<td>➔</td>
</tr>
<tr>
<td>1 0 0 1</td>
<td>➘</td>
</tr>
<tr>
<td>1 0 1 0</td>
<td>➘</td>
</tr>
<tr>
<td>1 0 1 1</td>
<td>➘</td>
</tr>
<tr>
<td>1 1 0 0</td>
<td>➞</td>
</tr>
<tr>
<td>1 1 0 1</td>
<td>➞</td>
</tr>
<tr>
<td>1 1 1 0</td>
<td>➞</td>
</tr>
<tr>
<td>1 1 1 1</td>
<td>➞</td>
</tr>
</tbody>
</table>

Which of these circuits is used, depends on the application. Without electronics, the maximum number of areas that can be driven is in the order of a ten separate areas (41 bondwires). When using simple indexing, this number can be increased to about 20 (41 bondwires). When using electronics, the number of areas is limited by the chip area that is needed for the electronic circuitry.
3.9 References


77
Micromotion positioning system


Micromechanical system fabrication

4.1 Introduction

As already discussed in section 3.4.5, the proper functioning of electrostatic actuators depends on the electrical field strength and thus on the size of the gap between the electrodes. The devices themselves also need to be very small if these are to function properly. When fabricating distributed micro systems, large numbers of devices must be processed at the same time. The fabrication process chosen must conform to all these demands.

As discussed in section 1.2, silicon micromachining process was chosen. Over the years many different micromachining technologies have been developed with a wide range of topologies. In section 4.2 a short overview is given of the different micromachining techniques. Section 4.3 discusses the possible use and applications of micromachined structures. In section 4.4 we will consider the demands on the processing, resulting from the desired structures and the compatibility with electronic processing, as well as the choices made. Finally, the surface micromachining process used to fabricate the devices is described in section 4.5.
4.2 Micromachining technologies

Many different technologies have been proposed over the years, each with its own advantages and disadvantages, designed for a specific device or to solve a specific problem. Although the number of technologies is too large to discuss in full, it is possible to distinguish a number of basic technologies with fundamentally different approaches and different advantages and disadvantages.

4.2.1 Bulk micromachining

This is the earliest form of micromachining [4.1],[4.2]. As the name implies, it shapes the bulk silicon of the wafer. The technique removes part of the silicon bulk using a mask on the backside of the wafer. For the etching of the thick silicon substrate, anisotropic etchants are used such as solutions of potassium hydroxide (KOH), ethylenediamine and pyrocatechol (EDP), tetra-methylammonium hydroxide (TMAH) and hydazine-water. These are anisotropic etchants, i.e. they have different etch rates in different crystal orientations of the silicon. The side walls are defined by the slowest etching crystal plane of the silicon; the \(<111>\) plane, and the angle of the walls is determined by the crystal orientation of the surface. When using the \(<100>\) wafers conventionally used for microelectronic processing, this angle is 54.74\(^\circ\).

![Bulk micromachining process using an electrochemical etch stop on a \(<100>\) silicon wafer.](image)

Fig. 4-1
Etching is in most cases done from the backside of the wafer, leaving the front of the wafer planar. Depending on the device, the wafer is usually not etched all the way through. A thin membrane is left which can serve to thermally isolate part of the wafer or to deform due to some external force that is to be measured. Figure 4-1 shows a typical example of a bulk micromachining process.

The membrane thickness (in effect the depth of the etch) can be controlled in several ways; The simplest way is to use a *timed etch stop*. When the etch speed and the original thickness of the wafer are known, the etch time needed to leave a membrane of a certain thickness can easily be calculated. This method is not very accurate due to variations in etch rate and results in a membrane of varying thickness due to the non-uniformity of the etch and the thin membranes needed (3 to 10 µm), as compared to the initial wafer thickness (525 µm). Nevertheless, a modified two step version of this technique is often used in industry with a fast initial phase up to around 20 µm, a thickness measurement, and a slow second phase.

To have more control over the resulting thickness of the membrane, special measures are taken. A slightly more complicated way is based on the *colour etch stop*, where the wafer is illuminated by a white light source. Observing the membrane from the other side, visible light starts getting transmitted through when the membrane reaches a thickness of approximately 10 µm and changes colour as the membrane gets thinner, due to the wavelength-dependent absorption of light in silicon. This provides a better reproducibility of the nominal membrane thickness, but the non-uniformity of the etch is still present.

Other etch stops make use of different etch characteristics of the membrane. The etching stops automatically and results in uniform membranes of predefined thickness. This can be done using a different type of material, such as silicon nitride [4.4], a high-boron etch-stop in which the stop layer is doped with very high concentrations of boron [4.5], or an electrochemical etch-stop in which the chemical etching process is stopped at a junction, such as the one between substrate and epi-layer, when it is biased above the passivation potential required for the electrolyte etchant [4.6].

The feature size is limited to features of full wafer thickness, of several hundred micrometres and beams of around 10 microns. The alignment of the mask on the back of the wafer, defining the etch hole that will release
the structure, with the features on the frontside can cause some problems, as normal optical alignment is not possible. However, special equipment such as infrared alignment, special jigs and double microscope aligners are available to reduce the alignment error to less than 5 μm [4.7],[4.8].

Bulk micromachining is now a widely accepted technique and is used in a wide range of commercially available pressure sensors, accelerometers and for thermal isolation of certain areas on a chip for applications such as infrared sensing and thermal flow sensing. All this notwithstanding, this technique is not capable of producing the vertical electrode structure needed for electrostatic actuators in a single wafer.

4.2.2 Surface micromachining

As the name implies, surface micromachining does not shape the bulk silicon, but instead builds structures on the surface of a wafer by stacking on additional layers [4.9], [4.10]. Conventional IC processing techniques, such as Chemical Vapour Deposition (CVD) techniques, evaporation and sputtering are used and IC compatible materials such as polycrystalline silicon, silicon nitride, silicon oxide, aluminium, etc. are used. Other IC compatible materials are used that are not in the standard repertoire of electronic fabrication facilities, such as silicon carbide, diamond like carbon, zinc-oxide, gold, thungsten, and organic compounds such as polyimide, but special care must be taken to maintain cleanroom compatibility.

The construction of free-standing structures in this technique is based on the selective etching of these materials. Figure 4-2 shows a cross-section of a simple surface micromachining process. On the chip surface a layer is deposited which will be removed selectively at a later stage in the process. This layer, which most commonly consists of silicon oxide, PSG or aluminium, is called the sacrificial layer. This layer is patterned using conventional photolithographic techniques to form the spacers for the free-standing structure. Subsequently, a structural layer is deposited and patterned on top of this to form a cantilever. This layer is often composed of polysilicon, although silicon nitride, silicon oxide, aluminium and a range of more exotic materials, such as organic polymers, are also used. Finally the wafer is etched using an etchant that has a high selectivity between the sacrificial layer and both the substrate and the structural material. This etch is usually done using a wet etch because of the need to under etch the structural layers. Figure 4-2 c.
shows the structure halfway through the etch, as some of the sacrificial layer is still present. After the etch is completed and all of the sacrificial material is removed, the structural layer is left free-standing and free to deform. Due to the large distances that need to be underetched this etch takes a considerable amount of time (in the order of 30 minutes to several hours) and the large selectivity is needed to avoid attack of the structural layer(s) and the substrate surface.

![Diagram showing the release of the structure by underetching of the beam.](image)

Fig. 4-2  *Simple surface micromachining process showing the release of the structure by underetching of the beam.*

Using this principle many different structures are possible. It is possible to repeat the procedure of sacrificial layer and structural layer deposition to fabricate more than one freestanding layer. This is not possible when using bulk micromachining.

However, surface micromachining has some drawbacks. Since the layers are stacked on the surface, this surface becomes non-planar. As this non-planarity can cause problems with the photolithography needed for the patterning of the layers, the total thickness must be limited. This limits the possible thickness of the layers to a maximum of about two microns, making side wall capacitances, as used in comb drive/detection structures, very small, thus complicating the design. When using thick photoresist layers, thicker layers can be patterned, but the resolution is lower.

### 4.2.3 Wafer-to-Wafer bonding

Wafer-to-wafer bonding makes use of several wafers, which are bonded together to produce one device [4.11], [4.12]. The wafers are either silicon, in which case it is possible to integrate electronics, glass or GaAs. The individual wafers are processed using bulk micromachining.
processes or simple etches. The stacking of the dies allows more freedom of design compared to bulk micromachining. As an example, it is possible to realize an accelerometer with sensing and driving electrodes to either side of a mass in a truly three dimensional structure. The wafers are bonded using anodic bonding, silicon bonding or eutectic bonding. The basic technique is shown in Figure 4-3.

Fig. 4-3 Basic wafer bonding. a shows the fabrication of the individual wafers, and b shows the complete structure.

Using this technology it is possible to fabricate e.g.: accelerometers, pressure sensors, fluid management systems such as pumps, valves and miniature mixing chambers and miniature gaschromatographs [4.13]-[4.16].

The technology is limited by the size of the lateral dimensions of the device features and it is difficult to change the thickness of the wafers. Alignment of the wafers to be bonded is difficult when using silicon wafers and not all processing equipment is able to handle the much thicker stack of wafers.

However, large and more sensitive devices are possible using this technique and often offers a kind of first level package, provided by the wafers, protecting the sensitive device during further processing and final packaging. The wafers to be bonded can be processed separately and bonded as a last step in the construction of the device.

4.2.4 SOI and Epi micromachining

Both Silicon On Isolator (SOI) micromachining [4.17], [4.18] and Epitaxial layer based micromachining [4.19]-[4.22] are newcomers at the micromachining scene. These combine some of the elements of surface
and bulk micromachining and offer new possibilities. In principle, they construct the moving structures out of the epilayer using deep plasma trench etching to pattern the epi and isotropic etching techniques to remove either the underlying oxide layer in the case of SOI micromachining, or to remove a doped buried layer, in the case of Epi micromachining, to free the structure. It has the advantage of using the epilayer, which is much thicker than the layers that can be used in surface micromachining. If serving as a seismic mass, or as the side electrodes in laterally driven actuators or sensors, it provides a considerably larger effect. The surface remains planar and can fairly easily be combined with on-chip electronics, although interconnect can present some problems.

However, inducing vertical movement is virtually impossible, making interaction with the outside world difficult and the thickness of the structure is determined by the epilayer thickness and can not be influenced by the designer. Figure 4-4 shows a simple Epi micromachining sequence.

Fig. 4-4  *Basic Epi-micromachining process, using a highly doped region to release the structure.*

4.2.5 Other micromachining techniques

Apart from the above mentioned techniques, there are many hybrid and exotic fabrication technologies. Most have been proposed in an attempt to solve a problem in a specific application, which can not be addressed by existing technologies. A few examples are given below to illustrate the many different approaches used in the construction of three-dimensional microscopic machines.

One such a technique is the wafer dissolving process (see Figure 4-5) [4.23], [4.24]. It uses the boron etch stop as used in bulk micromachining.
to define the structures. A silicon wafer is etched to form the anchor points of the structures (a). The etch depth defines the gap between the structures and the substrate in the final device. A deep boron diffusion is driven into to surface to define the thick beams and a second diffusion is used to produce a thinner beam (b). The silicon wafer is then bonded to a glass wafer (c). The whole is then etched in EDP. This dissolves the whole silicon wafer, with the exception of the deep boron diffusions, hence the name (d).

Fig. 4-5 Fabrication sequence of the wafer dissolving process.

Another techniques such as LIGA use selective deposition (see Figure 4-6) [4.25], [4.26]. On top of a sacrificial layer a seed layer is deposited (a) on top of which a thick photoresist layer is deposited and patterned that will function like a mold (b). Using selective deposition techniques like electroplating the mold is filled with a metal such as chrome (c) and finally the mold and sacrificial layer are removed (d). In many ways it is similar to surface micromachining, but has the possibility to fabricate layers of 10 to 100 micrometres thick, which is virtually impossible with other techniques.

Fig. 4-6 Cross section of a simple electroplating process
Flexible layers of organic materials are used, sometimes in combination with rigid films to construct structures that are able to assemble themselves into a three-dimensional shape. This is achieved by the use of bimetal effects and inherent stresses in the materials [4.27].

4.3 Requirements on the surface micromachining process

The micromachining process used for the fabrication of the positioning device, was designed for a more general use. Because of this objective, the entire process is more complicated than is strictly necessary for the construction of the actuators described in chapter 3. It must enable the fabrication of not only actuators but also sensors and micromechanical signal processing structures.

4.3.1 Surface micromachining applications

Sensors have been the first application of micromachining to be commercially produced and are now widely used for the measurement of pressure and acceleration. Most of these devices are bulk micromachined in combination with wafer to wafer bonding, but surface micromachined devices are now being introduced. The conversion of mechanical quantities into electrical signals can often be done using mechanical structures, which are now converted into micro mechanical designs.

Actuators have not seen many commercial applications. The difficulties of wear, dust and dirt, and the problems of connecting the movement or force produced to the outside world, have made practical applications difficult. Nevertheless, actuators have been used in sensor systems with self test and in servo sensors, which use an actuator to counteract the influence of the measurant. In these systems the actuation needed is a quantification of the measurant [4.28].

Micromachining has also been used to improve the performance of electronics and sensors by using mechanical shapes and structures to reduce thermal leakage in sensors, for thermal isolation, stress reduction, etc.

An application of micromechanics which has not had much attention, is the use of mechanical structures for signal processing
applications. Instead of using the physics involved in the electronics and electrical devices, a range of mechanical laws of physics could be used to this end [4.29]. This could lead to signal processing functions that were deemed impractical in electronics. E.g. Simple squaring circuits using the laws of electrostatic attraction, very low frequency filtering, some high temperature signal processing, etc.

### 4.3.2 Structural requirements

To enable the construction of most of the sensors fabricated using surface micromachining, both conducting and non-conducting layers are needed. One of the most complex sensor structures used is the differential capacitive sensor, in which the displacement of a conducting layer is determined by the change in capacitance between this plate and two other stationary plates. The basic geometric design requires three, stacked conducting layers. Of these, one can be a diffusion in the substrate, leaving a minimum of two freestanding structural layers to enable differential capacitance measurement. When considering the construction of actuators the process must be capable of constructing complex, free moving structures, such as micromotors and gears. Two structural layers are sufficient for the construction of these structures. One for the free floating parts, like rotors and one to keep these on the chip.

### 4.3.3 On-chip electronics

For many of the sensor systems, the incorporation of microelectronics on the same chip increases signal-processing speed, reliability and reduces the size and in some cases the cost. It also makes it possible to have a self-test capability, a digital output signal, an on-chip bus-interface, etc. The same advantages are true for micromachined actuators. However, the inclusion of electronics in the same wafer as the micromachining does present compatibility problems, unless the mechanical structures are made of layers present in the electronic processing [4.30], [4.31]. The reduction in cost that can be achieved by the integration of electronics is offset by more complex processing, lower yield, more expensive testing and packaging [4.32].

In many cases a solution can be found in the form of hybrid systems, in which an micromachined chip is combined with a microelectronic chip in the same package and connected by bond-wires
[4.33], flip chip techniques [4.34], [4.35] or in the form of a Multi-Chip Module [4.36]. However, for some applications the on-chip integration is essential for proper functioning of the device. Such as the amplification of small signals, which are sensitive to noise or parasitic capacitances of bond wires and the multiplexing and demultiplexing of sensors and actuators signals in arrays to reduce the number of lead wires [4.37].

As the surface micromachining module is designed to minimize the limitations on the designers of micro-electro-mechanical systems, it must be compatible with on-chip microelectronics. The microelectronic process used is the DIMES-01 BiFET process [4.38].

4.4 Surface micromachining process design considerations

The requirements presented in section 4.3 dictate many of the features of the surface micromachining process. However, other, more general, considerations must be taken into account as well.

4.4.1 Order of processing

The first question that must be answered when designing a surface micromachining process that is compatible with electronics, is whether the surface micromachining should take place before or after the processing of the electronic devices. Both approaches have advantages and disadvantages.

When the surface micromachining takes place before microelectronic processing (pre-processing), there are no electronic circuits present during the mechanical fabrication. This has the advantage that there are no restrictions to the processing temperatures, to which the electronics are sensitive (see section 4.4.3) and the performance of the microelectronic circuits is not changed by subsequent processing steps.

The disadvantages are; the problems with the photolithography during the electronic processing on a non-planar surface (see also section 4.4.2), the high temperature steps that could change the mechanical characteristics of the structural layers and will change the doping profiles of polysilicon layers by diffusion of dopants from highly doped layers such as PSG. Moreover, the materials and processes that
Micromechanical system fabrication

can be used during the micromachining, have to be restricted to ensure it does not contaminate the standard electronics line. It also makes it necessary to protect some structural layers from oxidation during the formation of the thermal oxide and will result in the formation of so-called bird beaks at the edge of mechanical layers on silicon.

When the surface micromachining takes place after all electronic processing has taken place (post-processing). The wafers are taken from the standard line before the interconnect is deposited, as this will not be able to withstand high temperature processing. As the wafers need not re-enter a standard processing line all possible sorts of materials can be used. The diffusions present in the electronics can be used in the structures as electrodes, underpasses, etc. without the need for an extra diffusion.

The disadvantages include the restriction on high temperature processing steps, the need to protect electronics with an extra ‘buffer’ layer and the extra difficulty of making contact holes on the electronics. The last restriction makes the use of washed emitters, in which a hole is present in the oxide, questionable. The buffer layer can influence the contact, thereby changing the transistor characteristics.

We have chosen for the second option; surface micromachining after the electronic processing has taken place. This provides the possibility to solve the problems that occur during the surface micromachining, by adapting the processing. This cannot be done if the surface micromachining is done first, as the processing of the electronics is standard and can not be changed. However, this choice limits the degree of freedom in the mechanical design.

4.4.2 Structural and material choice

The structural requirements force two free-standing structural layers (see section 4.3.2). These structural layers have to be electrically conductive, for which two possibilities are available in the processing facility used; doped polycrystalline silicon and metals. Metals like aluminium and even titanium, are structurally less strong and stable than polysilicon and after the deposition of one aluminium layer, no high temperature steps, such as CVD, can be used. Metals like tungsten have high melting points (3410°C), but can cause problems on silicon contacts during high temperature processing steps, if no barrier layers are used. This, in
Surface micromachining process design considerations

combination with the difficulty of finding a suitable sacrificial layer material, has led to the choice of polysilicon.

To free the structural layers, two sacrificial layers are needed, for which many possible materials are available. These include aluminium [4.39], resist or polymers [4.40] and undoped or doped silicon oxides [4.10], [4.41]-[4.46]. As both the aluminium and the organic layers can not withstand the moderately high temperatures during processing, a deposited silicon oxide is chosen as the sacrificial layer. These can be etched in hydrofluoric acid (HF) with a high selectivity to polysilicon and can withstand the high deposition temperatures of the subsequentially deposited structural layers. Of the many silicon oxides available, a doped oxide called Phosphorous Silicon Glass (PSG), is chosen as the doping of the silicon oxide raises the etch rate in HF, which is an advantage considering the long underetching necessary to free the structure.

To provide electrical insulation between the two polysilicon layers a non-conductive layer must be deposited either on top of the lower layer, or below the top layer. Because differential capacitance measurement needs a rigid upper and lower electrode and a flexible middle electrode, the insulating layer was placed underneath the top layer. The material used for this layer is silicon nitride. Silicon nitride is a very good insulator and has high structural strength. It can also be etched with a high selectivity to silicon oxide/PSG in HF. Other non-conductive materials, such as oxides and undoped silicon provide no selectivity between sacrificial layers and conducting layers when etching in (B)HF.

The compatibility requirement with the existing DIMES-01 BIFET electronic process, puts contraints on the processing, but very little demand on the choice of the structural and materials to be used for the surface micromachining structure. To protect the electronics present during surface micromachining, the wafers are taken out of the standard line before the contact holes are opened. The thermal oxide layer that covers the electronics at this point is protected agains oxide etchants used in the micromachining process, with a silicon nitride layer. This layer seals the active components from any chemicals used in the subsequent processing. When the contact holes are opened, the nitride is locally removed.

For the same reasons aluminium is not used as a structural material, the aluminium interconnect must be deposited as the last layer. The
Micromechanical system fabrication

freeing of the structures in the final sacrificial etch, takes place after all
depositions have taken place. A more thorough discussion of the reasons
behind this choice are presented in section 4.5.11.

Fig. 4-7  Basic structure necessary to enable differential capacitance
measurement (structure on the left) and containment of a
free floating structure (structure on the right).

These considerations lead to the layer structure shown in Figure 4-7. To
construct the mechanical devices, five of these layers are patterned and
stacked. This shows one more structural demand on the layers; The total
height of the structures above the surface of the wafer, in this case the
silicon nitride layer that seals the electronics, must be limited. When the
surface becomes very non-planar, problems can occur with the
application of the photoresist layer for the patterning of the top layers.
Step coverage of this layer is poor and thickness variations of the layer
result in width variations due to over exposure of the thin parts of the
resist and loss of resolution due to different focusing lengths. Figure 4-8
shows a SEM of an aluminium interconnect line going over a stack of the
five structural layers. As can be seen, the width on top of the stack is
considerably less than that at the bottom.

To avoid these problems the maximum height of the structure is
kept at 2 micrometres. From this, follows that the thickness of the five
layers that make up this stack should be around 400 nm. This is quite thin
and leads to problems with the release of these very flexible layers with
very small gaps between the layers. This problem is further discussed in
section 4.5.15. Furthermore, avoiding high temperature stress aneal of
polysilicon layers forces a deposition technique with a low deposition
rate (see section 4.5.5), which favours thin layers.
4.4.3 Temperature steps

Processing of the materials used in this process, take place at temperatures up to 1100°C. The high temperatures to which the structural layers and the electronics are subjected during additional processing, may change their characteristics. For the electronics in particular, this is unacceptable. If electronic device characteristics change, electronic designs that work when tested on the electronic process alone, may cease to function when combined with surface micromachining.

The reason for change in the electronics, is the diffusion of the dopants, changing the junction depth. The diffusion of the dopants is dependent on the type of dopant and the temperature as given by [4.47]:

\[ D = D_0 \exp\left(\frac{-E_a}{kT}\right), \]  

(4.1)

where \( D_0 \) is the Diffusion coefficient, \( E_a \) is the activation energy, \( k \) is the Boltzmann’s constant and \( T \) is the absolute temperature.

The lighter dopants, such as Boron, diffuse faster at the same temperature than heavy dopants such as Arsenic. As can be seen from the equation, the diffusion rate increases exponentially as a function of temperature. The diffusion rate is small up to temperatures of around

Fig. 4-8 Loss of line width of aluminium interconnect on high structures due to overdevelopment of thinner photoresist.
850°C and starts to rise quickly at higher temperatures. For phosphorus dopants, the diffusion is 60 higher at 1000°C. The base width of a bipolar transistor is particularly sensitive to these effects. The total temperature load that the electronics undergo after their fabrication is called the thermal budget. In the processing of the surface micromachining, this must be kept to a minimum to avoid damaging the electronics.

Following is a list of temperatures for the different processing steps of the surface micromachining layers. These are approximate values, as the processing temperatures can be varied to change the mechanical/electrical properties of the layers. However, these variations are small and have no significant influence on the diffusion rate of the dopants.

- **Silicon nitride**: deposition at 850°C
- **PSG**: deposition at 425°C, anneal at 850°C
- **Polysilicon**: deposition at 575°C, anneal at 605°C
- **Activation of dopants in Polysilicon** at 800°C

As can be seen from this list, these temperatures do not present a problem for the electronics. One process step however, does present a problem. As will be shown in the next section, standard polycrystalline silicon has a high residual mechanical stress [4.48], [4.49]. In other surface micromachining processes this stress can be removed by annealing at temperatures of 1050°C or higher for more than 30 minutes. This temperature will cause problems with the electronics and other methods must be found to reduce the stress in the polysilicon films.

All this notwithstanding, some high temperature processing could be allowed. The standard processing of the electronics ends with an additional drive-in step of 45 minutes at 1000°C. By reducing this drive-in time, the allowed thermal budget of the surface micromachining can be increased if necessary. However, only a small part of this thermal drive-in can be used, as the relative diffusion rates of the different dopants change with temperature. This makes it impossible to get the same diffusion depth for all dopants at a different temperature. This would also change the standard electronic processing and is therefore to be avoided if possible. Moreover, such break-in in the microelectronic process ruins the modular concept, although this is already somewhat compromised by contact holes and the interconnect.
4.4.4 Low stress mechanical layers

A requirement on all structural layers in micromechanical devices is a low mechanical stress. A large tensile stress can cause bridge like structures to break [4.50]. Large compressive stress causes bridges to buckle (see section 5.3) [4.51]. These curve up or down to release part of the stress. Another factor that might influence the performance of mechanical structures is the shift in resonance frequency and the spring constant of bridges and bridge like structures.

Another important factor is a gradient in the mechanical stress through the thickness of the layer. If, for example, the stress in the top of the layer is more tensile than the stress in the lower part of the layer, bridge structures would stay straight and flat, but cantilever structures, corners of membranes, etc. would curl upwards (see Figure 4-9). The

![Image of bendingcantileverbeams]  

Fig. 4-9 Bending of cantilever beams due to uneven stress distribution through the layer.

The effect is the same as that of a bimetal in which a change in temperature, introduces a stress, due to the difference in thermal expansion of the two metals. As a result the bimetal curls up or down.

As a total absence of stress is impossible to achieve, a small stress will always be present. To avoid any problems with the different effects on the mechanical structures for tensile and compressive stress, a predictable stress is desired of either type. For the stability of the structures, this stress should always be tensile. This keeps bridges of all lengths straight and taut and reduces the effects of sticking (see section 4.5.15). Problems due to stress gradients through the layer are

95
reduced by low stress levels, but can only be solved by symmetrical stress distributions through special consideration during processing.

Stress can build up in the layers due to many factors. The main reason being the difference in thermal expansion coefficient of the structural layer and the silicon substrate. As the whole wafer is cooled from the processing temperature to room temperature the layers will contract more or less than the substrate, leaving the layers with a tensile respectively, compressive stress. Another reason for stress is the internal structure of the layer. In polycrystalline silicon for example, the crystal size, shape and orientation has a strong influence on the resulting stress level [4.52], [4.53]. Also mentioned as a reason for stress is the matching of the crystal lattice on that of bordering layers. A difference in lattice size being cause for the stretching or compressing of the crystal structure at the interface. The presence of dopants in the layer, which disrupt the structure of the material also influences the stress level.

One problem in the design of a surface micromachining process is the difficulty of determining the stress in a layer. This problem remains when the process is complete. During fabrication of the devices, the stress needs to be monitored to check the processing.

Many different methods are available and are discussed in chapter 5, including a number of new measurement structures, which are designed to overcome the problems encountered when using existing stress measurement techniques to monitor stress in a surface micromachining module.

4.5 Surface micromachining module

The process presented, is set up as a module that can be added to the normal electronic processing, or used as a separate surface micromachining process. It provides designers the possibility of using mechanical structures for sensing, actuation, or signal processing functions. Because of the complexity and the influence of each process step on previous and following process steps, it is difficult to find an unambiguous order of reconstruction in the reasoning and clarification necessary of all steps. Here the resulting processing sequence is used as a basis to discuss the problems and choices encountered.
4.5.1 Electronic processing

The electronic processing that is done prior to the surface micromachining is called the DIMES01 BIFET process [4.38]. The process consists of six different transistors, viz. the bipolar BS-NPN, SP-PNP, BW-NPN and WP-PNP transistors, a lateral PNP transistor and a JFET. Of these transistors the WP-PNP and JFET make use of the Washed N-doped (WN) region. This region is defined using the thermal oxide on the wafer as a mask and uses the implantation window as a self aligned contact hole. Therefore, the WN region basically leaves an unprotected bare silicon. In the normal processing sequence it would be covered by the interconnect metal. However, as will be shown in the next section, a protection layer is required to cover the whole wafer before any surface micromachining takes place and the characteristics of these transistors could be changed. To avoid such problems, these transistors are not used in combination with the micromachining module. To enable the use of the JFETs an optional boron implantation is used to realize the p-channel instead of the WN implantation.

The implantations used for electronics can be used to form electrodes and electrical wiring underneath the mechanical structures. The electronic components must be kept at a lateral distance from the micromachined structures, as the aluminium interconnect must be protected during the etching of the sacrificial material as will be discussed in section 4.5.15.

During processing, a 330 nm thick thermal oxide is grown to prevent tunnelling and crystal damage during implantation. The dopant concentration of the SN diffusion in particular causes lateral differences in growth rate of this thermal oxide, resulting in an oxide thickness over the SN regions of 450 nm. This surface irregularity is small, but must be taken into account when construction of mechanical devices takes place on top of it.

The wafers are taken out of the electronic process after all the implantations and drive-in steps have taken place, and before the contact holes are opened in the thermal oxide and the aluminium interconnect is deposited. The wafer then looks as shown in Figure 4-10.
4.5.2 Nitride protective layer

The first step in the micromachining module is the protection of the underlying electronics and oxide with a layer of silicon nitride of approximately 150 nm in thickness. Nitride is chosen because a non-conductive layer is needed that can be etched with a high selectivity to polycrystalline silicon, which is used for the first mechanical layer and is not attacked by the sacrificial etchant, BHF.

A low-stress variant is used to avoid changing the electrical properties due to piezoresistive effects. Normal stoichiometric nitride has a high tensile residual stress (in the order of 1 GPa) and is therefore not suited. The stress can be reduced by changing the deposition parameters [4.54]. The low-stress silicon nitride used, is a silicon-rich nitride, deposited using LPCVD at 850°C, 150 mTorr, with gas flows 170 sccm DCS and 30 sccm NH₃. Growth rate is 6.8 nm/min. The resulting stress is approximately 210 MPa. The low-stress nitride has the added advantage that the etch rate in HF is lower, viz. 0.3 nm/min for low stress and 0.55 nm/min for stoichiometric nitride [4.55].

The layer is kept as thin as possible to avoid a large step in the contact holes for the aluminium interconnect. At the same time, a thick layer is undesirable if doped layers in the substrate are used to form capacitances with the mechanical poly layers, as this would reduce the capacitance. However, if the layer is too thin, overetching of the mechanical layers and a limited selectivity can cause the nitride to be etched all the way through. This would expose the underlying thermal oxide, which would be etched during the patterning and the later removal of the sacrificial layer. The electronics would be exposed and deposition of aluminium interconnect on the bare diffusions would short-circuit the entire wafer.
Keeping overetching, selectivity of etchants and variations in layer thickness in mind, a total thickness of in the order of 100 nm will be lost. A margin of 50 nm is sufficient to ensure breakthrough will not occur leading to the 150 nm thickness mentioned. The resulting structure is shown in Figure 4-11.

![Si₃N₄](image)

**Fig. 4-11 Cross-section of the wafer after deposition of the silicon nitride protection layer.**

### 4.5.3 First sacrificial layer

A layer which is suitable for sacrificial etching must be etched quickly and with a high selectivity to the layers between which it is sandwiched and must be deposited at a reasonable rate to enable relatively thick layers. As was discussed in section 4.4.2, the sacrificial material used is Phosphorous Silicon Glass (PSG). PSG is a LPCVD deposited layer of silicon oxide, in-situ doped with phosphorus to increase the etch rate in hydrofluoric acid (HF). Deposition is done at 425°C at 200 mTorr. Gas flows are: 30 sccm SiH₄, 20 sccm PH₃/SiH₄ resulting in 46 sccm SiH₄ and 4 sccm PH₃ [4.46]. The resulting layer has a phosphorus content of 3%. Tests have shown that a phosphorus content of 3% has a high etch rate, is stable and the phosphorus concentration is not high enough to introduces phosphorus dopants in the bordering polysilicon layers.

PSG etches quickly, because it is a very poor quality silicon oxide which is porous. However, this makes the layer mechanically unstable. At higher temperatures, the material reorganizes its internal structure and becomes more dense, which results in a reduction of the etch rate. To avoid the mechanical stress this might cause in layers on top of the sacrificial material, the deposition is followed by an annealing step to densify the layer before other layers are deposited. This step takes place after deposition of the PSG layer in a nitrogen atmosphere at 850°C for 30 minutes. The shrinking is dependent on the phosphor content of the PSG and the 3% phosphorus chosen here, results in approximate 5% shrinkage. This loss in thickness can be compensated for by a slightly longer deposition.
The first sacrificial layer determines the gap in the micromachined structures. This gap should not be large, as this would reduce the capacitance of sensing plates and the electrostatic force for actuators. As was mentioned in section 4.4.2, the total thickness of the stack should be limited to avoid problems with the photolithography. On the other hand, if the gap is very thin, the possible movement of beams and plates becomes very small, limiting the possibilities and the chance of the plates getting stuck during processing becomes larger. This last problem of sticking is further discussed in section 4.5.15. In addition, the under etch rate of the PSG is dependent on the thickness of the layer. As the layer becomes thinner, the reaction products and fresh etchant are slower to move in and out of the gap and the reaction slows down [4.56]. As a compromise, the thickness of the layer after deposition is chosen to be 400 nm. After annealing the resulting thickness is about 380 nm.

Patterning of the sacrificial layer is done using a wet BHF etch. This results in slight underetching of the mask. The sidewalls, however, are steep. As will be discussed in the sections describing the deposition of the structural layers, the step coverage of these layers is very good and the vertical sidewalls can cause problems during the patterning. The patterning of the polysilicon and silicon nitride layers is done using plasma etching, which is highly directional (see Figure 4-12). This leaves a small ribbon of material along the whole step when the layer is etched through, which is called a stringer (see Figure 4-13) [4.57]. To remove

![Fig. 4-12 Etching of structural layers using plasma etching which results in material being left along steps.](image)
Fig. 4-13 *SEM of the remains of the polysilicon layer (stringer) along the step of the sacrificial layer due to highly directional plasma etching.*

This stringer, long overetching would be needed. However, due to limited selectivity of the plasma etches, material underneath the rest of the wafer would be lost as well.

To reduce the overetching needed, the sidewalls of the sacrificial layer should be beveled. This is achieved by an argon implantation in the unpatterned but annealed PSG layer [4.58]. The implantation is done with an energy of 50 keV and a density of $10^{14}$ cm$^{-2}$. This implantation damages the structure of the top layer, making it less dense. This causes the top of the PSG to be etched quicker than the bulk of the PSG. This causes fast underetching of the mask. From this expanding opening the

Fig. 4-14 *Cross-section of the PSG etching process, and the resulting 45 degree bevel along the side of the sacrificial layer.*
slower etching PSG underneath is etched from the top. The top layer functions as a variable etch window for the bulk of the layer and causes a tapering of the sides. The result is shown in Figure 4-14. The angle of approximately 45° causes the plasma etch to attack the sidewalls more easily. However, the underetching of the mask is slightly more than normal. The result on the structure is shown in Figure 4-15.

Fig. 4-15 Cross-section after deposition and patterning of the first sacrificial layer.

4.5.4 Contact holes to polysilicon

The next step involves the opening of contact holes through the silicon nitride layer and the thermal oxide layer to provide contact between the conductive regions in the substrate and the first polysilicon layer. This enables the designer to use diffusions as interconnect underneath and in the direct vicinity of the mechanical structures.

This is done after the first sacrificial layer is deposited and patterned, because the contact holes would provide access to the thermal oxide layer would otherwise be etched during the patterning of the sacrificial PSG layer. Opening the holes after sacrificial layer patterning also makes sure the contact holes are not polluted by extra processing step between the opening of the contact holes and the deposition of the material that makes contact, in this case the first polysilicon layer.

The contact holes are opened in two steps. The holes are opened first by a timed plasma etch to remove the nitride layer and a considerable amount of the thermal oxide layer. This is a CF₄ based etch, which etches silicon nitride and thermal oxide at approximately the same rate. The etch is done at 0.05 mbar with 60 Watt with 25 sccm CHF₃, 50 sccm CF₄ and 40 sccm He. The same mask is then used for a short oxide etch in BHF to remove the last oxide from the silicon in the contact hole. The extended plasma etch reduces the underetching of the nitride by the wet etch, and the wet etch removes the oxide without damaging the silicon surface. This provides a clean undamaged contact. The steep sidewalls and considerable depth of the contact hole (in the order of
500 nm) is no problem for the LPCVD deposited polysilicon, which features an excellent step coverage.

When making electrical contact between doped regions in the substrate and a structural polysilicon layer, care should be taken to make contact only on n-doped regions (DN or SN), as the doping of the polysilicon layers is also n-type and an n-p junction would result otherwise. A cross-section of the wafer is shown in Figure 4-16.

Fig. 4-16 Cross-section after opening of the polysilicon-doped region contacts.

4.5.5 First polycrystalline silicon

Following the opening of the contact holes, the first structural layer is deposited. This is a conductive layer consisting of polycrystalline silicon (polysilicon). Polysilicon is normally deposited at temperature between 600 and 620°C. At these temperatures the polysilicon has a residual stress in the order of ~5 GPa [4.48], [4.49], [4.59]. This compressive stress is much too high to construct free-standing bridges and is usually reduced by annealing at 1050 to 1150°C for 30-60 minutes (see also chapter 4.4.4). In the presence of on-chip electronics this high temperature anneal is not possible. Rapid Thermal Annealing (RTA) can reduce the stress at a minimum addition to the thermal budget [4.60], [4.61]. However, temperatures up to 1000°C still leave the polysilicon under compressive stress and higher temperatures can cause blistering of the LPCVD nitride. Therefore, a solution was sought in the deposition parameters [4.62], [4.63].

The silicon is deposited at 575°C, 150 mTorr and 45 sccm SiH₄ to a thickness of 300 nm [4.64]. This is just below the temperature for the formation of polycrystalline silicon. At these temperatures the silicon consists of small crystals surrounded by amorphous silicon. This layer is then annealed in-situ, at 605°C for 60 minutes. At this temperature the small crystals grow as the amorphous silicon surrounding them crystallises. This results in a polysilicon layer that has a small tensile stress (170 MPa), as is desired for the construction of mechanical
structures, without the need for high temperature anneal and an even stress distribution through the layer due to the small silicon grains.

The intrinsic stress is believed to come from the lower density of the amorphous silicon or polycrystalline silicon. As the wafer is cooled from the processing temperature to room temperature, the crystalline silicon of the wafer shrinks slightly more than the less organised amorphous and polycrystalline silicon. This causes compressive stress in the polysilicon layer after deposition. In polycrystalline silicon the stress can be reduced by rearranging the crystal boundaries between the individual crystals at high temperatures. If there is amorphous silicon present between the crystals, this can be crystallized at much lower temperatures. As the amorphous silicon crystallises, its volume is reduced and the stress will change from compressive to tensile. A disadvantage is the lower deposition rate (2.4 nm/min), but this does not present a problem with the thin layers used.

The layer is then doped with phosphorus by implantation at 30 keV and a dose of 5×10¹⁵ cm⁻². To activate the dopants, an anneal is needed of 850°C for 30 minutes. However, activation is not done immediately, but is achieved by the next high temperature deposition step. The resulting layer has a resistance of 200 Ω/□. Although this is high, it is sufficiently low for the capacitive and electrostatic sensors and actuators with acceptable series resistance.

Patterning of the polysilicon is done using a plasma etch. Etching is done at 0.09 mbar at 50 Watt with 5 sccm O₂, 45 sccm SF₆ and 45 sccm He. This etch is slightly isotropic with a lateral etch rate of approximately 40% of the vertical etch rate. Some overetching is still needed to remove the whole stringer along the steps. The selectivity to silicon nitride is considerable (1:2.6) and only a thin layer of nitride is lost. Selectivity to the PSG is good (1:3) but the overetching does cause the thickness of the first sacrificial layer to be slightly lower around polysilicon structures than directly underneath thus increasing non-planarity. The resulting structure is shown in Figure 4-17.

Fig. 4-17 Cross-section of the wafer after depositing and patterning of the first polysilicon layer.
4.5.6 Second sacrificial layer

The second sacrificial PSG layer is deposited in a similar manner to the first sacrificial layer. Thickness is again 400 nm after deposition and 380 nm after anneal. This layer is also subjected to the argon implantation to create the bevel. In contrast to the first layer however, the second layer is not only used as a spacer, but also as an etch stop layer for subsequent etching of the structural layers to be deposited on top of it. Therefore, patterning follows a different approach. Here only anchoring points for the following structural layers are opened. Most of the wafer will remain covered with the PSG. This is done to avoid problems with overetching of underlaying layers during patterning of the structural layers. The parts of the structural layers that must be removed are situated on top of the PSG layer. However, the opening of the contact holes is done only once for both the structural silicon nitride layer and for the second polysilicon layer, so some problems remain as will be described in the sections on the structural layers.

Another problem when opening the anchor points by plasma etching, is the impossibility to stop on the first sacrificial layer as this consists of the same material. Although this is not an issue for the anchoring of beams and bridges, it can present problems when structures comprising of combinations of the first polysilicon and the second silicon nitride and polysilicon are to be used, as it is impossible to keep these structures level. To avoid some of these problems, the patterning of the second sacrificial layer is timed so that approximately half of the first sacrificial layer will remain. The material damage in the top of the first PSG layer has been removed by the 575°C poly deposition and 605°C anneal step. If this had not been the case, the rapid etching of this intermediate layer could cause underetching of the PSG layer, resulting in stringers trapped between the PSG layers. Measurements have shown that approximately 150 to 200 nm is lost from the first sacrificial PSG

![Cross-section after deposition and patterning of the second sacrificial layer.](image-url)
layer, when an opening is etched in the second sacrificial PSG layer. The second sacrificial layer and the anchor points are shown in Figure 4-18.

4.5.7 Silicon nitride

The structural silicon nitride layer is again a low-stress, silicon rich, nitride layer as was used for the protection layer covering the electronics (see section 4.5.2). This layer forms an insulation between the two conductive layers and can be used as a non-conducting mechanical layer. It is placed over the second sacrificial layer to enable the construction of a stiff upper electrode in differential capacitance measurements. The layer is deposited to a thickness of 300 nm. The high deposition temperature of 850°C at the same time serves to activate the dopants in the first polysilicon layer.

The silicon nitride layer is patterned using a different plasma etch as the one used to open the contact holes in the first nitride layer. The etch is done at 60 Watt, 0.05 mBar, 10 sccm SF6, 70 sccm CH4, and 10 sccm O2. This plasma etch is very anisotropic and again a problem can occur with the stringers along the side walls. To remove these stringers, considerable overetching is needed (in the order of twice the nominal etch time), but in contrast to the plasma etch used to etch the first nitride layer this plasma etch has some selectivity over silicon oxide (1:1.7). This limited selectivity causes some loss of the sacrificial layer, again causing differences in the gap between second nitride and second polysilicon. Although plasma etch chemistries are available with much higher selectivity [4.65], [4.66], these can cause problems with the contact resistance between the first polysilicon layer and the polysilicon layers following the nitride [4.67]. This is due to the formation of polymer films that can increase the contact resistance to several mega ohms. The O2 in the plasma used is believed to remove this layer as it is formed, reducing the selectivity, but ensuring a good electrical contact.

Another problem might arise that could affect the contact holes already present in the sacrificial layer for the anchoring of the second polysilicon layer, as these are not covered with the silicon nitride during etching. If these openings are on the protective nitride layer, there is no selectivity between this layer and the structural nitride layer during etching. Because of the long overetching needed, the protective layer would be etched through, revealing the oxide underneath. To avoid this, all the openings in the second sacrificial layer must be covered with silicon nitride, situated on the first polysilicon or on the first sacrificial
layer. These design rules put hardly any constraints on the device design, and solves overetching problems and problems with stringers along sidewalls. The structure is shown in Figure 4-19.

![Cross-section of the wafer after deposition of the structural silicon nitride layer.](image)

Fig. 4-19 Cross-section of the wafer after deposition of the structural silicon nitride layer.

### 4.5.8 Second polycrystalline silicon

The second polysilicon layer is a low-stress polysilicon layer deposited at the same temperature as the first polysilicon layer (see section 4.5.5). This layer is also 300 nm thick and implanted with phosphorus. To activate the dopants an extra anneal step is needed of 850°C for 30 minutes, as no high temperature steps follow this deposition.

Before deposition, the wafer needs to be cleaned to ensure good mechanical and electrical contact between the first polysilicon and the second polysilicon. To this end, a short dip etch in 1% HF is done to remove the native oxide from the first polysilicon layer. Obviously, care should be taken to minimize attack of the PSG. Contact resistance between the first and the second polysilicon with a contact area of 4 μm by 4 μm is in the order of 80 Ω.

The polysilicon is then patterned using the same, partially isotropic plasma etch used to pattern the first polysilicon layer. Again, a considerable over-etch is needed to remove the stringers along steps. These stringers not only appear along the steps in the sacrificial layer, but around all steps, including the structural layers. Therefore, the removal of stringers of the second polysilicon is very important, to avoid short circuits and mechanical connections between structures, etc. As some of these stringers would be free-standing, these can break off during processing and end up on a different part of the wafer, where it can cause short-circuits or get lodged in openings and obstruct movement of mechanical structures. Because of the presence of the second sacrificial layer, over-etching can be long. Selectivity of the etch between polysilicon and nitride is 2.6:1, which ensures a loss of no more than
approximately 50 nm of nitride. Overetching of the sacrificial layer is not relevant, as this is the last structural layer to be deposited on top of it.

This polysilicon layer must cover all anchoring points in the second sacrificial layer that have not been covered by the nitride layer. In any openings where polysilicon 1 is present but polysilicon 2 is etched, the first polysilicon will be etched almost all the way through, due to the long over-etching. Because of the underlaying nitride layer, this causes no other problems. However, the large step to a thickness of both polysilicon layers (760 nm) makes it unlikely that a reliable electrical contact can be formed with the interconnect layer, which has a poor step coverage. The structure and the anchoring of the second polysilicon layer is shown in Figure 4-20.

![Cross-section of polysilicon layer](image)

**Fig. 4-20 Cross-section after deposition of the second polysilicon layer, showing the anchoring on the left.**

### 4.5.9 Protection layer

Following the deposition of the structural layers, provisions are made for the deposition and patterning of the aluminium interconnect.

One problem that needs to be solved is caused by the 1% silicon that is present in the aluminium to avoid junction spiking [4.68]. During patterning of the aluminium this silicon needs to be removed as well, or it would leave grains of silicon on the surface, which could cause mechanical problems or short circuits. The process used to pattern the aluminium will be discussed further in section 4.5.11. No matter what type of etch used, the interconnect patterning etches both aluminium and silicon. Therefore, it would etch the polycrystalline structures present on the wafer as well.

To avoid any problems, an intermediate layer is needed to serve as protection for the polycrystalline structures. A PECVD low temperature silicon oxide layer is used for this purpose, as this material can protect the structural layers the same way an oxide protects the electronics during the same process step and it can easily be removed from the
Structural layers. However, before such a layer can be deposited, the sacrificial PSG that is still present on most of the wafer is removed. This ensures that the oxide layer will be the same thickness and composition all over the wafer, which is important when etching the contact holes and anchoring windows through this layer (see section 4.5.10 and 4.5.11). This is done by a short etch in BHF which will remove all PSG from the surface. This etch needs to be timed correctly to avoid large underetching of the structures, which could cause problems with resist or aluminium getting lodged underneath the structures, or, in the extreme case where it would completely underetched some structures, free-standing structures could break during subsequent processing. Figure 4-21 shows the structure after removal of the PSG from the surface of the wafer.

![Cross-section after removal of the PSG from the surface.](image)

After removal of the sacrificial layer, a 300 nm TEOS (Tetra-Ethyl-Ortho-Silicate) layer is deposited using PECVD at 350°C. This thickness is needed to make sure all polysilicon is covered with sufficiently thick oxide. TEOS has a good step coverage [4.69] and caused no problems during etching of the interconnect.

The TEOS layer is then patterned using a plasma etch. This is done to provide anchoring points for the aluminium interconnect on the nitride, to make contact between the polysilicon and the aluminium and to reduce the step height in contact holes. If this had not been done, the aluminium would be situated on top of the TEOS layer. As will be shown in section 4.5.15 on sacrificial etching, this can result in free-standing aluminium layers. If the oxide around contact holes to the silicon is not cleared, the depth of the contact holes would be close to 1 μm, which is too much for the aluminium to cover reliably. The oxide also needs to be cleared on parts of the polysilicon, where the aluminium makes contact, although care should be taken that the aluminium overlaps the oxide edges.

Etching is done using a plasma etch at 60 Watts and 0.05 mBar with 25 sccm CHF₃, 50 sccm CF₄ and 40 sccm He. This results in a high
selectivity on polysilicon (1:4), but a limited selectivity on silicon nitride (1:1.1). This causes the protective nitride layer covering the wafer, to be etched approximately 25 nm in addition to earlier etching. At critical places the nitride layer is now reduced to a thickness of only 50 nm. The structure after deposition of the TEOS layer is shown in Figure 4-22.

![Cross-section](image)

**Fig. 4-22 Cross-section of the wafer after deposition of a TEOS layer to protect the polysilicon structures.**

### 4.5.10 Contact holes

Contact holes are opened in the nitride/oxide layer to provide contact between the active devices and diffusions in the epitaxial layer and the interconnect. This etch is done in two steps, similar to the contact holes opened for the first polysilicon layer. The first step uses the same plasma etch as the one used to pattern the TEOS: 60 Watts at 0.05 mBar with 25 sccm CHF$_3$, 50 sccm CF$_4$ and 40 sccm He. It removes the nitride layer and part of the thermal oxide underneath. This is followed by a BHF dip etch to remove the remaining oxide from the silicon surface. This combination results in steep sides with very little underetching and a clean silicon surface. The thickness of the nitride in the area where the contact holes are situated is now approximately 100 nm due to overetching. The resulting structure is shown in Figure 4-23.

![Cross-section](image)

**Fig. 4-23 Cross-section showing the structure after opening of the contact holes to provide contact to the electronics.**
4.5.11 Interconnect

The interconnect consists of aluminium with 1% silicon to avoid junction spiking [4.68]. This layer is the same as that used for the standard microelectronic process and is sputtered to a thickness of 800 nm. This is thicker than the 600 nm used for electronic processing to ensure good contact in the deeper contact holes and on the polysilicon contacts.

Electrical contact with the mechanical polysilicon layers can be achieved through contacts on the first polysilicon or through doped regions in the substrate. The second polysilicon layer overlaps the edges of the anchoring points in the sacrificial layer and as a result leaves a gap of 380 nm along all edges in the form of wings. Although direct contacts on the second polysilicon are possible by stacking polysilicon one, silicon nitride and polysilicon two, contacts through the first polysilicon are the simplest.

Patterning of the interconnect in the standard microelectronic process is done by wet etching. The etchant etches the aluminium and the silicon present in the interconnect is left on the wafer as silicon granules. These granules need to be etched in a polysilicon etchant.

As was mentioned in section 4.5.9 a problem arises when this is done with the surface micromachined structures present. The wet etchant used to remove the silicon, also etches silicon oxide. Because of the limited step coverage of the TEOS used as a protective layer, break-through can occur on the edges where the oxide is thin. If there is polysilicon underneath, this will be etched and can result in breaking of the structures. A thicker TEOS layer could solve this problem, but might cause step coverage problems of the interconnect itself.

Therefore, in the surface micromachining module, the interconnect is patterned using plasma etching. The etch consists of two steps. The first is a fast etch with no selectivity on silicon or oxide. It consists of: 40 Watts, 0.03 mBar, 7 sccm Cl₂, 22 sccm BCl₃ and 100 sccm N₂. After most of the aluminium and the silicon granules are removed, the chemistry is changed to: 100 sccm N₂ and 60 sccm SF₆, which has some selectivity on oxide. Because this is highly directional, the etch will only attack the thick, level portions of the TEOS layer, so there are no problems with etching of the mechanical structures.
Fig. 4-24 *SEMs showing the remains of the aluminium along the steps of the structure and the structure after a short wet aluminium etch.*

However, this anisotropic etching leaves remains of the aluminium along the edges of steps, due to the effects similar to those resulting in the polysilicon and nitride stringers that were described in section 4.5.3. To remove these pieces, a short, wet aluminium etch is used (see Figure 4-24). In the standard microelectronic process, the mask is scaled up, to compensate for the loss in width caused by the underetching of the wet etch normally used. When using an extra wet etch to remove flakes of aluminium in the surface micromachining process, it can be timed so that the same underetching will occur and the same scaling is necessary. This enables us to use the same IC mask used for the electronics in normal microelectronic processing and when combined with the surface micromachining module. Figure 4-25 shows the cross-section of the wafer after patterning of the aluminium.

![Fig. 4-25 Cross-section after patterning of the interconnect. The resist used in the patterning is not removed.](image)
4.5.12 Oxide removal

After the patterning of the interconnect, the TEOS is removed. This is done using a short etch in BHF. This leaves the sides of the aluminium that covered the TEOS to form an overhang as shown in Figure 4-26.

Fig. 4-26 Wing shaped overlap of the interconnect on a contact on polysilicon.

Although aluminium can withstand a short HF etch [4.70], the aluminium present on the wafer is protected from the etchant by the resist used in patterning the interconnect, which is not removed. However, to ensure a good protection, the resist is baked in a furnace at 110°C for 30 minutes. This dries and hardens the layer after all the processing steps it has endured. The present state of processing of the structures is shown in Figure 4-27.

Fig. 4-27 Cross-section of the wafer after removal of the TEOS layer, using the IC mask to protect the aluminium from the etchant.

4.5.13 Sintering

After the TEOS is removed, the resist is stripped and the aluminium is sintered at 400°C for 30 minutes. Any native oxides on the silicon are
Micromechanical system fabrication

Fig. 4-28 Stress in aluminium interconnect before (left) and after (right) sintering.

removed by aluminium reduction leaving a low-resistance contact on both silicon and polysilicon. The sintering also changes the stress in the aluminium layer from compressive to tensile which is clearly shown in Figure 4-28. Notice the breaking of several stretches of interconnect due to the stress. Figure 4-29 shows the cross section of the wafer after removal of the resist and sintering.

Fig. 4-29 Cross-section of the wafer after deposition and sintering of the interconnect are completed.

4.5.14 Wafer dicing

Dicing of a wafer with free-standing surface micromachined structures requires special attention. With normal (electronic) wafers the wafer is placed on an adhesive plastic sheet, which will hold all the dies in place and sawed. This saw is cooled and lubricated using a jet of water. This water jet would destroy the micromachined structures and can cause sticking (see section 4.5.16).

Therefore the wafers are diced before the release of the structures with the resist layer still present to protect the aluminium interconnect from the etchant still applied (see section 4.5.15.1). The resist and the
fact that the structures are not yet released ensures these are not damaged by the sawing of the wafer. No processing steps involving resist patterning, deposition of layers or plasma etching follow, so all subsequent steps can be done using individual dies.

Alternatively, the sawing need not go all the way through the wafer. Then processing steps can be done on a whole wafer to speed up the overall processing and individual dies can be broken off without risk of damaging the structures. The latter approach has the advantage of maintaining batch processing.

4.5.15 Sacrificial etching

The last step in the surface micromachining is the release of the structures by the selective etching of the PSG sacrificial layers. This is the most important step, after which the structures become free-standing and no extra processing can be performed. This is also the step in which many problems arise and is therefore, in a sense, the most critical step in the entire surface micromachining process.

4.5.15.1 Aluminium protection

Etching of the PSG is normally done using hydrofluoric acid (HF), as this has a large selectivity between the oxide and the polysilicon and silicon nitride, which form the mechanical structure. Moreover, it has a high etch rate, which is required to underetch the structures in a reasonable time. However, HF also etches aluminium, so some sort of protection must be found for the interconnect layer present on the wafer. This can take the form of a protection layer covering the interconnect and preventing contact between the etchant and the aluminium.

This layer could be PECVD silicon nitride, which can be deposited at 300°C without damaging the aluminium interconnect. However, this would complicate processing considerably. It would require two extra depositions (an oxide intermediate layer to enable the selective removal of the layer and the nitride layer itself) and three extra masks (one to anchor the nitride on the wafer and to prevent the etchant from underetching the nitride, one to open a window in the nitride above the mechanical structures and to open a window over the contact pads, and one mask to protect the contact openings during etching). The problem encountered by the opening of the contact pads is again the same problem as that faced before: exposed aluminium during sacrificial etching. Furthermore, PECVD nitride shows many mechanical defects in
the form of pinholes. Through these pinholes the etchant could reach the aluminium and ruin the interconnect. In addition, the etch rate of PECVD nitride is much larger than that of LPCVD nitride [4.71]. An experiment showed an etch rate in 20% HF of 45nm/min and 7nm/min in BHF. This presents a problem along edges after only 5 minutes, as the step coverage of the PECVD nitride is poor. Therefore, the use of silicon nitride as a shielding layer does not provide an adequate solution.

The second possible layer that can be used to protect the aluminium is photoresist [4.72]. This layer requires just one mask and can be removed afterwards, without dangers of etching the mechanical structures or the aluminium. However, photoresist is not as strong as silicon nitride and several problems arise.

**Photoresist etching**

The first problem that arises is the etching of the photoresist itself in HF. The etch rate of the resist is in the order of 50 nm/min in 40% HF. In combination with leaking through the layer, this limits etch time to approximately 10 minutes.

Modern photolithography is optimized for microelectronic technologies where patterning is done exclusively by plasma etching. The bake at 90°C for 1 minute used, is insufficient to remove all the solvents present in the resist, leaving the resist soft and porous. This is sufficient for plasma etches, but not for long wet etches. An extra bake step of 110°C for 30 minutes is used to remove all solvents and harden the resist.

Combining this with the use of Buffered Hydrofluoric acid (BHF), which etches the resist slower than pure HF, the time the photoresist can withstand the etchant is increased to two hours. However, the BHF etches the PSG slower as well. If more time is needed, intermediate drying of the resist by a short bake, will lengthen this time even further.

**Photoresist adhesion**

The adhesion of the photoresist layer to the surface of the chip is another problem [4.73]. The weak bond can cause the resist to separate from the surface, allowing etchant to seep underneath and attack the underlaying
Fig. 4-30 Etching of the interconnect due to separation of the resist and the substrate (a) and further separation due to expansion of attacked aluminium (b). Aluminium (see Figure 4-30 a). The porous nature of the aluminium seems to absorb the etchant and the aluminium expands. This pushes the resist upward, allowing new etchant to reach the aluminium. Figure 4-30 b shows this effect. The areas around the attacked aluminium are lighter in colour, indicating a separation between the photoresist and the wafer.

Furthermore, as the bond between resist and aluminium is as good as the bond between aluminium and nitride, the peeling of the resist can cause the interconnect to be ripped off the wafer, resulting in broken wires (see Figure 4-31).

Fig. 4-31 Peeling of the interconnect due to separation of the resist.
Micromechanical system fabrication

Separation of the photoresist and the wafer can be observed after about 10 minutes and is a critical problem, as this time interval is insufficient to release most structures in a universal surface micromachining process.

To improve adhesion the wafer is dehydrated by a short bake before the resist is deposited. This shows very little improvement as most of the surface is covered with silicon nitride.

Different adhesion promoters were tried in order to improve the adhesion. The promoter HMDS that is used as standard in the facility used, did not show any difference with wafers without the adhesion promoter. An alternative promoter (TCPS) was also used. Although this showed a small improvement, separation of the resist was observed within 20 minutes. The adhesion promoters used are designed to bind to the resist to the dangling hydrogen bonds of a silicon oxide layer, which is most commonly used as a patterning layer in the fabrication of electronic devices. However, the layer on which the resist is deposited after the micromachining has taken place, consists of either nitride or silicon, which lack these hydrogen bonds.

Tests have been done using a very thin TEOS layers on the entire wafer to promote adhesion. No peeling of the resist has been observed when using this oxide. However, underetching of this layer can cause interconnect layers close enough to the etch cavities to be attacked, even when using layers as thin as 30nm. Underetching of this thin layer is as fast as, or even faster than the etching of the sacrificial layer due to the poor quality of the plasma oxide.

To reduce the underetching, the oxide layer was etched for a short time, removing the oxide from the etch window and allowing a slight underetching of the edges. Heating at 150°C for 30 minutes causes the resist to reflow and seal the access to the oxide (see Figure 4-32). This increases the possible etch time, but is insufficient to protect the oxide and the aluminium indefinitely, as the seal will also start to leak. This will occur after around 15 minutes, but it reduces the underetching considerably due to the very limited admission of fresh etchant in the cavity. The underetching rate of the resist mask is in the same order as the underetching rate of the structures.

Even though the etchants reach the aluminium at a certain moment, etching of the interconnect is much slower when the oxide layer is present. By the time the etchant reaches the aluminium, it is almost
saturated with reaction products of the etching of the oxide and therefore, much less concentrated. Transport of the reaction products takes place by the slow process of diffusion. Because of this, the aluminium will be unaffected even after several minutes in contact with the BHF. Etches of 30-45 minutes are possible, depending on the distance between the interconnect and the edge of the etch mask.

4.5.15.2 Alternative etchants

The above mentioned photoresist layer is not sufficient to provide adequate protection for the interconnect, unless the limited etch times indicated are sufficient. Another possible solution is to increase the selectivity between aluminium and the sacrificial oxide by modifying the HF entchant or changing the type of etchant.

Ways to increase the selectivity of BHF reported in literature are by the addition of glycerine or glycerol. Etch rates reported in literature of 0.55 nm/min. for aluminium and an etch rate of 95 nm/min. for thermal oxide when using the glycerol (40g NH₄F + 60ml H₂O + 20ml HF + 40ml glycerol) [4.75]. The etch rate when using glycerine is reported to be 6 nm/min. for aluminium and 200 nm/min. for thermal oxide (4 parts NH₄F + 1 part HF (47%) + 2 parts glycerine (87%)) [4.74].

Experiments done using these mixtures did not reproduce these results and considerable etching of the interconnect was observed.
Therefore, another possible sacrificial etchant was investigated. This is the, so called, pad-etchants\(^1\) used to open contact holes in oxides on contact pads [4.76]. It consists of 1/3 CH\(_3\)COOH (acetic acid) and 1/3 NH\(_4\)F and 1/3 H\(_2\)O. Etch rates of 5 nm/min. for aluminium and 200 nm/min. for PSG are found.

All etchants mentioned above etch the sacrificial material slower than HF or BHF and thus longer etch times are required. Even though the etch rates of the aluminium are very low, the long underetching needed causes some loss of aluminium thickness. For the pad etchant for example, an underetching of 10 \(\mu\)m of PSG will lead to 250 nm of aluminium being etched.

![Resist](image)

Fig. 4-33 Cross-section after patterning of the resist layer that is to protect the aluminium interconnect during sacrificial etching.

To make sure the protection is sufficient for the underetching of large structures (in the order of 30 - 50 \(\mu\)m underetching) a combination of the methods mentioned above is used. The wafer is dehydrated and the resist deposited, baked and patterned. The cross-section of the process wafer is shown in Figure 4-33. Subsequently, the entire wafer was etched in pad-etchant for up to 90 minutes, without problems with the interconnect. The wafer with the resist mask still present is shown in

![Wafer](image)

Fig. 4-34 Cross-section of the wafer after sacrificial etching.

1. Riedel-de Haen AEW 33-33-33
Figure 4-34. Special TEOS adhesion layers and underetching and reflow of the resist is not needed, as 90 minutes is sufficiently long to enable the freeing of the structures.

4.5.15.3 Anchoring of the interconnect

As described in section 4.5.9, the interconnect is situated on top of an silicon oxide layer to protect the mechanical structures during a polysilicon dip etch, which is used to remove the silicon ganules that remain after the etching of the aluminium.

This oxide layer will be etched if the etchant seeps this far under the resist protection or if no photoresist protection is used. This can result in freestanding aluminium wires when an etchant with sufficient selectivity is used (see Figure 4-35). The large tensile stress present in the aluminium after sintering can cause these connections to break. As there is already a mask to etch openings in the oxide layer around contact holes (see section 4.5.9), this must be used to anchor the interconnect layer to the substrate. For the electronics this can be done automatically from the interconnect pattern, if circuits from libraries are used or if polysilicon is not present, all TEOS can be removed.

Fig. 4-35 Free-standing interconnect, due to underetching of the TEOS protection layer.
4.5.16 Rinsing and drying

After the etching of the sacrificial layer, the wafer must be rinsed thoroughly in DI-water to remove all traces of the etchant, even underneath the resist layer. This must be done carefully to make sure the structures, which are now free-standing, are not damaged by the flowing water.

After the rinsing the wafer must be dried. This can cause considerable problems due to an effect known as sticking, which causes the structures to become permanently stuck to the substrate. As this ruins the devices, special care must be taken to avoid this.

4.5.16.1 Sticking

Sticking causes free-standing beams, bridges and plates to become permanently stuck to the substrate, or to each other, which inhibits their movement. Several causes for sticking have been mentioned in literature. These include; glueing by dirt, surface charges, hydrogen bridging, solid and liquid bridging and Van der Waals forces between the very flat surfaces [4.77]-[4.79]. No real understanding of the sticking effect has yet been found and observations indicate a combination of several of the aforementioned causes.

Whatever causes the effect, two surfaces that come into close contact will not separate again. Therefore, precautions must be taken to avoid such contact whenever possible. The main reason for the structures to come into contact, is the drying of the liquid in which these have been submerged [4.80], [4.81]. After the wafer has been removed from the liquid, pockets of the liquid will be trapped underneath and between structures. When the droplets evaporate, they shrink. However, the surface tension of the liquid minimizes the surface and introduces a force

![Diagram](image)

Fig. 4-36 Drying of the liquid trapped underneath the structure, causing sticking.
Surface micromachining module

acting on the surfaces between which it is trapped. These forces pull on
the structure under which the droplet is trapped. As the droplet dries, the
structure is pulled towards the substrate, until the droplet has evaporated
completely and the structure is stuck against the substrate. The process is
illustrated in Figure 4-36.

Several factors determine the probability of a structure to survive this
drying. First of all, the distance between the structure and the substrate.
The larger this gap, the larger the force necessary to pull the structure all
the way down and the lower the number of structures that will stick. The
more flexible the structure (thinner, longer), the more easily it can be
pulled down. The higher the surface tension of the liquid, the more force
is created to pull the structure down. The shape of the structure
influences the way the droplet dries and therefore, how the forces caused
by the drying droplet are distributed. The smaller the contact area, the
lower the sticking force.

These factors also give us the means to reduce the sticking through
use of special techniques. Several techniques have been used in
literature. These include replacing the liquid with one with a lower
surface tension [4.77]-[4.79], drying at elevated temperatures to reduce
the surface tension [4.82], freezing the liquid and then removing it by
sublimation in vacuum or in a nitrogen atmosphere [4.31], [4.83]-[4.85],
holding the structures in position with temporary beams [4.86], [4.87] or
a photoresist grid pattern which is then removed by an oxide plasma strip
etch after drying [4.88]-[4.91], adding shapes to the designs to force
more advantageous drying patterns [4.92], the inclusion of bumps and
dimples in the design to reduce the contact area [4.83], [4.93] and special
surface treatment to remove/reduce the sticking force [4.77], [4.78],
[4.94]. The latter techniques have the advantage that if also prevents
sticking of a device when in operation.

Two methods were examined for this surface micromachining module.
The most simple one; replacing the water underneath the structures by a
liquid with a lower surface tension, resulted in the following procedure:

After rinsing in water, the liquid underneath the structures
is replaced by rinsing, first in acetone and then in
iso-propyl alcohol (IPA) without allowing if to dry in
between. The IPA is then dried on a hotplate. The best
result was achieved at 46°C. The surface tension of the IPA

123
is smaller than that of water and the temperature lowers the surface tension even more and reduces sticking.

Although an improvement was observed, this simple technique did not yield the desired results with the often large structures and the small gap. As a next step, the liquid phase was avoided altogether, using a simple freeze drying technique, suggested by Legtenberg [4,84]:

After rinsing in water, the liquid is replaced by cyclo-hexane. This is then cooled to approximately -5°C on a Peltier element, at which temperature the cyclo-hexane freezes. The chip is then placed in a nitrogen flow in which the cyclo-hexane sublimes, leaving the structures free-standing.

This technique is faster than alternative freeze drying techniques and does not require a vacuum chamber.

From these two the second gives the best results. With this technique it is possible to release polysilicon bridges (beams supported on both sides) of 5 μm width and up to 150 μm in length with a thickness of 300 nm and a gap of 380 nm. Figure 4-37 shows a set of nitride cantilevers (beams supported on one side) of increasing length. As can be seen in SEM a, beams up to a length of 55 μm are freestanding using the

![Fig. 4-37 The influence of drying techniques on sticking. SEM a shows a freeze dried sample. SEM b shows the same structure rinsed in water.](image)
first method. SEM b shows the same structure simply rinsed in water, in which case beams of 30 \( \mu \text{m} \) are already sticking.

In addition to special processing techniques the drying pattern of the liquid can be influenced by the structure design [4.92]. If the liquid dries underneath a cantilever beam there are several ways this can take place. The liquid can dry from the sides and reduce in volume evenly along the length of the beam (Figure 4-38 b). In this case, the surface tension force acts on the entire beam and sticking becomes very likely. The liquid can also dry from the side and break up into two droplets (Figure 4-38 c). One at the base of the beam and one at the end. The droplet at the base is not strong enough to cause any problems, as the beam is very stiff at this point. The droplet at the end of the beam, however, acts on the most flexible part of the beam and can easily pull down the cantilever causing it to stick. The last pattern in which the liquid can dry, is from the sides and the end of the beam (Figure 4-38 d). The droplet then pulls away from the tip of the beam towards the base of the cantilever, as the volume of the droplet reduces. This only leaves a droplet at the base of the beam, which does not cause much sticking.

![Fig. 4-38 Possible drying patterns of the trapped liquid.](image)

By shaping the structures, it is possible in a number of cases, to influence the drying pattern of the liquid and thereby reduce the sticking. E.g. the drying of the droplet underneath a cantilever can be forced into the last pattern (droplet shrinking towards the base of the cantilever), by placing a thin beam at the end of the cantilever. The drying of the thin beam will be of the desired kind and once the droplet dries in this manner, it will continue to do so underneath the wider beam.
Micromechanical system fabrication

In Figure 4-39 the completed structure is shown after sacrificial etching, removal of the resist mask and drying of the wafer.

Fig. 4-39 Cross-section of the finished, free-standing surface micromachined structures with electronics present on the same wafer.

4.6 References


References


References


References


Micromechanical system fabrication


Micromechanical system fabrication
5.1 Introduction

As was indicated in section 3.4.1.3 mechanical stress in a thin film can have a large influence on the behaviour of a mechanical structure. To be able to predict the mechanical characteristics of the devices, the amount of stress must be known. The stress level in the mechanical layers must be minimised during processing and preferably be tensile.

Determination of the level of stress in the final structure proved vital in the development of the compatible micromachining process described in chapter 4. Therefore, considerable time was devoted to the development of accurate measurement or this mechanical stress.

5.2 Stress measurement techniques

With respect to the design and realisation of surface micromachined structures and compatible processes, two different functions can be distinguished; first, the determination of stress in thin layers during process development and second, during fabrication of the devices. The first application is not dimensionally constraint and does not put many
special demands on the mechanisms used. The method can be relatively complicated and slow, but must be reliable and valid for all layers used. For the second application it must be capable of measuring the stress in all mechanical layers on the same wafer the structures are situated on, to monitor the stress. This is similar to the test structures generally used to check electrical quantities in microelectronic processing. In this case, the technique must be simple and convenient, as many chips must be checked and it can not take up a large area on the wafer or chip in question.

For the measurement of stress in thin films several methods can be used. These make use of three different effects caused by the stress, viz. changes in mechanical properties, such as spring constants and resonance frequencies, changes is size and buckling; the deformation that occurs when a beam is compressed above a certain level.

5.2.1 Resonance frequency technique

In this technique, changes in resonance frequency are detected [5.1]-[5.4]. Similar to the shift in tone of a guitar string as it is tuned by stretching, the resonance frequency of a beam under tensile stress will be higher than that of the unstressed beam. In the same manner, compressive stress will cause the resonance frequency to drop. For a simply supported beam, the resonance frequency, $\omega$, can be calculated as a function of strain $\varepsilon$ [5.5]:

$$\omega = \sqrt{\frac{(n\pi)^2 EI}{\rho AL} - \frac{\varepsilon AL^2}{I}}$$

where $n = 1, 2, 3, \ldots$ gives the mode of vibration, $\rho$ is the density of the material, $E$ is the Young's modulus, $I$ is the moment of inertia and $A$ is the cross section of the beam. For a double clamped beam an analytical solution can not be found.

For this method, the resonating structures need to be excited, which can be achieved using a vibration table, acoustically, thermally, optically or, for conducting structures, electrically. Measurement of the deflection can be done optically using laser interferometry, through a microscope with stroboscopic illumination, or by means of capacitance changes in the case of electrically excited structure.

Apart from the excitation necessary, a problem arises when the structure is in air. The air will function as a damping layer at lower
frequencies and as a spring at high frequencies [5.6] unless special
designs are considered [5.7]. Therefore, the only way accurate frequency
measurements can be obtained is in a vacuum. The need for an extensive
measurement setup in vacuum, makes this a very complicated method for
the measurement of stress.

5.2.2 Load deflection

As was shown in section 3.4.1.3, beams under stress behave differently
from unstressed beams when loaded. Tensile stress in the beam causes
the deflection due to a load to decrease, whereas compressive stress
increases the deflection. By measuring the deflection of a beam or
membrane under a known load, the stress of the beam can be calculated
[5.2], [5.8]-[5.10].

To determine the stress accurately, all parameters except the stress
must be known and a specifically determined load must be applied. In
addition, the deflection must be measured accurately. For this special
structures and special equipment are needed and material properties such
as the Young’s modulus must be known. All these conditions make this
method of measuring stress difficult and inadequate for the monitoring of
stress during processing.

5.2.3 Wafer curvature

The method most commonly used in IC technology, is the method of
wafer curvature [5.11]-[5.13]. When a layer is deposited on top of a clean
silicon wafer, the stress in the thin film will cause the whole wafer to
curve up or down, depending on the stress. Figure 5-1 shows the result.

Fig. 5-1 Curvature of a wafer due to compressive stress (a) and tensile
stress (b).

When the film is compressive, it causes a convex surface, if the stress is
tensile, the surface is concave. By measuring the curvature, the stress in
the thin film can be calculated. As the wafer will have a curvature before
the film is deposited, the curvature must be measured on the same wafer
Stress measurement

both before and after deposition of the film. The stress in the film can then be calculated using the Stoney formula [5.11]:

\[ \sigma_0 = \frac{E_f t_f^2}{6(1-v_s)t_s} \left( \frac{1}{R_0} - \frac{1}{R_f} \right). \]  

(5-2)

where \( E_s \) and \( v_s \) are the Young's modulus and Poison's ratio of the substrate respectively, \( t_s \) is the thickness of the substrate, \( t_f \) is the thickness of the film, \( R_0 \) is the radius of the curvature before deposition of the film and \( R_f \) is the curvature radius after deposition.

The disadvantage of this method is that it only works on an unpatterned sheet of film and not on the structures. Moreover, material properties and thicknesses must be measured to accurately determine the stress. Extra processing and different layer interfaces can cause changes in stress that are not apparent from the wafer curvature, which is measured on a separate process control wafer. It also makes it difficult to determine the stress distributions over the wafer. It can therefore be useful during process development, but can not be used to monitor stress on process wafers during fabrication.

5.2.4 Beam curvature

A technique using a method similar to the wafer curvature method described previously, is the bending of cantilever beams [5.14], [5.15]. Beams constructed by micromachining techniques out of the bulk material are covered by the thin film that is to be measured. In the same manner as the wafer is curved, the beam will bend upwards, if the stress is tensile, or downwards, if the stress is compressive (see Figure 5-2). Due to the much thinner base, the bending due to the stress is much more pronounced than that of a wafer. Therefore, much smaller structures can be used.

The stress level can be determined by measurement of the deflection of the tip of a beam of known length, or by direct measurement of the curvature. The latter method is more accurate, but more complicated. The advantages over full wafer curvature is obvious; measurements can be done on the same wafer the structures are constructed in and the measurement structures are comparatively small. However, the thickness of both the film and the base must be known, as well as the material properties. To measure the curvature or deflection,
Stress measurement techniques

Fig. 5-2  *Curvature of a cantilever beam due to compressive stress (a) and tensile stress (b).*

often the wafer is broken, so that the structure can be viewed from the side. This would make it impossible to use this method to monitor stress in a production run, or the change in stress due to subsequent processing.

5.2.5 Buckling

This method makes use of the effect caused by the expansion or contraction of structures when they become free-standing [5.16], [5.17]. When a double supported beam is released by removing the sacrificial layer, stress will partially be released by expansion or contraction. When trying to expand, a beam that is clamped at both sides, will release some of its stress by bending out of the horizontal plane. This is called buckling and will occur at a certain level of stress for a certain length of beam. Buckling structures are essentially measuring increases in length with a built-in threshold and thus do not measure stress, but the strain that is the result of the stress.

The strain at which the beams buckle can be determined using several methods including the deflection under a load of a stressed beam [5.18] and the zero-frequency resonance mode of the beam [5.16]. Here an energy method as this provides more insight in the buckling behaviour.

The potential energy in the beam is minimized to determine its static situation for a certain strain. This is different for beams with different boundary conditions. Figure 5-3 shows simply supported and clamped beams that will be considered here, in a buckled situation.

The potential energy is contained in the axial strain, i.e. the shortening of the beam, and the bending strain. The axial strain of a beam which is unbuckled is $\varepsilon_0$. The axial strain of the buckled beam can be
Stress measurement

Fig. 5-3  Buckling of a beam with simply supported anchoring points (a) and clamped anchoring points (b).

found by calculating the elongation of the beam due to the buckle. This elongation is given by:

\[ \Delta L = \frac{1}{2J} \int_0^L \left( \frac{dv}{dx} \right)^2 dx, \]  \hspace{1cm} (5-3)

where \( \Delta L \) is the elongation, \( L \) is the original length of the beam and \( v(x) \) is the deflection of the beam.

The function \( v(x) \) depends on the boundary conditions and two different situations are considered here; the simply supported beam shown in Figure 5-3 a, where the anchoring points can rotate freely, and the clamped beam in Figure 5-3 b, where the anchoring points are fixed. The deflection curves due to buckling are given by [5.18]:

\[ v_{supp}(x) = v_{max} \sin \frac{\pi x}{L} \]

\[ v_{clamp}(x) = \frac{v_{max}}{2} \left( 1 - \cos \frac{2\pi x}{L} \right) \]  \hspace{1cm} (5-4)

where \( v_{supp}(x) \) is the deflection curve of the simply supported beam, and \( v_{clamp}(x) \) of the clamped beam.

The energy in the beam due to the strain is given by:

\[ U_{strain} = \frac{1}{2} \left( \varepsilon_0 - \frac{\Delta L}{L} \right)^2 ELA, \]  \hspace{1cm} (5-5)

where \( E \) is the Young’s modulus of the material and \( A \) is the cross-sectional area.
The energy in the beam due to bending is given by:

\[ U_{\text{bending}} = \int_0^L \frac{EI_y}{2} \left( \frac{\partial^2 v}{\partial x^2} \right)^2 dx, \]  

(5-6)

where \( I_y \) is the moment of inertia.

Combining Eq. (5-4), (5-5) and (5-6) gives us the total potential energy in the buckled beams. For the simply supported beam this results in:

\[ U_{\text{pot}} = \frac{1}{2} \left[ \varepsilon_0 \left( -\frac{v_{\text{max}} \pi^2}{2L} \right)^2 \right] E L A + \frac{E w h^3 v_{\text{max}}^2 \pi^4}{12 L^3}, \]  

(5-7)

where \( w \) is the width of the beam and \( h \) is the thickness of the beam.

The minimum of this equation \((dU_{\text{pot}}/dv_{\text{max}} = 0)\), gives us the deflection \( v_{\text{max}} \) for a certain level of strain. For the simply supported beam this is:

\[ v_{\text{max}} = \pm 2 \sqrt{\frac{3 \varepsilon_0 L^2 - h^2 \pi^2}{3 \pi^2}} \]  

(5-8)

From this equation we can see that for certain strains \( \varepsilon_0 \), there is a solution for \( v_{\text{max}} \). The critical strain at which \( v_{\text{max}} > 0 \) is given by:

\[ \varepsilon_{\text{crit}} = \frac{\pi^2 h^2}{12 L^2} \]  

(5-9)

A similar calculation can be done for the clamped beam. This results in a critical strain of:

\[ \varepsilon_{\text{crit}} = \frac{\pi^2 h^2}{3 L^2} \]  

(5-10)

As can be seen from these equations, the boundary conditions set by the supports have a large influence on the stress at which a beam buckles. Therefore, it is important to accurately determine these boundary conditions in the structures used.
Stress measurement

If we analyse the potential energy equation, it can be seen that up to the critical strain, all strain is present as axial strain. Any strain above the critical strain level will be present as buckling strain. The buckling height, \( v_{\text{max}} \) is shown in Figure 5-4. Although the deflection increases rapidly when the strain increases, the total buckling height is still very small.

![Graph](image)

Fig. 5-4  Maximum deflection of a simply supported beam of length 50 \( \mu \text{m} \) and thickness 0.5 \( \mu \text{m} \), as a function of the intrinsic strain in the layer. The critical buckling strain is 329\( \times 10^{-6} \).

5.2.5.1 Beam arrays

To determine the stress in a layer, an array of beams of different length is used. At a certain stress level all beams of a particular length and longer, will buckle. The length of the shortest beam to buckle can be used to calculate the strain using Eq. (5-9) or (5-10). The number of beams in the array determines the range of the stress that can be measured and the difference in length between the beams determines the resolution. The resolution \( \delta \) is given by:

\[
\delta = \frac{\Delta \varepsilon}{\varepsilon} = -2 \frac{\Delta L}{L} \tag{5-11}
\]
Stress measurement techniques

The number of beams in an array can be quite high. The minimum number of beams needed for a certain measurement range and accuracy is given by:

\[
N_{min} = \frac{\log\left(\frac{\varepsilon_{max}}{\varepsilon_{min}}\right)}{\log(1 + \delta)} - 1
\]  

(5-12)

For example; to determine the strain over a range of \(100\cdot10^{-6}\) to \(2000\cdot10^{-6}\), with an accuracy of 5%, 61 beams are needed.

Buckling structures can be used to determine the strain in any layer that can be fabricate free standing. Determination of which beams are buckled can be done using an optical microscope. Light interference patterns can be see as the gap beneath the beam becomes greater. However, the height of the buckle is very small, so identifying that shortest buckling beam is not easy. Therefore, the inaccuracy is usually larger than the theoretical resolution.

![SEM of part of a strain measurement array. showing beams under compressive strain.](image)

Figure 5-5 shows part of a strain measurement array. As can be seen in the SEM, the beams buckle with the deflection profile of a double clamped beam. Practically all bridge structures fabricated using micromachining techniques can be considered to be clamped. The shape of the anchoring points has little effect on this and stepped connections as
shown in Figure 5-6 can also be considered clamped as long as the step is not high and the film thin in comparison to the length of the beam [5.19], [5.20].

Fig. 5-6  Curved anchoring of beam ends (a) showing clamped beam deflection profile.

5.2.5.2 Ring structure

As the buckling will only occur in bridges under compressive stress, special conversion structures are needed to convert tensile stress into compressive stress in a certain part of the structure.

Fig. 5-7  Ring structure (a) converting tensile strain into compressive strain in the middle beam, causing it to buckle (b).

A ring structure was designed by Guckel et al. [5.17], [5.21] (see Figure 5-7 a). When under tensile stress, the circle will deform into an ellipsoid and the stress in the central beam becomes compressive (see Figure 5-7 b). However, the boundary conditions of the beam are not well defined and as was shown in section 5.2.5, this has large influence on the critical buckling strain. Because of this uncertainty and the fact that the conversion is not very efficient, this structure has a large inaccuracy. Although it can be used to measure high levels of strain, it is not suited for the measurement of the low strain levels that are needed in the mechanical thin films useful for surface micromachining.
5.2.5.3 Buckling method evaluation

The fact that buckling structures determine the strain instead of the stress poses no problems. The stress can be calculated using the Young's modulus if needed and many problems in mechanical structures are directly related to strain, not stress.

Using the buckling method does however, have several disadvantages. The determination of which beam buckles is not always easy, as the buckle height only increases slowly at higher strains. This inaccuracy, plus the limited resolution, which is determined by the difference in length between the beams in the array, makes it difficult to determine the strain accurately. Uncertainty in the boundary conditions of the beam can introduce large errors.

In addition to this measurement problem, a large array is needed to reduce the inaccuracy and obtain a sufficiently large range in the measurement. This consumes considerable die area, in particular when conversion structures are needed for the measurement of tensile strain. This method can therefore be used to measure strain during development of fabrication techniques and micromachining materials, but takes up to much die area to be incorporated on all chips to monitor strain in all layers.

5.2.6 Displacement technique

This technique makes use of the conversion of the stress into a deformation of the structure. It has the advantage of only requiring one small micromachined structure. Two structures have been proposed in literature [5.22]. The so called T and H-structures.

5.2.6.1 T-structure

This structure measures tensile strain only and is shown in Figure 5-8. When the structure is released from the substrate the bar A becomes shorter in an attempt to release the strain and deforms cross-bar B into a curve. The offset in the curve, \( \delta \), is a measure of the strain:

\[
\varepsilon = \delta \cdot \left( \frac{1}{L_A} + \frac{32w_b^3}{w_A(2L_B^3 - 2w_A L_B + w_A^3)} \right),
\]

(5-13)

where the variable are given in Figure 5-8.
Stress measurement

Fig. 5-8 T-structure for the measurement of tensile and low compressive strain.

The theoretical maximum of the deflection is; $\delta = \varepsilon L_A$ which is in the order of hundredths of a micron and very difficult to measure and is therefore not very accurate. Especially for the measurement and monitoring of low strain levels which are needed in micromachining.

In addition, the measurement of compressive strain, though possible, can cause the beam A to buckle, rendering the structure useless.

5.2.6.2 H structure

The H-structure is shown in Figure 5-9 and is also designed to measure tensile strain. The wide beam A generates a larger force on the connection between the beams A and B. This causes a displacement of the connection point due to the strain. The strain, $\varepsilon$, can be expressed in terms of the displacement, $\delta$, as:

$$\varepsilon = \delta \frac{A_A L_B + A_B L_A}{L_A L_B (A_A - A_B)},$$

(5-14)

where $A_A$ is the cross section beam A, $A_B$ is the sum of the cross sections of all beams B and all other variables are indicated in Figure 5-9.

Similar to the T-structure, the H-structures yields only very small displacements, making accurate measurement of small stresses in particular, very difficult, and buckling can easily occur when compressive strain is present.
5.3 New buckling structures

To improve the applicability of the buckling structures for the measurement of tensile strain, two new structures were designed. These structures both have a more efficient conversion of the tensile strain into compressive strain than the ring structure, so the area required by the structure, used for the detection of a certain strain, is reduced.

Fig. 5-10 Diamond shaped structure under tensile (a) and compressive (b) strain showing the (exaggerated) deformation.
5.3.1 Diamond structure

To improve the accuracy of the ring structure discussed in section 5.2.5.2, the structure was modified to a diamond shaped structure which works in the same manner as the ring structures. Under tensile strain, the square is deformed and the central beam is compressed and can buckle (Figure 5-10 a). When the strain is compressive, the outside beams are compressed and can buckle (Figure 5-10 b). Extra beams between the support and the diamond improve the conversion of strain and thin beams at the sides ensure proper clamped boundary conditions of the central beam.

5.3.1.1 Modelling

In the calculations the connecting points between the beams are considered to rotate freely. Figure 5-11 shows the representation of the structure. For the complete calculation see Appendix B.

![Diagram of Diamond Structure](image)

Fig. 5-11 Geometry of the diamond structure.

The strain in side beam (3) and middle beam (4) can be calculated as:

\[
\varepsilon_3 = \varepsilon_0 \left[ 1 + \frac{(A_4-A_3)L_4}{N} + \frac{L_4A_1A_2}{N} \left( \frac{L_1N+L_4(L_3A_3+L_3A_4+L_4A_3)}{N(L_2+\frac{1}{2}\sqrt{2}L_4A_1)+L_3L_4A_1A_2} \right) \right] \quad (5-15)
\]
and

$$
\varepsilon_4 = \varepsilon_0 \left( 1 - \left( \frac{2(A_4 - A_3)L_3}{N} \right) \left( \frac{L_4A_1A_2}{N} \cdot \frac{L_1N + L_4(A_3 + A_4 + 2A_3A_4)}{N(L_1A_2 + \frac{1}{2}\sqrt{2}L_4A_1) + L_3L_4A_1A_2} \right) \right)
$$

(5-16)

where $N = 2A_4L_3 + A_3L_4$.

To illustrate the dependence of the conversion factor, $\varepsilon_4/\varepsilon_0$, Figure 5-12 shows the calculated strain in both beams 3 and 4 ($\varepsilon_3$ and $\varepsilon_4$ resp.) as a function of the cross-sectional area of one of the beams. The cross-section of the other beams are 2 $\mu$m$^2$ for beams 1 and 2, and 1 $\mu$m$^2$ for beams 3 and 4.

![Graph showing the strain in beams 3 and 4 as a function of cross-sectional area of beam x.](image)

Fig. 5-12 The strain in beams 3 and 4 as a function of the cross-sectional area of beam x. Intrinsic strain $\varepsilon_0 = 1000 \cdot 10^{-6}$, $L_1 = L_3 = L_4 = 20 \mu$m, $A_1 = A_2 = 2 \mu$m$^2$ and $A_3 = A_4 = 1 \mu$m$^2$.

Figure 5-13 shows the strain in the beams 3 and 4 ($\varepsilon_3$ and $\varepsilon_4$ resp.) as a function of the length of one of the beams. The length of the other beams remains 20 $\mu$m. From these figures we can see that the conversion factor is strongly dependent on the stiffness of beams 1 and 2 in comparison to 3 and 4. The wider these beams, the larger the conversion
factor. Increasing the length of beam 1 and beam 2 also increases the conversion factor. Although the dependence on the length of beam 2 is not shown in the graph, the length of beam 2 is automatically increased by increasing beam 4 in length and therefore shows the same dependence.

![Graph showing strain vs. beam length](image)

**Fig. 5-13** The strain in beams 3 and 4 as a function of the length of beam x. Intrinsic strain $\varepsilon_0 = 1000 \cdot 10^6$, $L_1=L_3=L_4=20 \, \mu m$, $A_1=A_2=2 \, \mu m^2$ and $A_3=A_4=1 \, \mu m^2$.

### 5.3.1.2 Finite element simulation

To check the assumption that the connection points between the beams are flexible is valid, a finite element simulation was done using ANSYS. Figure 5-14 shows the resulting strain in beams 3 and 4 as a function of the beam length $L_4$, of three different models. The wire model using flexible connections, a wire model using rigid connections and a solid model. As can be seen from this graph, for short $L_4$, the stiffness of the connections in the diamond, reduces the conversion factor and the wire model is not valid. For longer beams, however, the wire model can be used.

### 5.3.1.3 Conclusions

If the inner or outer beams are to buckle, the length of these beams must be greater than the critical length for the strain. If the ratio of the beam
New buckling structures

![Graph showing strain conversion]

Fig. 5-14 *The strain in beams 3 and 4 as a function of beam length $L_4$. Intrinsic strain $\varepsilon_0 = 1000 \times 10^{-6}$, $L_1 = L_3 = 20 \, \mu m$, $A_1 = A_2 = 2 \, \mu m^2$ and $A_3 = A_4 = 1 \, \mu m^2$.*

lengths in the diamond structure remains the same, the conversion factor remains constant. By scaling up the whole structure, we can make an array that can be used to determine the strain.

Although this structure works (see Figure 5-15) and is slightly more reliable than the ring structure, conversion of tensile strain into compressive strain is still not very efficient, so large structures, with long central beams are needed, making arrays very large. Although they can be used for test chips, strain monitoring using these structures on every surface micromachined chip is impractical.

These structures have been used to compare the performance with that of other structures, such as the rotation structures discussed in section 5.4.

### 5.3.2 Push-pull structure

Another conversion structure which was designed to improve the conversion of tensile into compressive strain and to make the overall structure smaller, is the push-pull structures shown in Figure 5-16. This structure can only measure tensile strain. When the strain is tensile, the two wide beams A, become shorter and compress the central beam B,
Stress measurement

Fig. 5-15 SEM showing the buckle of the middle beam of a diamond structure under tensile strain.

Fig. 5-16 Geometry and working of the push-pull structure.

which will buckle at the critical strain (see Figure 5-17). Strain is the central beam, \( \varepsilon_B \), is given by:

\[
\varepsilon_B = \varepsilon \cdot \frac{A_A(2L_A - L_B)}{2L_AA_B - L_BA_A},
\]

(5-17)

where \( \varepsilon \) is the strain in the film, \( A_A \) and \( A_B \) are the cross-sectional areas of beams A and B resp. and the other variables are as given in Figure 5-16.
As can be seen from Eq. (5-17), if $A_A << A_B$ and $2L_A > L_B$, the strain will change sign. Conversion is very efficient and the structures need not be as large as the ring or diamond structures. However, the structure should remain flat. When the initial buckle is downwards instead of upwards, the connection points are pushed up. This can also occur when stress variations through the thickness of the layer are present. Both situations will make the connection points between beams A and B curl up, making accurate measurements impossible (see Figure 5-18). To keep this to a minimum, flexible connections are placed at the end of the A-beams to keep the structure flat. Other possibilities if the technology allows, are to make the A-beams thicker. Figure 5-19 shows an example of the push-pull structure in a wafer dissolving process [5.23] with beams A much thicker than the buckling beam B.

The conversion of the tensile strain into compressive strain is very efficient using this structure. Structures need only be a little longer than simple beams buckling at the same level of compressive strain. However, due to curling upwards of the connection points and uncertainty in the
Stress measurement

Fig. 5-18 Bending of the push-pull structure making measurements impossible.

Fig. 5-19 Push-pull structure fabricated using the wafer dissolving process, with thick side beams and thin central beam.

boundary conditions, use of this structure is only useful in some processes where the side beams can be made rigid enough.
5.4 Rotating beam structure

To accurately and easily measure the stress in thin films, without the need for large arrays, a new measurement structure was designed. With this structure, which is shown in Figure 5-20, the expansion or contraction due to strain is converted into a rotation. When beams A and B shrink when the strain is compressive, both beams push on the central beam C. Due to an offset between the connection points, the beam C will rotate (as indicated by the dotted line). The rotation acts as an amplification of the expansion or contraction, providing an easily measurable displacement of the tip, which is proportional to the strain. When the strain is tensile, beams A and B will pull instead of push and beam C will rotate in the other direction. The same structure can be used to measure both tensile and compressive strain and only one structure is needed.

Fig. 5-20 Top view of the rotating beam structure which rotates left when the strain is compressive and right when the strain is tensile.

Since the development of the rotating structures, other structures using similar techniques have been designed changing the read-out [5.25], using beam curvature [5.22] and bent beams [5.23].

5.4.1 Modelling

If the rotation points are assumed to be ideal, i.e. they introduce no forces on the arms, a mathematical model of the rotation of the structure can be derived using simple geometry. Figure 5-21 shows the geometry of the
rotating structure and Figure 5-22 shows a section of the rotated structure.

![Diagram of rotating beams structure](image)

Fig. 5-21  *Geometry of the rotating beams structure.*

The angles noted in this structure are given by:

\[
\tan \varphi_1 = \frac{O}{-\varepsilon L_A + L_B + W} \tag{5-18}
\]

\[
\cos \varphi_2 = \frac{[(1+\varepsilon)^2 + 1]O^2 + (1+\varepsilon)^2W^2 - (1+\varepsilon)^2L_B^2 + (-\varepsilon L_B + L_B + W)^2}{2\sqrt{[(1+\varepsilon)^2O^2 + (1+\varepsilon)^2W^2] \cdot [O^2 + (-\varepsilon L_B + L_B + W)^2]}} \tag{5-19}
\]

\[
\tan \varphi_3 = \frac{W}{O} \tag{5-20}
\]

\[
\alpha = \varphi_1 + \varphi_2 + \varphi_3 - 90^\circ \tag{5-21}
\]

where the variables are as shown in Figure 5-21 and 5-22.

The strain can not be expressed in a simple form in terms of angle \(\alpha\). However, if the width, \(W_C \ll L_A \) and \(W_C \ll L_B \) and \(\cos \theta = 1\), which is true if beams A and B are much longer then the distance between the turning points \(O\), the strain \(\varepsilon\) can be approximated by:

\[
\varepsilon = \frac{O \tan \alpha}{L_A + L_B} \tag{5-22}
\]

and the deflection of the tip of the pointer, \(y\), is given by:

\[
y = \left(\frac{L_C}{2} + \frac{1}{2}O\right)\tan \alpha \tag{5-23}
\]
Rotating beam structure

Fig. 5-22 Close up of a schematic representation of the rotated structure.

The strain can now be expressed as a function of the displacement of the tip:

$$\varepsilon = y \cdot \frac{O}{(L_A + L_B)(L_C + \frac{1}{2}O)}$$  \hspace{1cm} (5-24)

To make the deflection of the tip as large as possible at a certain strain, distance $O$ needs to be small and lengths $L_A$, $L_B$ and $L_C$ need to be long.

Figure 5-23 shows the curves of both the complete and the simplified model for several geometries. As is clear from this graph, the error made by using the simplified model is very small in the range that the structure is used.

### 5.4.2 Finite element analysis

To check the model, the structure was simulated using the finite element program ANSYS. The width of the rotating points is varied to show the influence of this stiffness on the deflection. Figure 5-24 shows the result. The curve with a rotating point width of 0, is that of the model given by the mathematical model of Eq. (5-18) through (5-21).

From this simulation we can see that the deflection of the tip decreases as the width of the turning points increases. As this decrease is independent of the rotation angle, we can easily calculate the strain using Eq. (5-24) and then compensate for the true width of the turning points.
Stress measurement

Fig. 5-23 Error in strain measurement as a function of the rotation angle for four different rotation point widths. \( L_A = L_B = L_C = 50 \mu m \), \( O = 5 \mu m \) and \( W_A = W_B = W_C = 5 \mu m \).

Fig. 5-24 Deflection of the tip as a function of the intrinsic strain for five different rotation point widths. \( L_A = L_B = L_C = 50 \mu m \), \( O = 5 \mu m \) and \( W_A = W_B = W_C = 5 \mu m \).
using a correction factor from the simulated structure. Figure 5-25 shows the correction factor for four different rotation point width as a function of the arm width $W$. As can be seen, the correction factors is virtually independent of the arm width if the arm is not to stiff. This makes it possible to use the correction factor as a dependent of the rotation point width, not of the other geometric values.

Fig. 5-25 Correction factor as a function of the beam width for four turning point widths. $L_A=L_B=L_C=50 \mu m$ and $O=5 \mu m$.

Figure 5-26 shows three rotating beams structures in the same nitride layer with different rotation point widths. Table 5-1 shows the measurements of these three indicators and the correction factors that

Fig. 5-26 Three set of rotation structures with rotation point widths of (from left to right): 0.54 $\mu m$, 1.62 $\mu m$ and 2.70 $\mu m$. 

159
were calculated for those widths. As can be seen in the last column, after correction, the calculated strain is practically the same.

Table 5-1. Strain calculation from three rotation structures with different rotation point width $K$. Geometry: $L_A=L_B=45 \ \mu m$, $W=5 \ \mu m$ and $L_C=32.5 \ \mu m$

<table>
<thead>
<tr>
<th>$K$</th>
<th>$y$</th>
<th>$\varepsilon_{uncorrected}$</th>
<th>$C_F$</th>
<th>$\varepsilon_{corrected}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.54 $\mu m$</td>
<td>1.32 $\mu m$</td>
<td>$1.98 \times 10^{-3}$</td>
<td>0.95</td>
<td>$2.09 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.62 $\mu m$</td>
<td>0.92 $\mu m$</td>
<td>$1.38 \times 10^{-3}$</td>
<td>0.68</td>
<td>$2.03 \times 10^{-3}$</td>
</tr>
<tr>
<td>2.70 $\mu m$</td>
<td>0.70 $\mu m$</td>
<td>$1.05 \times 10^{-3}$</td>
<td>0.52</td>
<td>$2.02 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

From the structure with the 0.54 $\mu m$ wide turning point we can see, that the uncorrected strain error is only 5%. In most applications this is adequate, making it possible to use the structures without the need to correct for the turning point width. By increasing the flexibility of the turning point, for example by using a short, thin beam instead of just a narrowing or the arms, the error becomes even smaller. However, this does make the whole structure more fragile and prone to breaking.

When the indicator arm is long and the turning points are thin, the weight of the arm can cause the arm to drag along the substrate. This can cause measurement errors, if the stress changes after release and drying. This can easily be avoided by making the indicator symmetrical, with arms of equal length on either side of the turning points as was done with the structures shown in Figure 5-26.

5.4.3 Measurement

To determine the strain level in the layer, the deflection of the tip must be determined. This can be done using an optical microscope or a scanning electron microscope (SEM). As the deflection of the tip is small (In the order of several tenths of a micron to several micron, depending on the design and stress level), special measurement arrangements are possible to simplify the measurement.

This can be as simple as a pointed tip (see Figure 5-27) for structures that are large and rotate over large distances. Due to overetching and photolithographic effects, the points become rounded which makes accurate measurement difficult.
A better configuration consists of two side references (see Figure 5-28). By measuring the distance between the tip and both references, the difference can be calculated, giving the deflection. No extensions are needed and any changes in width during processing are compensated.

By placing two similar structures tip to tip, as shown in Figure 5-26, the displacement is doubled.
Stress measurement

To make fairly accurate measurement possible through a simple optical microscope, Vernier structures can be added to the tip of the structure as shown in Figure 5-29.

![Fig. 5-29 Symmetrical rotation structure with Vernier structures at the tips to simplify read-out.](image)

When thick turning points are used (larger than 1 μm) the width of the turning point must be measured, so the correction factor can be determined. This can be done using a SEM.

The rotation structure is easy to use and simple to make. The strain in any layer that can be patterned and freed from its substrate, can be measured using this structure. Figure 5-30 shows the rotation structure using the wafer dissolving process mentioned in section 4.2.5.

The thickness of the layer does not have any influence on the deflection and buckling due to compressive strain is unlikely due to the ease with which the structure can release the strain by rotating. This makes large, accurate structures possible. Only one structure is necessary for the measurement of both tensile and compressive strain, so no large die area is needed, and the structure can be incorporated on all micromachined chips to monitor the strain in all free-standing layers.

Sticking that may occur during release of the structure, as mentioned in section 4.5.16.1, does not present any problems. The
structure will rotate as soon as it is freed from its surrounding. When sticking occurs, the structure will stick in the rotated position, with measurements still possible.

A problem may arise when the strain varies through the thickness of the layer. This can cause the tip of the indicator beam to curl up, making measurements extremely difficult, or even impossible.

The structures have been applied to many measurements. Strains have been measured in polysilicon, silicon nitride and monocrystalline silicon. The deflection of the tip is determined by the average contraction or expansion of the arms. This averaging effect can be used to determine the strain in materials that can not be shaped in a rotating structure on their own [5,24]. An example is the strain in silicon oxide, which would be removed during the sacrificial etch of the wafer. By sandwiching the oxide between other materials, e.g. polycrystalline silicon, a strain measurement of the average of the strain in the oxide and the polysilicon could be determined (see Figure 5-31). By also measuring the strain in a rotating structure made entirely of polysilicon, the strain in the oxide can be calculated.

In a similar way, the strain distribution through the thickness of the layer can be determined. The displacement of the whole structure is
Stress measurement

Fig. 5-31 Top view and cross section of a rotation structure constructed of two materials to determine the strain in a layer that can not be formed into a freestanding structure.

measured, then part of the layer is removed by plasma etching and the displacement is measured again. This is repeated several times and the strain can be calculated as a function of depth.

5.5 Conclusions

Many different stress and strain measurement techniques are available. However, several have the disadvantage of requiring entire wafers, either because the test is destructive or unpatterned layers are needed, or complex measurement set-up. To be able to use the measurement technique during fabrication to monitor or check the stress levels in several mechanical layers, easy optical determination of the stress is needed. This can only be accomplished using either buckling or deformation structures.

Buckling structures are convenient only for the determination of compressive strain. The conversion structures are large and inaccurate. Determining which beams in an array buckle, is not easy due to the very small buckling height and to get enough measurement range and resolution, the arrays need to be large.

The deformation structures suggested in literature only measure compressive strain and are difficult to read, due to very small displacements.

Better results have been accomplished using the rotating beam structure. The structure is small, easy to use and accurate enough for
most micromachining purposes (±10%). It measures both compressive and tensile strain and only mask defined lengths are needed for the calculation. No thickness or width information is needed, so deviations due to processing do not influence the accuracy.

5.6 References


References


Stress measurement
6.1 Actuator design

Now that the basic shape of the actuator has been established and a theoretical model of the actuator is available, a specific geometry can be determined. As is clear from the theoretical model presented in chapter 3, the performance of the actuator depends on many parameters. Some of these parameters are determined by the surface micromachining process used to fabricate the devices. This process is described in chapter 4 and specifies the thicknesses of all the layers used, including the gap size. This leaves us with the length of the lever, $L_1$, and the length of the electrode, $L_2$ to determine the step size and the width and hinge configuration to determine the minimum activation voltage.

The actuator used is designed to have the largest possible step size with the largest possible load. The angle of rotation is determined by the gap size of 0.53 μm (380 nm PSG 2 and 150 nm polysilicon 2 as the rotation point of the hinge can be considered to be halfway through the thickness of the layer) and electrode length $L_2$. To get the largest possible step, the electrode is kept as small as possible, which is 6 μm. This is the smallest distance possible while ensuring that all steps do overlap when taking overetching and misalignment into account. The maximum angle of rotation $\varphi$, is: 0.1 (6.1 degrees).
Realization and measurement

The vertical offset between the hinge and the point touching the objects is 1.13 μm (380 nm PSG 1, 300 nm polysilicon 1, 300 nm silicon nitride and 150 nm of polysilicon 2), which is the maximum that can be achieved. The arm length $L_1$ is chosen to be 6 μm. Although this is very short, it ensures that the load the actuator is able to lift is considerable, despite the small drive electrode. This provides an initial angle θ, of 0.22 (12.7 degrees). By substituting these parameters in Eq. (3-37) we find a step size $\Delta x = 186$ nm for a single lift. In one cycle, the object will be transported over 245 nm.

The initial electrostatic couple is determined by the parameters given above and the width, $w$, of the plate. Together with the couple caused by the spring action of the torsion hinge, this determines the activation voltage. As the actuators are to be driven using the BiFET process DIMES-01, the activation voltage must be below 20 V. The width of the plate is 40 μm. The plates are sufficiently stiff to make them long, as this saves chip area due to fewer hinges and lead wires per unit area. The hinge is chosen to be 10 μm long and is masked at 2 μm wide. It consists of only the polysilicon layer, not the silicon nitride layer in order to keep it as flexible as possible. Underetching of the polysilicon causes the width of the hinge beams to be in the order of 1.5 μm. This gives a theoretical activation voltage of 21 Volts (the voltage at which the electrostatic couple is larger than the hinge couple for all angles of rotation.). This should be low enough to enable normal microelectronic devices to switch the drive signals even under load.

6.2 Fabrication of the actuator

The micromachining process described in chapter 4 was used for the fabrication of the actuator described in chapter 3. The process sequence of the device is shown here to illustrate the use of all the layers when applied to this particular structure. The positioning device itself, only makes use of the micromachining module and can be constructed on a wafer without any electronics. However, the circuits suggested in section 3.8 can be incorporated without changing the basic layout of the actuator. In many implementations the diffusions are used as under-crossings to connect the electrodes into a matrix. The structure shown here uses such an under-crossing to connect the grounded moving plate/top electrode to
an interconnect line. The structure is shown in both top view and two cross-sections.

The wafer is protected by the first silicon nitride layer, after which the first sacrificial layer is deposited. This sacrificial layer is used to form the bumps underneath the electrostatic plate and the contact point to increase the initial angle. Patterning of the sacrificial layer takes place using the PSG1 mask.

Fig. 6-1  *Nitride deposition and deposition and patterning of the first sacrificial PSG layer.*
Realization and measurement

Opening of the contact holes to the substrate is done using the NT1 mask, to enable a contact between the polysilicon and the under-crossings.

Fig. 6-2  Opening of contact hole to the substrate.

The first polycrystalline layer is deposited, doped and then patterned with the PS1 mask, to form the lower electrode on one side and to lift the contact point to the same level on the other side. This also serves as an electrostatic shield to avoid any electrostatic forces on the other side of the electrodes.

Fig. 6-3  Deposition and patterning of first polysilicon layer.
The second sacrificial layer is deposited and patterned with the PSG2 mask to form the anchor points for the supports of the torsion hinges.

Fig. 6-4 *Deposition and patterning of the second sacrificial layer.*

The mechanical silicon nitride layer is deposited to form an insulation between the two electrodes and to stiffen both the moving plate and the support of the torsion hinges. Due to the first polysilicon underneath the supports, the nitride can end halfway in the anchor points. The pattern is defined using the NT2 mask.

Fig. 6-5 *Deposition and patterning of the mechanical nitride layer.*
The second polycrystalline silicon layer is deposited and patterned with the PS2 mask, to form the upper electrode and the rest of the moving plate, the torsion hinge and the electrical contact over the support onto the first polysilicon layer.

Fig. 6-6  deposition and patterning of the second polysilicon layer.

The PSG is partially stripped and the TEOS layer is deposited. This is patterned using the OX1 mask, to provide contact on the first polysilicon, anchor the interconnect layer and clear the oxide around contacts to the substrate.

Fig. 6-7  Stripping of PSG and deposition and patterning of TEOS protection layer.
Contact holes are opened to allow excess to the under-pass. The CO mask is used in this step, which is combined with the standard electronic processing.

Fig. 6-8 *Etching of the contact holes.*

The interconnect is sputtered, patterned and the TEOS stripped from the surface using the IC mask. This mask is also part of the standard electronic process.

Fig. 6-9 *Sputtering and patterning of the interconnect.*
The surface is dried and a resist layer spun. This is patterned using the SAC mask to protect the aluminium interconnect.

**Fig. 6-10 Resist protection mask.**

The wafer is etched in pad-etchant, rinsed in wafer and the resist mask is removed without drying the wafer. The water is replaced by cyclohexane, which is frozen and sublimated in a nitrogen flow.

**Fig. 6-11 Top and cross-sectional views of the finished actuator.**
This leaves the structures free-standing and ready for use. Two SEMs of the resulting actuators as described in the previous section are shown in Figure 6-12.

Fig. 6-12  *SEM giving a side and top view of the completed actuator.*
6.3 Measurements

Several one and two-dimensional arrays where fabricated to test both the processing and the mechanical and electrostatic model of the devices. In addition, the electrical characteristics of the materials and transistors where tested.

6.3.1 Testing of the actuators

The arrays where tested by applying a square wave as shown in Figure 3-3 to a set of actuators with the same orientation. The frequency of the signal was 3 Hz and the amplitude of the signal was varied.

The actuators that where tested showed considerable variation in the actuation voltage. Actuation voltages of actuators fabricated to the specifications mentioned in Figure 6.1, varied between 5 V and 15 V. At 5 V one or more actuators rotated. At the actuation voltage all functioning actuators rotated. Increasing the voltage up to 30 V did not increase the number of actuators rotating. Approximately 80% of the actuators in the array worked. The voltage was not increased above 30 V during testing of the arrays, as breakdown of the cross-overs between different signal lines was observed. The actuation voltage of the array, i.e. the voltage at which all functioning actuators rotated, varied between arrays on the same wafer. The lowest was measured at 7 V, the highest at 15 V.

The measured voltages are considerably lower than those predicted by the theoretical model. These differences can be caused by many small variations in the fabrication process, such as changes in the thickness of the second PSG layer, which forms the gap, the thickness of the second polysilicon layer of which the hinge is constructed, the Young’s modulus of the polysilicon and extra underetching of the polysilicon, resulting in a slimmer and longer hinge beam.

During operation, a number of actuators stopeed working. After turning off the signal for a short while, these actuators did function again. This was thought to be caused by charging of the isolation layer of the actuator. This theory was further supported by the observation that the actuators started working again when exposed to humid air.

Testing of the transport function was not possible, as all arrays that where tested exhibited considerable cross-talk between signal lines.
When actuators of one orientation where activated, many actuators with different orientations where activated as well. However, tests with an object placed on the array did show that the actuators are strong enough to lift small items and could function as a positioning device. The object placed on the array was a grain of sugar, which has flat crystal surfaces.

### 6.3.2 Electrical measurements

The measurements of the sheet resistance of the polysilicon layer show values between 150 $\Omega/\square$ and 200 $\Omega/\square$. This is in good agreement with the desired value of 200 $\Omega/\square$.

To test the compatibility of the electronics with the surface micromachining, test circuits where placed on the same chip and subjected to the full surface micromachining process. Figure 6-13 shows the Gummel plot of an NPN transistor of the DIMES01 BiFET process. The plot shows the characteristics of a standard transistor and one which has been subjected to the surface micromaching. The curves show that the extra processing during the surface micromachining, has no significant effect on the characteristics of the NPN transistor.

![Gummel plot of a DIMES01 NPN transistor before and after the surface micromachining process.](image)

**Fig. 6-13** Gummel plot of a DIMES01 NPN transistor before and after the surface micromachining process.
Realization and measurement
Conclusions

A positioning device has been fabricated and integrated with a standard bipolar microelectronic process (DIMES-01) using a compatible surface micromachining module.

The positioning device is constructed as a distributed system to provide the necessary force and motion needed to manipulate macroscopic objects. It consists of an array of actuators with limited motion and power, which are divided into two sets that are driven alternately. This provides extra functionality, which can be increased further by dividing the array in several parts, making reference lines and points, rotations and relative motion between two objects possible. The addition of electronic circuits can increase the possibilities even further and is only possible in this kind of compatible micromachining.

The individual actuators are driven electrostatically and can be designed to provide high power or high movement speeds. These actuators have dimensions in the order of $20 \times 20$ \(\mu\text{m}\) and the arrays that make up the system consist of several thousand actuators per square millimetre. Maximum movement speed is limited by the object being moved. The objects that can be moved must be vary flat and smooth or special support surfaces need to be used.
Conclusions

A surface micromachining process has been developed that provides two free-standing polysilicon layers, separated by a silicon nitride layer that can also be used as a mechanical layer. The layers are all under low tensile strain to avoid buckling of large structures and reduce the chance of sticking. The process is cleanroom and material compatible with standard processing and is thermally compatible with BiFET microelectronic process DIMES-01. The highest temperature needed in the processing is 850°C. Electronic components on the same chip as the surface micromachining still operates within the standard specifications. However, not all devices available in the DIMES-01 process can be used in combination with the surface micromachining module. These include those transistors that use a washed emitter. This emitter can not be protected from the micromachining and these transistor’s characteristics will change. To change the micromachining process to accommodate these transistors, more testing and research is necessary.

For the development of the low-stress mechanical layers in the micromachining process, on-chip stress measurement structures have been designed. New buckling structures have been designed for the measurement of tensile strain and a rotating arm structure was designed to measure both compressive and tensile strain. Of these the rotating structure functions are the best. Although it is not very accurate (inaccuracy is approximately 10%), it is small and easy to used and has been used successfully in the development of the low stress polysilicon process. It is also placed on all chips that include surface micromachining, to monitor the strain in all mechanical layers.

Structures have been fabricated and tested and have actuation voltages varying between 7 V and 15 V. These actuators are strong enough to lift small objects. Cross-talk between signal lines makes transportation impossible. The electronics present in the wafer are not influenced by the surface micromachining process making the incorporation of adressing circuits possible.
Appendix A
Mechanical calculations

The shape of the diamond shaped structure causes the strain in the beams to vary, depending on the residual strain in the film. To calculate the strain in the beams, Figure A-1 shows the shape and dimensions used.

Fig. A-1 Diamond shaped strain conversion structure for the measurement of tensile and compressive strain.
Appendix Mechanical calculations

To model the structure, a simplified wire model is used. In effect, the beams are considered to have no width and the connection points are considered to rotate freely.

![Diagram of forces acting on the connection point of beams 3 and 4.]

Fig. A-2  Forces acting on the connection point of beams 3 and 4.

The strain in beams 3 and 4 is calculated by first determining the displacement of the connection point between beam 3 and 4 due to the residual strain in beams 3 and 4. The forces pulling on the connection point and the resulting displacement are shown in Figure A-2. The force equilibrium gives:

\[ F_4 = F_3 \Rightarrow \left( \varepsilon_0 - 2 \frac{\delta}{L_4} \right) A_4 = \left( \varepsilon_0 - \frac{\delta}{L_3} \right) A_3, \quad (A-1) \]

where \( A_3 \) and \( A_4 \) are the cross-sectional area of beams 3 and 4 respectively.

From this the displacement can be calculated:

\[ \delta_{3,4} = \frac{(A_4 - A_3)L_3 L_4}{A_3 L_4 + 2A_4 L_3} \quad (A-2) \]

The strain in beams 3 and 4 can be expressed as a combination of the residual strain and the strain resulting from the displacement:

\[ \varepsilon'_3 = \varepsilon_0 + \frac{\delta_{3,4}}{L_3} = \varepsilon_0 \left( 1 + \frac{(A_4 - A_3)L_4}{A_3 L_4 + 2A_4 L_3} \right) \quad (A-3) \]

\[ \varepsilon'_4 = \varepsilon_0 - \frac{2\delta_{3,4}}{L_4} = \varepsilon_0 \left( 1 - \frac{2(A_4 - A_3)L_3}{A_3 L_4 + 2A_4 L_3} \right) \]

Figure A-3 shows the model of the structure. Only one half is modelled as the structure is symmetrical along beams 3 and 4. If the end of beam 1
Fig. A-3  Wire frame model of one half of the the diamond shaped stress measurement structure.

is considered to be not connected and free to move, the displacement due to the strain is comprised of the change in length of beam 1 itself due to the strain $\varepsilon_0$, and the displacement of the connection point to beam 2. This displacement is very small and can be expressed as:

$$\Delta \delta_{1,2} = \left. \frac{d}{d\varepsilon} \delta(\varepsilon) \right|_{\varepsilon = 0},$$

(A-4)

where the function $\delta(\varepsilon)$ is given by:

$$\delta^2(\varepsilon) = (L_2(1-\varepsilon))^2 - \left( \frac{L_4}{2} - \delta_{3,4}(\varepsilon) \right)^2$$

(A-5)

$L_2$ can be expressed in terms of $L_4$ as: $L_2 = \frac{1}{2}\sqrt{2}L_4$. Combining Eq. (A-2) with Eq. (A-4) and Eq. (A-5) gives an expression for the displacement:

$$\delta_{1,2} = -\varepsilon_0 \frac{(A_4 L_3 + A_3 L_4 + A_3 L_3)L_4}{A_3 L_4 + 2A_4 L_3}$$

(A-6)

The displacement of the end of beam 1 can be expressed as a combination of the change in length of beam 1 and the displacement of the connection point.

$$\delta_{e1} = -\varepsilon_0 L_1 + \Delta y$$

(A-7)
The end point of beam 1 is anchored to the substrate which results in a reaction force $F_1$, which results in a change of length of beam 1 of:

$$
\Delta L_1 = \frac{F_1 L_1}{EA_1},
$$

where $E$ is the Young’s modulus of the film and $A_1$ is the cross-sectional area of beam 1. This result in a force on the end of beam 2.

$$
F_2 = \frac{F_1}{2 \cos \frac{\pi}{4}} = \frac{1}{2} \sqrt{2} F_1,
$$

which in turn results in a change in length of beam 2:

$$
\Delta L_2 = \frac{F_2 L_2}{EA_2},
$$

where $A_2$ is the cross-sectional area of beam 2.

This force then results in a force on the connection point between beams 3 and 4 along beams 3 and 4 of:

$$
F_{3,4} = \frac{1}{2} \sqrt{2} F_2 = \frac{1}{2} F_1,
$$

which in turn must comply with the displacement of the connection point, $\delta_{3,4}$.

$$
F_{3,4} = F_3 + F_4 \Rightarrow F_{3,4} = \delta_{3,4} E \left( \frac{A_3}{2L_3} + \frac{A_4}{L_4} \right)
$$

Here only half the cross-sectional area of the beams 3 and 4 are used, as only half the structure is modelled. The displacement $\delta_{3,4}$, which is the same as the change in length of both beams 3 and 4, can be expressed as:

$$
\delta_{3,4} = \Delta L_3 = \Delta L_4 = \frac{2L_3 L_4}{F_{3,4} E (2A_4 L_3 + A_3 L_4)}
$$
From these changes in length, the strain energies in the beams can be calculated:

\[ U_{\alpha 1} = \frac{F_1 \Delta L_1}{2} = \frac{F_1^2 L_1}{2EA_1} \]  \hspace{1cm} (A-14)

\[ U_{\alpha 2} = \frac{F_2 \Delta L_2}{2} = \frac{F_2^2 L_2}{4EA_2} \]  \hspace{1cm} (A-15)

\[ U_{\alpha(3,4)} = \frac{F_{3,4} \delta_{3,4}}{2} = \frac{F_{3,4}^2 L_3 L_4}{4E(2A_4 L_3 + A_3 L_4)} \]  \hspace{1cm} (A-16)

The force \( F_1 \), will result in a total displacement \( \delta_1 \), of the end of the beam 1. The work \( W \) of the reaction force \( F_1 \), is equal to the strain energy in the beams:

\[ W = \frac{F_1 \delta_{F1}}{2} = U_{\alpha 1} + 2U_{\alpha 2} + 2U_{\alpha 3,4} \]  \hspace{1cm} (A-17)

By substituting the strain energies given by Eq. (A-14), Eq. (A-15) and Eq. (A-16), the displacement can be calculated as a function of force \( F_1 \):

\[ \delta_{F1} = \frac{F_1}{E} \left( \frac{L_1}{A_1} + \frac{L_2}{A_2} + \frac{L_3 L_4}{2A_4 L_3 + A_3 L_4} \right) \]  \hspace{1cm} (A-18)

Because the end of beam 1 is anchored on the substrate, it will not move and the displacement due to the reaction force \( F_1 \) will be the same as the displacement due to the residual strain \( \varepsilon_0 \):

\[ \delta_{\varepsilon 1} + \delta_{F1} = 0 \]  \hspace{1cm} (A-19)

From this an expression for the reaction force \( F_1 \) can be determined.

\[ F_1 = \varepsilon_0 E \frac{L_1(A_4 L_3 + A_3 L_4 + A_3 L_3) L_4}{L_1 L_2 L_3 L_4} \frac{A_4 L_4 + 2A_4 L_3}{A_1 A_2 + 2A_4 L_3 + A_3 L_4} \]  \hspace{1cm} (A-20)
Appendix Mechanical calculations

The resulting strain in beam 3 and (half of) beam 4 is the sum of the strain calculated as a result of the residual strain and resulting displacement (see Eq. (A-3)) and the strain resulting from the displacement caused by force $F_{3,4}$:

$$
\varepsilon_3 = \varepsilon_3' + \frac{\delta_{3,4}}{L_3}
$$
(A-21)

$$
\varepsilon_4 = \varepsilon_4' - 2\frac{\delta_{3,4}}{L_4}
$$

By combining the equations, Eq. (A-3), Eq. (A-11), Eq. (A-13) and Eq. (A-20), the strain in beams 3 and 4 can be expressed as a function of the residual strain.

$$
\varepsilon_3 = \varepsilon_0 \left(1 + \frac{(A_4 - A_3)L_4}{N} + \frac{L_4A_1A_2}{N} \left( \frac{L_1N + L_4(L_3A_3 + L_3A_4 + L_4A_3)}{N\left(L_1A_2 + \frac{1}{2}\sqrt{2L_4A_1}\right) + L_3L_4A_1A_2} \right) \right)
$$

(1-22)

$$
\varepsilon_4 = \varepsilon_0 \left(1 - \frac{2(A_4 - A_3)L_3}{N} \left( \frac{L_4A_1A_2}{N} \left( \frac{L_1N + L_4(L_3A_3 + L_3A_4 + L_4A_3)}{N\left(L_1A_2 + \frac{1}{2}\sqrt{2L_4A_1}\right) + L_3L_4A_1A_2} \right) \right) \right)
$$

(1-23)

where $N = 2A_4L_3 + A_3L_4$. 

188
Summary

Design and fabrication of a surface micromachined positioning device.

Chapter 1

The development of microelectronics has initialised the development of the construction of mechanical devices using micromachining techniques. Many sensors have been designed using this technology, but actuators have only been used to demonstrate the possibilities of the technology, not for any particular purpose.

The objective has been to design a positioning device using surface micromachining. The systems consists of many simple actuators arranged to cooperate so their overall strength and range is many times that of the individual actuators. Such systems are called distributed systems.

Chapter 2

Many different kinds of actuators and actuation principles are feasible when using micromachining. Electrical power is the most useful as it can also be used to carry a signal, and is consistent with the IC technology used in its manufacture. Electrical power can be converted into motion in many different ways, using many intermediate energy forms. These include; thermal by electrical heating, using the expansion of materials, mechanical using piezoelectric materials or the increase of pressure by heating, magnetic using coils to move magnetic materials and electrostatic using the attracting forces between capacitor plates. Of these methods the electrostatic method is the most useful as it can generate large forces without the need for exotic materials, is efficient, fast and is compatible with silicon micromachining.

Two electrostatic drive principles that can be used are variable capacitance and charge relaxation. With variable capacitance actuation, the attracting and aligning forces between parallel capacitor plates are used when an electrical field is present between the plates. Charge
relaxation makes use of the force on induced charges in an electrical field. When the field changes, the charges will be displaced in comparison to the potential of the field, which will result in a force on the conducting material. This is a dynamic effect and can only be used in some actuators.

Individual actuators are small and produce little power and their motion is limited. By combining many actuators in parallel and series into so called, distributed systems their range and power can be combined.

Designing a microactuator using micromachining in silicon is different from large scale machines and design considerations such as; electrostatic efficiency, friction and wear, binding effects, mechanical tolerances and the mechanical properties of silicon and silicon related materials must be taken into account.

Chapter 3

The positioning device consists of an array of small electrostatically actuated levers, which alternately lift and lower the object to be moved, passing it from actuator to actuator. The distance moved in one cycle is determined by the length of the lever and the angle over which it rotates. The actuator consists of a plate, supported along two sides above an electrode. When a voltage is applied between the electrode in the base and the plate, the plate will rotate around the support, the other side acting as the lever needed to move the object. To optimize this distance, the drive efficiency and the load distribution, both the electrode side of the plate and the lever side are bent upwards.

To support the plate and at the same time enable it to rotate, three configurations present themselves. The bending of cantilevers, the deflection of bridges and the twisting of beams. Mechanical analysis shows that the twisted beams allow a rigid support and the lowest resistance to the rotation.

Electrostatic analysis shows that the fringe field can be neglected when considering the electrostatic forces on the plate electrode. The actuation forces will not only cause the plate to rotate, the force will also cause the support to bend, reducing the gap and the distance over which the object is moved. To keep the electrodes from shorting out when pull-in occurs at a certain voltage, the electrodes are separated by a non conducting layer, or by separate landing electrodes. To avoid sticking,
when the electrodes touch, the actuators must be driven by an alternating positive and negative voltage, to avoid charge built up.

The objects that can be moved using this system must be flat and smooth or must be placed on a movement tray. The arrays of actuators can be constructed to perform linear or circular movement. By dividing the arrays in one or more parts, reference lines and points can be created and rotations are possible. By introducing electronics on the same chip, many functions can be created, making the system easy to use and flexible.

Chapter 4

Silicon is a good mechanical material and many micromachining technologies, such as bulk micromachining, surface micromachining, wafer-to-wafer bonding, SOI and epi micromachining are available. The structures can be used to make sensors, actuators or for signal processing, using mechanical laws of physics. The process used for the fabrication of the actuators was designed for general use and consists of a two layer surface micromachining process, compatible with the BiFET DIMES-01 microelectronic process.

The electronics are first fabricated and the wafer is taken out of the standard line before the opening of the contact holes and the deposition of the interconnect. The wafer is covered by a protective, low stress silicon nitride film, on which the first sacrificial PSG is deposited. This is implanted with Ar to create a bevelled edge during patterning, which reduces the formation of stringers. A low stress polysilicon layer is deposited and in-situ annealed. This is followed by the second sacrificial PSG layer, which also serves to protect the underlying layers. Then a low stress LPCVD nitride layer is deposited, followed by the second low stress polysilicon layer. Patterning of the polysilicon and nitride layers is done using specially designed plasma etches to avoid the formation of stringers along steps. The sacrificial layer is removed from the surface, after which a TEOS layer is deposited to protect the polysilicon layers during the patterning of the interconnect. Contact holes are opened in the TEOS layer and through the underlaying nitride layer to the electronics. The aluminium interconnect is patterned using a combination of dry and wet etching. All the areas with aluminium are covered by resist, as this would be attacked by the sacrificial etchant. Problems such as poor adhesion of the resist and peeling of the aluminium have been seen. To overcome these problems, thin plasma oxide layers have been deposited,
Summary

resist has been hard baked and reflowed. To increase the selectivity of the sacrificial etchant BHF with respect to Al, glycerol and glycerine additions have been tried, and so called pad etchant has been used. Best results have been achieved using hard baked resist in combination with the pad etchant. To avoid sticking, the structures are rinsed in cyclo-hexane, which is frozen and sublimated under a nitrogen flow.

Chapter 5

To determine the intrinsic stress level in thin layer many different methods are available. These include resonance frequency measurement of a beam, load deflection of a beam, wafer and cantilever curvature, buckling of bridges and deformation. Buckling is widely used for surface micromachining as it is easy to use. At a certain strain, all beams of a certain length and longer will buckle. To determine the strain an array is needed of beams of increasing length. Buckling only occurs when the strain is compressive, so conversion structures are needed to measure tensile strain by converting the strain in part of the structure from tensile to compressive.

To overcome the disadvantages of the existing methods, new structures where designed. Two new strain conversion structures where designed using the buckling methods. The diamond shaped structure is more reliable than the existing ring structure, but is very inefficient and therefore large. The push-pull structure is very efficient and small, but tend to curl upwards making them useless. A rotating beams structure amplifies the deformation due to strain using a rotating pointer resulting in a displacement of the tip which is directly proportional with the strain. This structure is easy to use and measures both tensile and compressive strain, without the need for a large array. It has been used successfully in the development of a low stress polysilicon process.

Chapter 6

The actuators have been fabricated, which have plates measuring 10 μm long and 40 μm wide. Hinges are 10 μm long and 1.5 μm wide. This results in an actuation voltage of 21 V and a movement of the object in one cycle of 245 nm.

One and two-dimensional arrays have been fabricated and tested. Activation voltages where between 7 V and 15 V. Cross talk between signal lines makes testing of the transport function impossible. The
transistors have been tested and show no change in their characteristics after the surface micromachining.

Chapter 7

The positioning device has been fabricated in a two layer surface micromachining process that is compatible with a standard BiFET microelectronic process (DIMES-01). New stress measurement structures have been designed, which are sufficiently accurate and small enough to be incorporated on all surface micromachined chips to monitor the strain.
Summary
Samenvatting

Ontwerp en fabricage van een surface-micromachined positionerings systeem.

Hoofdstuk 1

De ontwikkelingen in de microelektronica hebben geleid tot constructietechnieken voor de vervaardiging van mechanische structuren op basis van micromachining. Veel sensoren zijn al ontworpen die gebruik maken van deze technieken. Actuatoren die van dezelfde technologie gebruikmaken worden uitsluitend gebruikt om de mogelijkheden van deze technologie te demonstreren. Niet voor praktische toepassingen.

Het doel van het onderzoek was dan ook om een positionerings-systeem te ontwerpen dat gebruik maakt van surface-micromachining. Het systeem is opgebouwd uit vele simpele actuatoren die zodanig samenwerken dat de resulterende kracht en bereik van het gehele vele malen groter is dan dat van de losse actuatoren. Zulke systemen worden gedistribueerde systemen genoemd.

Hoofdstuk 2

Een grote verscheidenheid van actuatieprincipes en actuatoren gafbriceerd met behulp van micromachining is denkbaar. Elektrische energie is het meest praktische, omdat het tevens kan worden gebruikt als drager van informatie en in overeenstemming is met de gebruikte IC-technologie. Elektrische energie kan op vele wijzen worden omgezet in beweging, eventueel door gebruik te maken van andere verschillende vormen van energie: Thermisch door elektrische verhitting, waarbij gebruik gemaakt wordt van de uitzetting van materialen, mechanisch met behulp van piezo-elektrische materialen of gas-druk door verhitting, magnetisch met behulp van spoelen warme magnetische materialen kunnen worden aangetrokken of afgestoten en elektrostatisch met behulp van de aantrekkende kracht tussen twee platen van een capaciteit. Van deze methodes is de elektrostatische de meest geschikte, aangezien deze grote krachten kan opwekken zonder dat
daarvoor bijzondere materialen nodig zijn. Het is efficient, snel en
compatibel met micromachining in silicium.

Twee verschillende electrostatische aandrift principes kunnen
worden onderscheiden, namelijk op basis van variabele capaciteit en
ladingsrelaxatie. Bij variabele capaciteit actuatie worden de
aantrekkelijke krachten gebruikt tussen twee capaciteitplaten, waartussen
een elektrisch veld aanwezig is. Ladingsrelaxatie maakt gebruik van de
krachten op geïnduceerde ladingen in een elektrisch veld. Als het veld
verandert, zullen de ladingen verplaatst zijn ten opzichte van de
potentiaal, wat resulteert in een kracht op het geleidende materiaal. Dit is
een dynamisch effect en kan alleen in een beperkt aantal actuatore
worden toegepast. Voor de gebruikte actuator is het variable capaciteit
principe gebruikt.

De individuele actuatore zijn klein en produceren maar weinig
vermogen en een zeer beperkte vrije slag. Door vele actuatore parallel
en in serie te zetten, in een zogenaamd gedistribueerd systeem, wordt hun
vermogen en bereik gecombineerd.

Het ontwerpen van een microactuator gemaakt met behulp van
micromachining in silicium, verschilt van het ontwerpen van
macroskopische machines. Er dient rekening te worden gehouden met
electrostatische efficiëntie, wrijving en slijtage, vastloop effecten,
mechanische toleranties en de mechanische eigenschappen van silicium
een gerelateerde materialen.

**Hoofdstuk 3**

Het positioneringssysteem bestaat uit een matrix van elektrostatisch
aangedreven hefboomjjes, die afwisselend het object dat verplaatst
moet worden, optillen en laten zakken en het daarbij doorgeven van
actuator naar actuator. De afstand waarover het object in één cyclus
wordt verplaatst, wordt bepaald door de lengte van de hefboom en de
hoek waarover deze draait. De actuator bestaat uit een plaat die
halverwege is opgehangen en aan één kant boven een elektrode hangt.
Als een spanning wordt aangebracht tussen de elektrode en de plaat, zal
die plaat roteren rond de ophanging. De andere kant van de plaat fungeert
als de hefboom die het object verplaatst. Om deze afstand, de aandrift
efficiëntie en de belasting te optimaliseren, is zowel de elektrode als de
hefboom naar boven gebogen.

Om de plaat te ondersteunen en tegelijk te kunnen laten rotieren,
zijn drie ophangingen mogelijk. Het buigen van een balkje, het verbuigen
van een bruggetje of het torderen van een balkje. Mechanische analyse laat zien dat de torderende balkjes een stevige ophanging mogelijk maken en een lage rotatieweerstand hebben.

Een elektrostatische analyse laat zien dat de randvelden kunnen worden verwaarloosd bij de bepaling van de elektrostatische krachten op de elektrode. De aantrekkende krachten zullen de plaat niet alleen laten draaien, maar ook enige doorbuiging van de ophanging tot gevolg hebben. Hierdoor wordt de afstand tussen de platen kleiner en daarmee ook de afstand waarover het object wordt verplaatst. Om te voorkomen dat er kortsluiting ontstaat tussen de elektrode en de plaat als de plaat is verdraaid, is een niet-geleidend tussenlaag aangebracht of een losse contact elektrode. Om kleven (sticking) te voorkomen als de elektrodes elkaar raken, moeten de actuatoren worden aangestuurd met een afwisselend positieve en negatieve spanning, om ladingsopbouw tegen te gaan.

De objecten die met dit systeem kunnen worden verplaatst, dienen vlak en glad zijn of op een speciaal plateau worden geplaatst. De matrix van actuatoren kan zo worden samengesteld dat rechtlijnig of cirkelvormige bewegingen mogelijk zijn. Door de matrix op te delen in los te besturen gedeelten, ontstaan referentielijnen en punten waarom rotaties mogelijk zijn. Door de elektronica op dezelfde chip te plaatsen kunnen andere functies worden uitgevoerd, wat het systeem veelzijdig maakt en eenvoudig in het gebruik.

**Hoofdstuk 4**

Silicium is een materiaal met goede mechanische eigenschappen en veel micromachining technieken, zoals bulk micromachining, surface micromachining, wafer-to-wafer bonding, SOI en epitaxiale micromachining, zijn beschikbaar. De structuren kunnen gebruikt worden voor het maken van sensoren, actuatoren en signaalverwerkings-systemen die gebruik maken van natuurwetten in het mechanische domein. Het proces dat gebruikt is voor de fabricage van de actuatoren is ontworpen voor algemeen gebruik. Het bestaat uit een surface micromachining proces met twee mechanische lagen dat compatibel is met het BiFET DIMES-01 microelektronica proces.

De elektronische schakelingen wordt eerst gerealiseerd, waarbij de plak uit de standaard productielijn wordt gehaald, alvorens de contactgaten worden geopend en de bedrading gedepongeerd. De plak wordt dan bedekt met een beschermende siliciumnitridelaag, waarna de
eerste sacrificiaal PSG-laag wordt gedeponeerd. Deze wordt geïmplanteed met argon, zodat er tijdens het aanbrengen van het patroon een afgeschuinde rand ontstaat, waardoor de vorming van stringers wordt gereduceerd. Een polysiliciuimlaag met een lage mechanische spanning wordt gedeponeerd en in-situ geannealed. Hierna wordt de tweede sacrificiaal PSG-laag aangebracht, die tevens dient om de onderliggende lagen te beschermen. Dan wordt een LPCVD siliciumnitride met lage spanning gedeponeerd, gevolgd door de tweede polysiliciuimlaag. Het patroon wordt op de polysiliciuim- en nitridelagen overgebracht met behulp van een speciaal ontworpen plasma-ets procédé, om de vorming van stringers langs de stappen tegen te gaan. De sacrificiaal-laag wordt van het oppervlak verwijderd, waarna een TEOS-laag wordt gedeponeerd om de polysiliuciumlagen te beschermen tijdens het aanbrengen van de bedrading. De contactgaten worden geopend in de TEOS-laag, het onderliggende nitride en oxide tot op de elektronica. De aluminium bedrading wordt gedeponeerd en met een combinatie van droog en nat etsen wordt het patroon aangebracht. Alle gebieden waar aluminium bedrading aanwezig is worden bedekt met fotolak, om deze te beschermen tegen het sacrificiaal etsmiddel. Hierbij treden problemen op zoals slechte hechting van de lak en losscheuren van het aluminium. Om deze problemen op te lossen is geëxperimenteerd met dunne oxidelagen en het bakken en laten vloeien van de fotolak. Om de selectiviteit van het sacrificiaal etsmiddel BHF ten opzichte van Al te vergroten, zijn glycerine en glycerol toevoegingen geprobeerd, evenals een zogenaamde pad-etchant. De beste resultaten zijn behaald met uitgebakken fotolak en de pad-etchant. Om sticking te voorkomen worden de structuren gespoeld in cyclo-hexaan, dat bevroren wordt en daarna sublimeert in een stikstof stroom.

Hoofdstuk 5

Om de intrinsieke mechanische spanning in een dunne film te bepalen zijn er verschillende methodes beschikbaar, zoals; bepaling van de resonantie frequentie van een bruggetje, doorbuiging van een bruggetje, meting van de kromming van een plak of balkje, opbollen (buckling) van bruggetjes en vervorming. Het opbollen is een veel gebruikte methode voor surface-micromachining en is eenvoudig in het gebruik. Bij een bepaalde spanning, zullen alle bruggetjes die langer dan een bepaalde lengte, krom gaan staan. Om de spanning te bepalen is een rij van bruggetjes nodig van verschillende lengte. Dit opbollen zal alleen
optreden als een drukspanning aanwezig is. Voor de bepaling van trekspanning zijn conversie structuren nodig die de trekspanning lokaal omzetten in een drukspanning.

Om de nadelen van de bestaande methodes te vermijden zijn een aantal nieuwe structuren ontworpen. Twee nieuwe structuren voor de conversie van de spanning maken gebruik van opbollen. De ruitvormige structuur is beter betrouwbaar dan de bestaande ring-structuur, maar is inefficiënt en daardoor erg groot. De trek-duw structuur is efficiënt en klein maar heeft de neiging op te krullen, waardoor deze onbruikbaar wordt. Een vervormende rotatie-structuur versterkt de vervorming ten gevolge van de spanning met behulp van een roterende arm. De uitwijking van het uiteinde van deze arm is een directe maat voor de spanning. Deze structuur maakt het mogelijk op eenvoudige wijze zowel druk- als trekspanning te meten, zonder een grote rij structuren. Het is met succes gebruikt bij de ontwikkeling van het proces voor polysilicium met lage spanningen.

Hoofdstuk 6

De actuatoren die zijn gemaakt hebben platen met afmetingen van 10 μm lang en 40 μm breed. De ophanging is 10 μm lang en 1,5 μm breed. Dit heeft een activerings spanning van 21 V tot gevolg en een verplaatsing van 245 nm, in één cyclus.

Eén en twee dimensionale arrays zijn gemaakt en getest. Activeringsspanningen lagen tussen de 7 V en 15 V. Door signaal overdracht tussen verschillende signaallijnen, was het onmogelijk de transport functie te testen. De transistoren zijn getest en de karakteristieken zijn niet veranderd door het surface-micromachining process.

Hoofdstuk 7

Het positioneringssysteem is gemaakt in een twee-laags surface-micromachining proces dat compatibel is met een standaard microelektronica proces (DIMES-01). Nieuwe structuren voor het meten van mechanische spanningen zijn ontworpen, die nauwkeurig en klein genoeg zijn om gebruikt te kunnen worden op alle surface-micromaching chips om de mechanische spanningen in de gaten te houden.
Samenvatting
Acknowledgements

In the years that it took to complete the work presented in this thesis, many people contributed, both professionally and personally. First of all I would like to thank all members of the Electronic Instrumentation Laboratory for the friendly atmosphere. Working there has been a pleasure. I would like to thank all the people of both the Electronic Instrumentation lab. and DIMES, for their help and the use of their facilities.

Although it is impossible to mention all people that contributed, I would like to thank the following people in particular:

Professor Dr. Ir. S. Middelhoek, my promotor, for providing me with the opportunity to perform this research.

STW and FOM for providing the necessary financial support

Dr. Ir. R.F. Wolffensuttel, for his advice and the informal working relationship.

P.J. French for his help in solving many technological problems and time spent processing.

My roommates, Bert van Drieënhuizen, Daniel Poenar and Yuang Xiong Li, for the relaxed atmosphere and the many conversations, of which at least a few were scientific.

Bert for his invaluable help with the development of the surface micromachining process, solving the many problems encountered and modelling of the stress measurement structures. Daniel for the characterization of the PSG sacrificial layers and Yuang Xiong for the development of the plasma etches and the many hours of processing.

Charles de Boer who did much of the processing and whose attention to detail resulted in the first working devices.
Acknowledgements

P.M. Sarro for her advice and letting our funny processing into the cleanroom.

Professor K.D. Wise of the University of Michigan, Ann Arbor, MI, U.S.A. and his staff for the opportunity to gain some experience in IC-processing and Huibert-Jan Verhoeven and Rob van Kampen for making our time in Ann Arbor and the trip afterwards an enjoyable one.

Finally I would like to thank my family for their support and encouragement and my friends for their support in difficult times, patience and the many relaxing hours when not on the job.
List of publications and presentations

Related to this thesis:


List of publications and presentations

Transducers '93, 7-10 June, Yokohama, Japan, 1993, pp. 783-786.


About the Author

Johannes F. L. Goosen was born in Rotterdam, The Netherlands on December 11, 1966. After finishing his pre-university education (VWO) at the OSG Walburg in Zwijndrecht, he started his study in Electrical Engineering at the Delft University of Technology, Delft in 1985. In 1991 he received the Master's degree, on his work on electrostatic bearings for micromotors using permanent charges in non-conducting layers. In October 1991 he joined the Electronic Instrumentation Laboratory as a PhD. student. The research concerned the design of a positioning device using a distributed system of electrostatically driven, surface micromachined actuators. He was also closely involved in the development of a two layer, surface micromachining process compatible with standard microelectronic processing, for the fabrication of surface micromachined sensors and actuators. At the 1994 Eurosensors conference he received the 'best oral presentation' award. During his PhD. research he spent three months at the University of Michigan, Ann Arbor, MI, U.S.A., as a visiting scholar, to work in their IC laboratory.